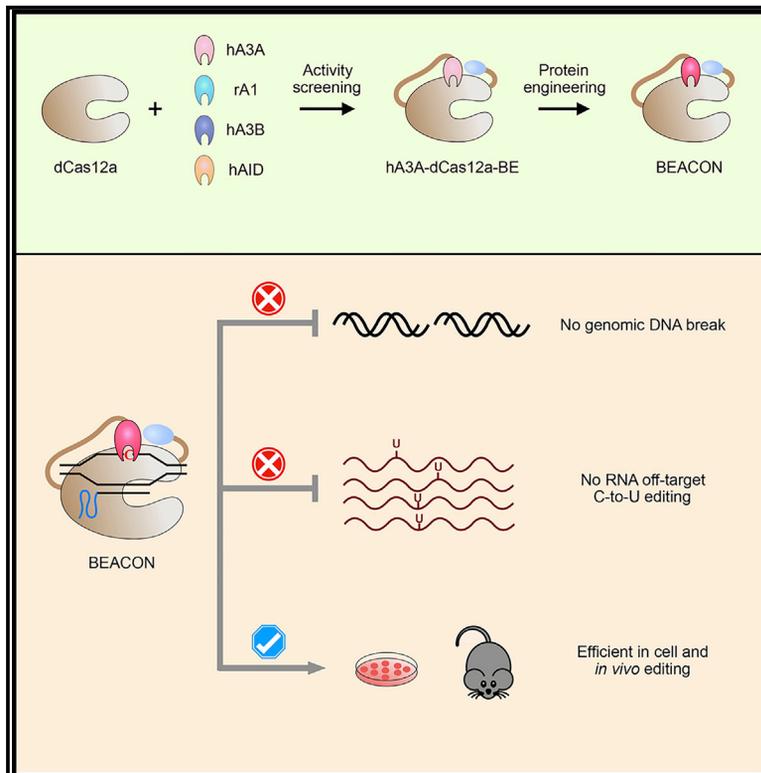


Cas12a Base Editors Induce Efficient and Specific Editing with Low DNA Damage Response

Graphical Abstract



Authors

Xiao Wang, Chengfeng Ding,
Wenxia Yu, ..., Zhen Liu, Li Yang, Jia Chen

Correspondence

zliu2010@ion.ac.cn (Z.L.),
liyang@picb.ac.cn (L.Y.),
chenjia@shanghaitech.edu.cn (J.C.)

In Brief

Wang et al. develop a BEACON base-editing system by combining dCas12a with human APOBEC3A and its engineered versions. BEACON induces efficient editing in cells and mouse embryos with basal levels of DNA damage response, RNA off-target mutations, and unintended side products.

Highlights

- BEACON induces basal levels of DNA breaks and DNA damage response
- BEACON induces a basal level of RNA off-target mutations
- BEACON induces *in vivo* base editing with high product purity



Article

Cas12a Base Editors Induce Efficient and Specific Editing with Low DNA Damage Response

Xiao Wang,^{1,2,3,7} Chengfeng Ding,^{1,2,3,7} Wenxia Yu,^{1,2,3,7} Ying Wang,^{4,7} Siting He,^{3,5,7} Bei Yang,^{6,7} Yi-Chun Xiong,⁴ Jia Wei,⁴ Jifang Li,^{1,2,3} Jiayi Liang,^{1,2,3} Zongyang Lu,^{1,2,3} Wei Zhu,¹ Jing Wu,¹ Zhi Zhou,¹ Xingxu Huang,^{1,2} Zhen Liu,^{5,*} Li Yang,^{4,*} and Jia Chen^{1,2,8,*}

¹School of Life Science and Technology, ShanghaiTech University, Shanghai 201210, China

²Shanghai Institute of Biochemistry and Cell Biology, CAS Center for Excellence in Molecular Cell Science, Chinese Academy of Sciences, Shanghai 200031, China

³University of Chinese Academy of Sciences, Beijing 100049, China

⁴CAS Key Laboratory of Computational Biology, CAS-MPG Partner Institute for Computational Biology, Shanghai Institute of Nutrition and Health, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Shanghai 200031, China

⁵Institute of Neuroscience, CAS Key Laboratory of Primate Neurobiology, CAS Center for Excellence in Brain Science and Intelligence Technology, Chinese Academy of Sciences, Shanghai 200031, China

⁶Shanghai Institute for Advanced Immunochemical Studies, ShanghaiTech University, Shanghai 201210, China

⁷These authors contributed equally

⁸Lead Contact

*Correspondence: zliu2010@ion.ac.cn (Z.L.), liyong@picb.ac.cn (L.Y.), chenjia@shanghaitech.edu.cn (J.C.)

<https://doi.org/10.1016/j.celrep.2020.107723>

SUMMARY

The advent of base editors (BEs) holds great potential for correcting pathogenic-related point mutations to treat relevant diseases. However, Cas9 nickase (nCas9)-derived BEs lead to DNA double-strand breaks, which can trigger unwanted DNA damage response (DDR). Here, we show that the original version of catalytically dead Cas12a (dCas12a)-conjugated BEs induce a basal level of DNA breaks and minimally activate DDR proteins, including H2AX, ATM, ATR, and p53. By fusing dCas12a with engineered human apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 3A (APOBEC3A), we further develop the BEACON (base editing induced by human APOBEC3A and Cas12a without DNA break) system to achieve enhanced deamination efficiency and editing specificity. Efficient C-to-T editing is achieved by BEACON in mammalian cells at levels comparable to AncBE4max, with only low levels of DDR and minimal RNA off-target mutations. Importantly, BEACON induces *in vivo* base editing in mouse embryos, and targeted C-to-T conversions are detected in F0 mice.

INTRODUCTION

The CRISPR-Cas9 system has been successfully applied in various living organisms for genome editing (Hsu et al., 2014; Knott and Doudna, 2018; Komor et al., 2017). The Cas9 nuclease generates DNA double-strand breaks (DSBs) at specific genomic loci under the direction of guide RNAs (gRNAs) (Cong et al., 2013; Jinek et al., 2013; Mali et al., 2013), which are further processed by downstream DNA repair pathways to induce gene editing outcomes. Therefore, the formation of DSBs is required for CRISPR-Cas9-mediated genome editing. Although Cas9 nickase (nCas9) only generates DNA single-strand breaks (SSBs), apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like (APOBEC)/activation-induced cytidine deaminase (AID) family members (Harris and Liddament, 2004) together with base excision repair proteins (Chen et al., 2014) can lead to DSBs during the repair of these SSBs (Lei et al., 2018). In addition, base editors (BEs) that link nCas9 with APOBEC1 are also

associated with DSBs, demonstrated by the induced insertions or deletions (indels) of nucleotides around target sites (Komor et al., 2016; Lei et al., 2018; Nishida et al., 2016).

In most cases, DSBs are repaired through the non-homologous end joining (NHEJ) pathway (Ceccaldi et al., 2016) to yield indels (Chakrabarti et al., 2019; van Overbeek et al., 2016), which lead to knockout of protein-coding genes. Alternatively, DSBs can be resolved through homology-directed repair (HDR) in the presence of a donor DNA (Ceccaldi et al., 2016), which can be used to induce precise sequence replacement. However, DSBs are highly toxic lesions (Chapman et al., 2012) and can trigger downstream DNA damage response (DDR) signaling pathways to disturb cellular homeostasis (Roos et al., 2016; Zhou and Elledge, 2000), e.g., cell proliferation (Haapaniemi et al., 2018; Ihry et al., 2018). Specifically, the generation of DSBs can trigger the autophosphorylation of the protein kinase ataxia-telangiectasia mutated (ATM) (Shiloh and Ziv, 2013). Meanwhile, during the repair of DSB, an end resection process



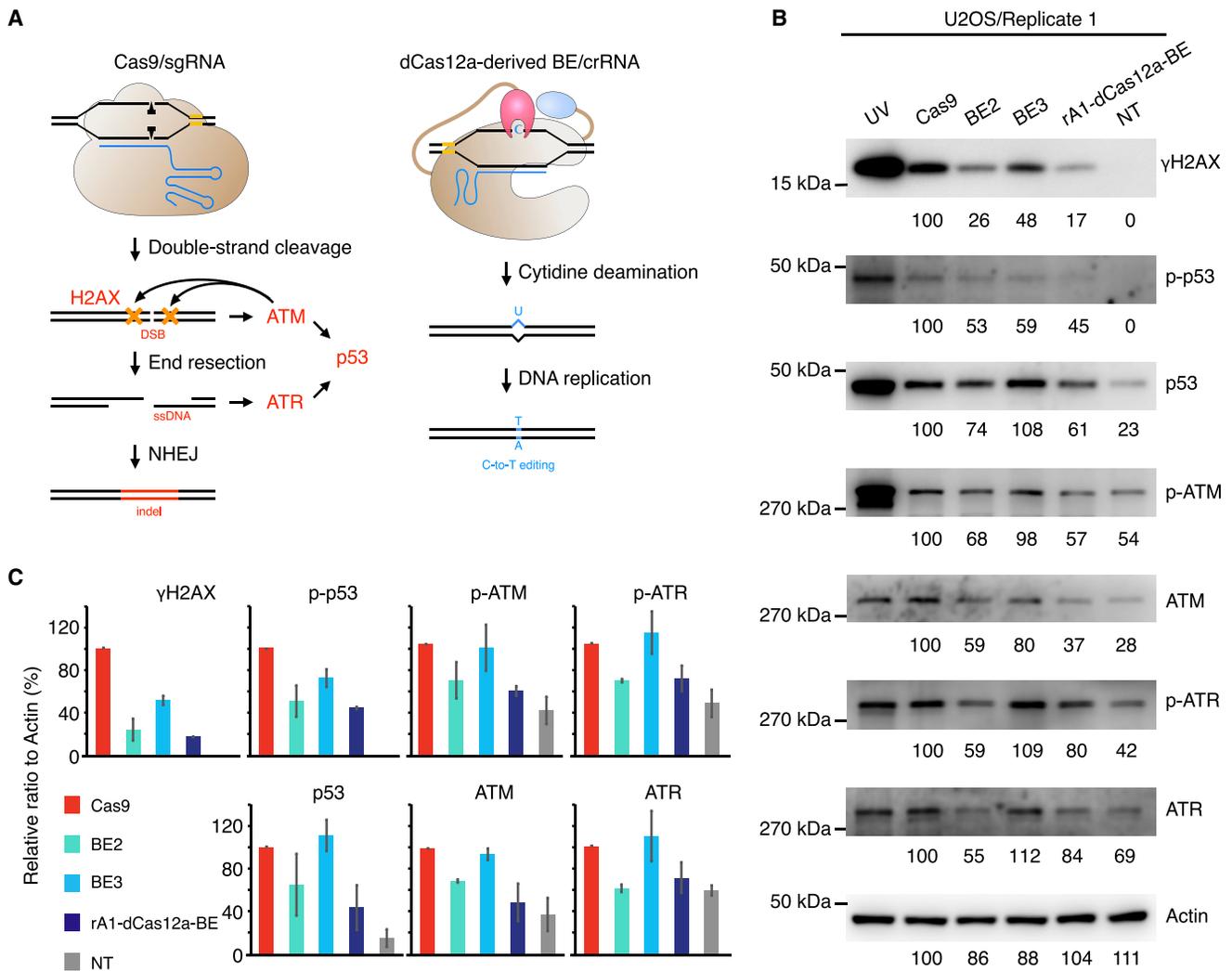


Figure 1. dCas12a-Derived BEs Triggered DDR Minimally

(A) Schematic diagrams illustrate the gene editing mediated by Cas9 (left), the base editing mediated by dCas12a-derived BEs (right), and the relationship with DNA damage response. DNA double-strand breaks generated by Cas9 nuclease can trigger the phosphorylation and activation of a series of proteins involved in DNA damage response signaling pathways, such as ATM, ATR, H2AX, and p53.

(B) Immunoblots of DDR proteins and their phosphorylated forms triggered by Cas9-mediated gene editing or BE2-, BE3-, and rA1-dCas12a-BE-mediated base editing. The numbers represent individual protein contents relative to those triggered by Cas9 (setting as 100).

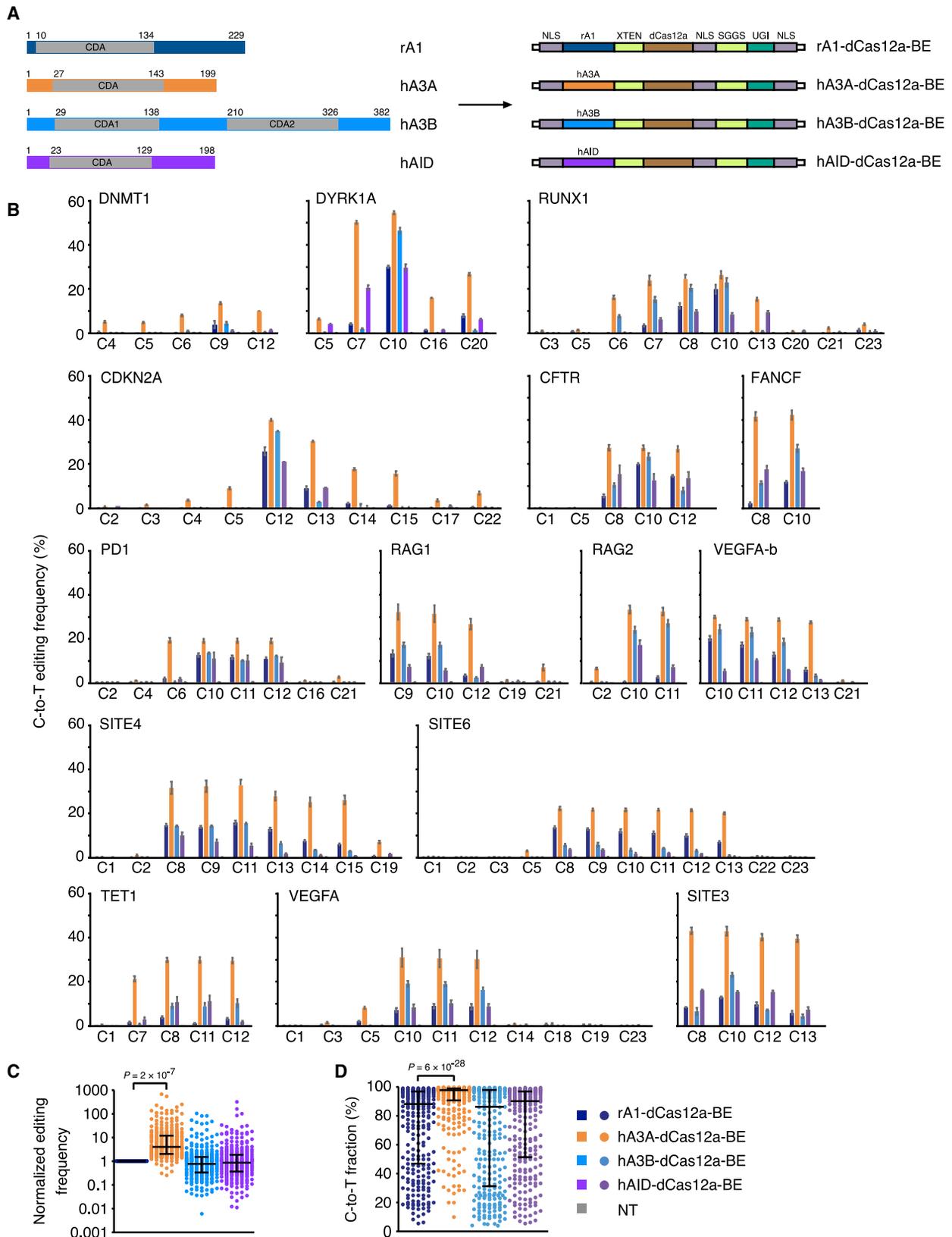
(C) Quantification of the relative protein contents normalized with actin in (B) and Figure S1A. Means \pm SD are from two independent experiments.

See also Figure S1.

generates single-stranded DNA (ssDNA) regions (Ceccaldi et al., 2016) and further activates another protein kinase, ATM and RAD3-related (ATR) (Cimprich and Cortez, 2008). ATM, and to some extent ATR, can phosphorylate the histone variant H2AX and tumor suppressor protein p53, which signal DSB repair (Mah et al., 2010) and then regulate the cell cycle (Biegging et al., 2014).

Cas12a (as known as Cpf1) is another CRISPR-Cas protein that is distinct from Cas9 in many aspects (Zetsche et al., 2015). Cas12a recognizes a T-rich PAM sequence, requires a short gRNA (CRISPR RNA, crRNA) and has been reported to have a generally higher targeting specificity than Cas9 (Kim et al., 2016; Kleinstiver et al., 2016, 2019; Yan et al., 2017). These

characteristics render Cas12a a promising gene editing platform. We have recently developed catalytically dead Cas12a (dCas12a)-derived BEs that were conjugated with rat APOBEC1 (rA1) (Li et al., 2018). In theory, dCas12a-derived BEs are unlikely to cause DSBs; therefore, unwanted DDR could be largely avoided. Nevertheless, dCas12a-BEs also induced much lower editing efficiencies than nCas9-BEs did (Gehrke et al., 2018; Huang et al., 2019; Koblan et al., 2018; Li et al., 2018; Thuronyi et al., 2019; Wang et al., 2019), and *in vivo* base editing by dCas12a-BEs has not been achieved in animals. In this study, we confirmed that the early version of dCas12a-derived BEs induced very low levels of DDR. Furthermore, by screening and engineering different APOBEC/AID family members, we



(legend on next page)

identified that highly efficient and specific C-to-T editing could be induced by human APOBEC3A (hA3A)-dCas12a-BEs, referred to as BEACON (base editing induced by human A3A and Cas12a without DNA break). With no toxic DSB and basal levels of DDR and RNA off-target (OT) effects, BEACON also induced *in vivo* base editing successfully in mouse embryos, and targeted C-to-T conversions were detected in F0 mice.

RESULTS

Low Levels of DDR Induced by dCas12a-BE

Different from Cas9-mediated genome editing, dCas12a-derived BEs only deaminate one or a few cytosines in the ssDNA of the R-loop formed by gRNA at target sites, which will be converted to thymines after DNA replication. Given the fact that dCas12a is used as the Cas moiety, dCas12a-derived BEs theoretically are unlikely to generate DNA breaks or activate DDR cascades (Figure 1A). In contrast, nCas9-derived BEs triggered indel formation as the nCas9-generated nick would be converted to a DSB (Lei et al., 2018), which activates DDR proteins. Indeed, much higher levels of phosphorylated H2AX (γ H2AX) were observed in cells treated with Cas9 or a nCas9-derived BE3 than those in cells treated with rA1-dCas12a-BE (Figures 1B, 1C, and S1A). Correspondingly, the levels of phosphorylated p53 (p-p53), ATM (p-ATM), and ATR (p-ATR) in rA1-dCas12a-BE-treated cells were also lower than those in the cells treated with Cas9 or BE3 (Figures 1B, 1C, and S1A). As shown in Figure S1B, rA1-dCas12a-BE induced almost no indel at the target site, whereas high levels of indels were observed with the Cas9 or BE3 treatment. In addition, the catalytically dead Cas9 (dCas9)-derived BE2 also triggered low levels of DDR (Figures 1B, 1C, and S1A) and indels (Figure S1B), but it only induced base editing with limited efficiencies (Figure S1C). These results demonstrated that rA1-dCas12a-BE induced base editing with the least damage on genomic DNA.

Screening of dCas12a-Derived BE

We next compared the editing efficiencies of rA1-dCas12a-BE and the commonly used BE3 at disease-associated target sites (Wang et al., 2019; Figure S2). Although rA1-dCas12a-BE induced purer C-to-T editing products and fewer indels than BE3, it induced ~4-fold lower C-to-T editing frequencies (median, $p = 2 \times 10^{-11}$; Figure S2E) than BE3 (~13% to 48%; Figure S2B) at seven selected disease-associated genomic loci. The fact that currently available rA1-dCas12a-BEs induce significantly low editing efficiency impedes its broad applications. We sought to develop new dCas12a-derived BEs with high C-to-T editing efficiencies by linking dCas12a with different APOBEC/AID family members. Among three newly constructed BEs

(hA3A-dCas12a-BE, hA3B-dCas12a-BE, and hAID-dCas12a-BE) and one previously reported rA1-dCas12a-BE (Figure 2A), hA3A-dCas12a-BE induced the highest editing efficiency, which is ~4-fold higher than the original rA1-dCas12a-BE (median, $p = 2 \times 10^{-7}$; Figure 2C), across 15 genomic loci (Figure 2B). Further analysis showed that hA3A-dCas12a-BE induced the purest C-to-T editing as well (Figure 2D).

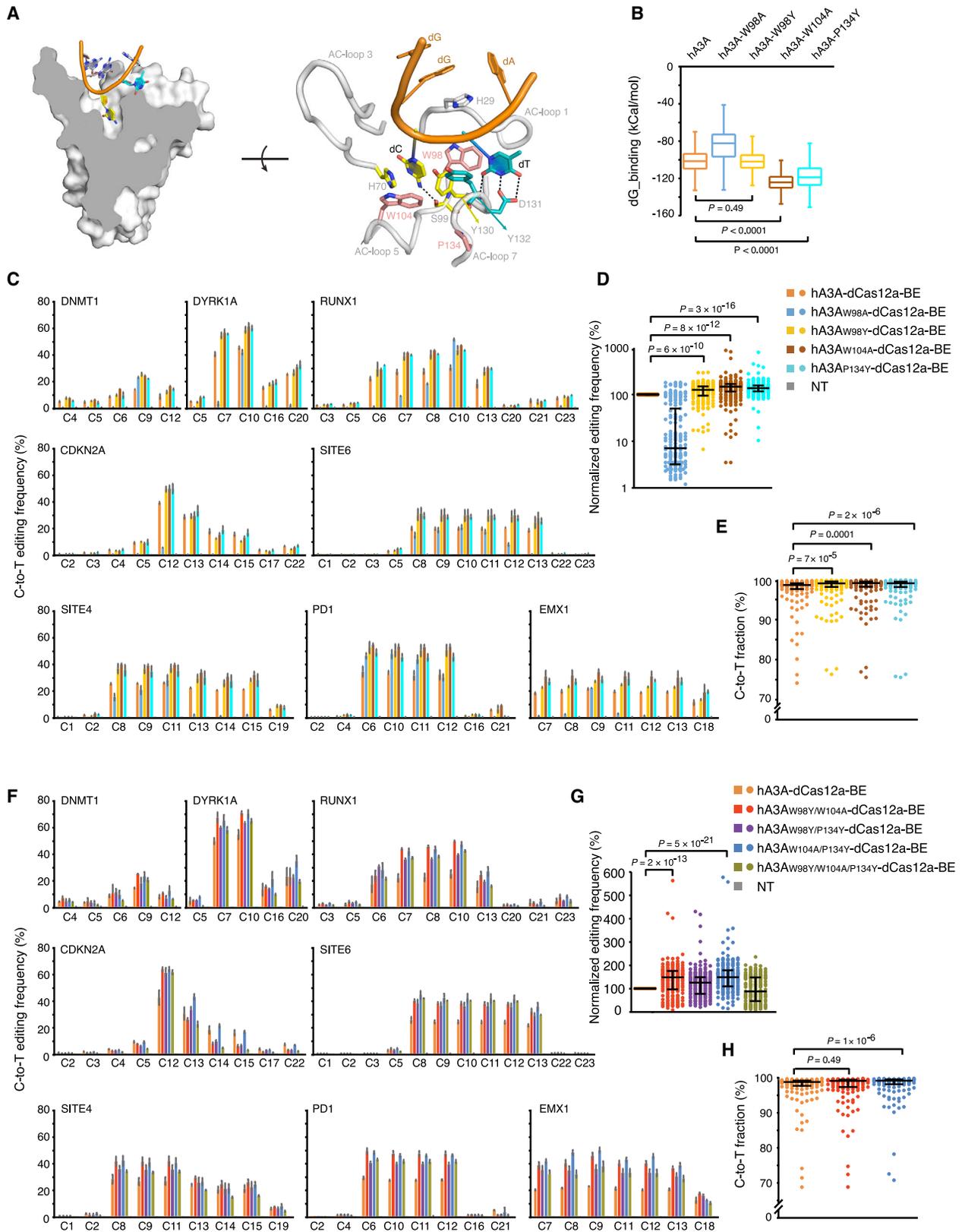
Improvement of hA3A-dCas12a-BE

We then aimed to further improve the editing efficiency of hA3A-dCas12a-BE by engineering its hA3A moiety. According to the structure of hA3A in complex with a 5'-TC-containing ssDNA (Shi et al., 2017), the targeted C was usually accommodated in a groove that is formed by hA3A residues from active center loops (AC-loops) 1, 3, 5, and 7 (Figure 3A). Of the residues on these loops, His29, His70, Tyr130, Ser99, Asp131, and Tyr132 are directly involved in substrate coordination (Figure 3A, close up view in right panel). Thus, we opted to avoid mutating these residues and instead to choose aromatic AC-loop residues that are located on the peripheral of the ssDNA binding groove, i.e., Trp98 and Trp104, for subsequent engineering. Meanwhile, Pro134 was also selected for engineering, as proline is generally considered a secondary structure disruptor and the change of Pro134 may have effects on the flexibility of AC-loop 7. Molecular simulation of the corresponding hA3A mutants in complex with ssDNA ligands (Figure 3B) indicated that mutations of W104A and P134Y significantly strengthened the interaction between hA3A and ssDNA, whereas W98A largely destabilized the interaction (Figure 3B). Although the W98Y mutation alone seemed to not affect the binding of the ssDNA ligand (Figure 3B), it might loosen the nucleotide preference at -1 positions (Shi et al., 2017). Thus, we introduced the amino acid changes of W98Y, W104A, and P134Y into the hA3A moiety of hA3A-dCas12a-BE, with the mutation W98A serving as the negative control (Figure 3C). In total, four engineered BEs (hA3A_{W98A}-dCas12a-BE, hA3A_{W98Y}-dCas12a-BE, hA3A_{W104A}-dCas12a-BE, and hA3A_{P134Y}-dCas12a-BE) were obtained for subsequent comparison.

Among them, three engineered BEs (hA3A_{W98Y}-dCas12a-BE, hA3A_{W104A}-dCas12a-BE, and hA3A_{P134Y}-dCas12a-BE) induced higher editing frequencies than the wild-type hA3A-dCas12a-BE (median, 1.32-, 1.54-, and 1.40-fold; and $p = 6 \times 10^{-10}$, 8×10^{-12} , and 3×10^{-16} , respectively), whereas hA3A_{W98A}-dCas12a-BE showed decreased editing efficiency (Figure 3D). As a control, two mutations at amino acids (C101S and C106S) close to the enzymatic active site eliminated editing efficacy (Figure S3A). Next, we further combined the amino acid changes of W98Y, W104A, and P134Y to test whether these alternations can increase the editing efficiency in a synergistic manner (Figure 3F). The dual-change combinations of W98Y/W104A and

Figure 2. Screening of APOBEC/AID-dCas12a Combinations for Efficient Base Editing

(A) Characteristics of the APOBEC/AID deaminases (left) and the schematic diagram illustrating the relevant dCas12a-BEs (right).
 (B) Comparison of C-to-T editing frequencies induced by indicated APOBEC/AID-dCas12a-BEs at different genomic target sites. Means \pm SD are from three independent experiments. NT, non-transfected.
 (C) Statistical analysis of normalized C-to-T editing frequencies shown in (B), setting the ones induced by rA1-dCas12a-BE as 100%.
 (D) Statistical analysis of C-to-T product purity yielded by indicated BEs shown in (B).
 n = 294 samples from three independent experiments in (C) and (D). p value, one-tailed Student's t test. The median and interquartile range (IQR) are shown. See also Figure S2.



(legend on next page)

W104A/P134Y induced higher editing frequencies than W98Y/P134Y (Figure 3G). Interestingly, the triple-change combination of W98Y/W104A/P134Y induced lower editing frequencies than the original hA3A-dCas12a-BE (Figure 3G), suggesting that the triple amino acid changes may compromise the cytidine-deaminase activity of hA3A. Among all engineered BEs, three of them (hA3A_{W104A}-dCas12a-BE, hA3A_{W98Y/W104A}-dCas12a-BE, and hA3A_{W104A/P134Y}-dCas12a-BE) induced high levels of editing frequencies, i.e., 1.54-, 1.57-, and 1.64-fold of editing efficiencies relative to hA3A-dCas12a-BE (median, $p = 8 \times 10^{-12}$, 2×10^{-13} , and 5×10^{-21} , respectively; Figures 3D and 3G), and were used for further engineering.

Next, we also optimized the codons of these four BEs (Figure 4A) for mammalian expression to enhance their editing efficiencies (Koblan et al., 2018). As expected, editing efficiencies of these codon-optimized BEs (hA3A-dCas12a-BE-op, hA3A_{W104A}-dCas12a-BE-op, hA3A_{W98Y/W104A}-dCas12a-BE-op, and hA3A_{W104A/P134Y}-dCas12a-BE-op) were all significantly improved (Figure 4B). Among them, the engineered and codon-optimized ones (hA3A_{W104A}-dCas12a-BE-op, hA3A_{W98Y/W104A}-dCas12a-BE-op, and hA3A_{W104A/P134Y}-dCas12a-BE-op) induced higher levels of editing frequencies (i.e., 1.55-, 1.71-, and 1.76-fold) than hA3A-dCas12a-BE-op with optimized codons only (Figure 4B, the far right panel). Of note, the engineering of hA3A or codon optimization did not substantially affect the product purity induced by corresponding BEs (Figures 3E, 3H, and 4C).

Editing windows of hA3A_{W104A}-dCas12a-BE-op, hA3A_{W98Y/W104A}-dCas12a-BE-op, and hA3A_{W104A/P134Y}-dCas12a-BE-op were shown to be ~15 bp long (Figures 4D and S3B, positions 6–20, setting the PAM-proximal nucleotide as position 1). Although large editing windows are useful for intended mutagenesis, e.g., creating stop codons to knock out genes (Billon et al., 2017; Kuscu et al., 2017), they were otherwise too wide for precise editing, as multiple cytosines might exist in the editing window (Rees and Liu, 2018). In order to narrow editing windows, we further introduced Y130F and/or Y132D, which were previously shown to narrow the editing windows of hA3A-conjugated BEs (Wang et al., 2018) into the engineered hA3As. Either Y130F or Y132D narrowed the editing windows while maintaining (or slightly affecting) editing efficiencies (Figures 4D, 4E, 4F, and S3B). Interestingly, the introduction of Y130F or Y132D further reduced the formation of indels (Figures S3C–S3F) with an unknown mechanism. However, despite an even further narrowed editing window, simultaneous introduction of Y130F and Y132D led to a reduced editing efficiency (Figures 4F and S3B). Thus,

we chose hA3A-dCas12a-BEs containing either Y130F or Y132D for subsequent analysis and designated them as BEACON (base editing induced by human A3A and Cas12a without DNA break). Hereafter, hA3A_{W104A/Y132D}-dCas12a-BE-op and hA3A_{W98Y/W104A/Y130F}-dCas12a-BE-op are referred to as BEACON1 and BEACON2, respectively. Both BEACONS exhibit narrowed editing windows and induce low levels of indels (Figures 4D, S3D, and S3E).

Comparison of BEACON and Other BEs

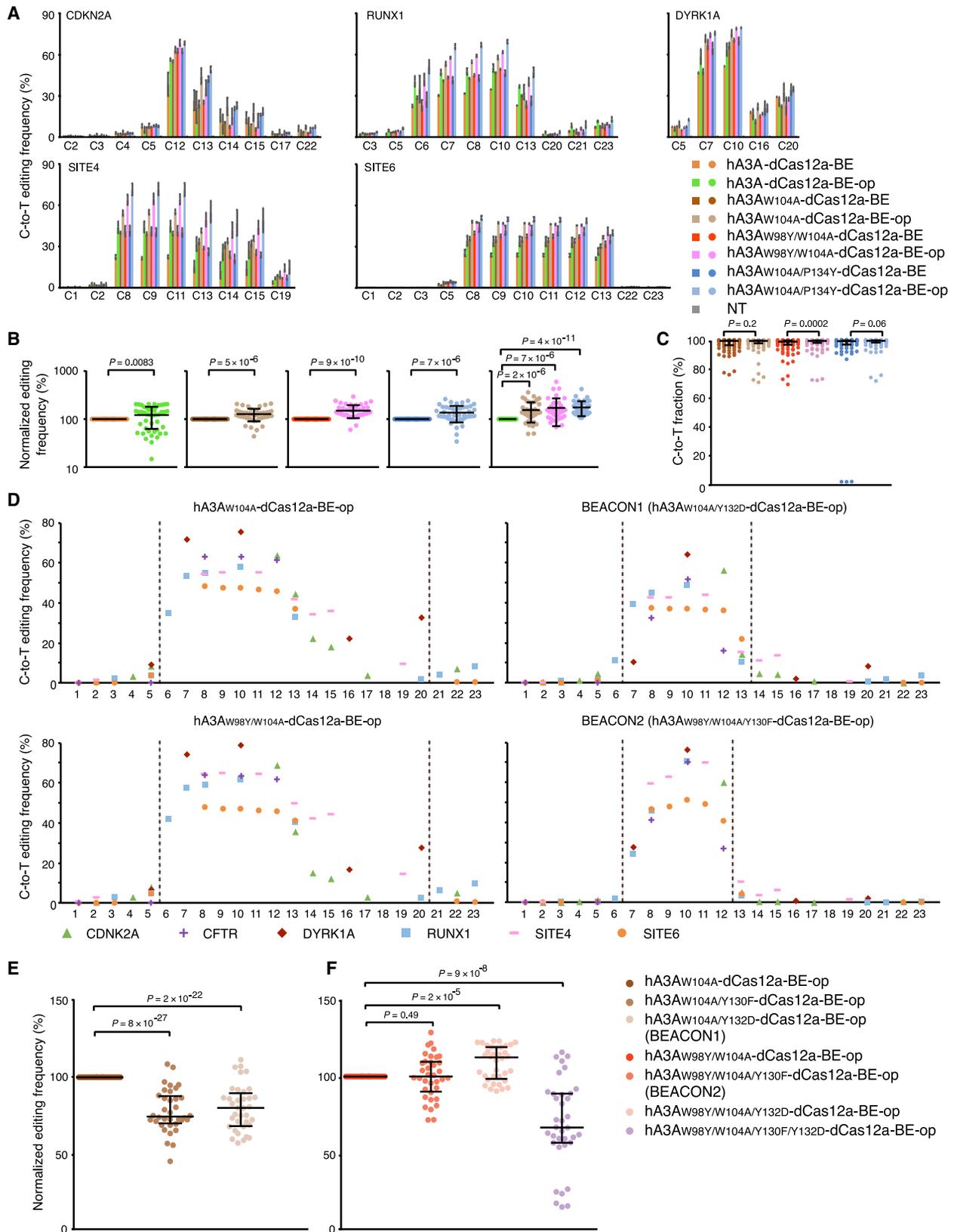
With efficiency-maximized and editing-window-narrowed dCas12a-BEs in hand, we next compared base-editing outcomes induced by BEACON1 and BEACON2 with those by BE2, BE3 (Komor et al., 2016), YE1-BE3 (Kim et al., 2017), and AncBE4max (Koblan et al., 2018) at seven disease-related genomic loci (Figure 5A). Across these sites, both BEACON1 and BEACON2 induced significantly higher editing than BE3 (median: 1.88-fold, $p = 2 \times 10^{-5}$; and median: 2.14-fold, $p = 1 \times 10^{-5}$, respectively; Figure 5E) and reached levels similar to those of AncBE4max (Figure 5E). At some sites, BEACON2 induced even more efficient C-to-T editing than AncBE4max, e.g., sites BMPR2, PDE6C, and PMS2 (Figure 5B). In addition, the product purity yielded by BEACON2 was higher than that by BE3 (median C-to-T fraction, 99.26% versus 87.28%, $p = 1 \times 10^{-6}$; Figures 5C and 5F) and similar to that by AncBE4max (median C-to-T fraction, 99.26% versus 97.80%, $p = 0.0002$; Figures 5C and 5F). As BEACONS use dCas12a as the Cas moiety, they generated no DSB and only background levels of indels (Figure 5D), which were much lower than those by BE3 and AncBE4max (Figure 5G). Importantly, the levels of DDR proteins in BEACON-treated cells were also lower than those in BE3- and Cas9-treated cells (Figures 5H, S4A, and S4B), consistent with the results that BEACON caused almost no toxic DSB (Figures 5D and 5G). Correspondingly, BEACON2 led to less DSBs (Figure S4C) and fewer negative effects on cell proliferation than AncBE4Max and Cas9 (Figure S4D), as DSBs can activate p53 and then lead to negative effects on cell proliferation (Haapaniemi et al., 2018; Ihry et al., 2018). Of note, although they did not trigger DDR robustly (Figures 5H, S4A, and S4B), BE2 and YE1-BE3 (an engineered BE3 with narrowed editing windows) induced lower levels of editing efficiencies than BEACON2 (Figures 5B and 5E).

RNA OT Effects

It has been recently reported that BE3 can generate random transcriptome-wide RNA OT editing (Grünwald et al., 2019;

Figure 3. Rational Engineering of hA3A

- (A) Cut-out view of hA3A-ssDNA structure (PDB: 5KEG) (left) and close-up view of the ssDNA-hA3A interface (right).
 (B) Computational binding energies (dG_{binding}) between each hA3A mutant and ssDNA were calculated along the 200-ns molecular simulation trajectory with a step size of 2. $n = 4,000$ frames.
 (C) Comparison of base-editing efficiencies induced by indicated BEs at different genomic target sites.
 (D) Statistical analysis of normalized C-to-T editing frequencies shown in (C), setting the ones induced by hA3A-dCas12a-BE as 100%.
 (E) Statistical analysis of C-to-T product purity yielded by indicated BEs shown in (C).
 (F) Comparison of base-editing frequencies induced by indicated BEs at different genomic target sites.
 (G) Statistical analysis of normalized C-to-T editing frequencies shown in (F), setting the ones induced by hA3A-dCas12a-BE as 100%.
 (H) Statistical analysis of C-to-T product purity yielded by indicated BEs shown in (F).
 Means \pm SD are from three independent experiments in (C) and (F). $n = 132$ samples from three independent experiments in (D), (E), (G), and (H). p value, one-tailed Student's t test. The median and IQR or $1.5 \times$ IQR (B) are shown.



(legend on next page)

Zhou et al., 2019). Thus, we performed RNA sequencing (RNA-seq) to determine whether BEACONs also trigger RNA OT mutations with RADAR (RNA-editing analysis pipeline to decode all twelve types of RNA-editing events) (Figures 6 and S4E).

In agreement with previous findings that A-to-I is the most common type of RNA editing event (Eisenberg and Levanon, 2018; Ramaswami et al., 2013; Zhu et al., 2013), only A-to-I RNA editing was found to be predominantly distributed in RNA-seq datasets from untransfected control cells or Cas9-treated cells (Figure 6A). However, C-to-U RNA editing was dramatically increased in RNA-seq datasets from BE3-treated cells, demonstrating widespread and unintended RNA OT effects induced by the rA1 moiety of BE3 (Figures 6B and 6C; Zhou et al., 2019). In contrast, much less RNA OT effects were observed in cells treated with AncBE4max, BEACON1, or BEACON2 (Figure 6B). Of note, the hA3A-Y130F-containing BEACON2 only induced RNA OT mutations similar to the background level (Figures 6B and 6D), which is consistent with a basal level of RNA OT effects induced by previously reported hA3A-Y130F-conjugated BEs (Wang et al., 2018; Zhou et al., 2019).

In Vivo Base Editing Induced by BEACON

To test whether the BEACON system can induce base editing *in vivo*, one target cytosine in the mouse hydroxysteroid 17-beta dehydrogenase 3 (*Hsd17b3*) gene was chosen for editing (Figures 7A–7C). We co-injected the mRNA of BEACON1, BEACON2, or rA1-dCas12a-BE with the *Hsd17b3*-targeted crRNA into one-cell mouse embryos and then analyzed base-editing efficiencies at the blastocyst stage (Figure 7A). Among 8 and 14 embryos injected with BEACON1/crRNA and BEACON2/crRNA, respectively, 1 and 9 were found to bear edited cytosine at the *Hsd17b3* locus (Figures 7B, S5C, and S5E). Deep-sequencing analysis confirmed the editing efficiencies at targeted C11 in these embryos (21.04% for BEACON1 and 17.04%–56.84% for BEACON2; Figure 7C). Meanwhile, only a basal level of editing was observed at four other nearby cytosines outside the editing window (C6, C14, C15, and C17; Figure 7C). In contrast, no edited embryo was detected in 11 embryos injected with rA1-dCas12a-BE/crRNA (Figures 7B and S5A). These results further demonstrated the improved *in vivo* editing efficiency by fusing engineered hA3A with dCas12a.

In addition to comparing BEACONs and rA1-dCas12a-BE in mouse embryos, we also transplanted the embryos into surro-

gate mothers (Figure 7A). A total of 24 pups were obtained from BEACON1-injected embryos (Figure 7D), and among them, 2 pups were edited (#7 and #11; Figure S5D) with the C11 editing frequencies at 57.73% or 23.10% in tails (Figure 7E). Consistent with the results from mouse embryos, no edited mouse pup was found from the rA1-dCas12a-BE injected embryos (Figure S5B). To further confirm the base editing in mouse offspring, we dissected one BEACON1-edited offspring (#7) and determined the editing frequencies in different tissues. The editing frequencies in heart, liver, spleen, lung, kidney, brain, and testis ranged from 51.34% to 71.24% (Figure 7E), similar to the editing frequency in the tail (Figure 7E). Finally, minimal indels were detected in the embryos or offspring treated with BEACONs (Figure S5F) and only a background level of distal OT editing (the base editing at OT sites where the crRNA potentially binds to) was detected in the tissues of the C11-edited mouse offspring (Figure S6), demonstrating the editing specificity induced by BEACON.

To further compare *in vivo* editing induced by BEACON2 with BE3 and AncBE4max, we also examined their efficiencies in mouse embryos at a target site within their overlapped editing windows (Wang et al., 2019; Figure S7). Both BE3 and AncBE4max achieved very efficient C-to-T editing in injected mouse embryos (16 of 16 embryos edited and 14 of 15 embryos edited, respectively), and BEACON2 induced a relatively lower efficiency (8 of 15 embryos edited) (Figures S7A–S7D). These results suggested that the nickase activity is beneficial for *in vivo* editing efficiency. However, both BE3 and AncBE4max induced significant levels of indels in mouse embryos but BEACON2 did not (Figures S7E–S7G), consistent with the *in vitro* results (Figures 5D and 5G).

DISCUSSION

Great efforts have been made to develop more efficient and specific BEs for correcting pathogenic-related point mutations to treat relevant diseases (Ranzau and Komor, 2019; Rees and Liu, 2018; Yang et al., 2019), e.g., using nCas9 to boost editing efficiency and introducing alternations in the APOBEC moiety of BEs to suppress unwanted RNA editing (Grünwald et al., 2019; Komor et al., 2016; Nishida et al., 2016; Rees et al., 2019; Zhou et al., 2019). Besides the unwanted OT effects, nCas9-derived BEs also caused the formation of DSBs (Komor

Figure 4. Codon Optimizing and Editing Window Narrowing of hA3A-dCas12a-BE

(A) Comparison of C-to-T editing frequencies induced by indicated BEs and their codon-optimized versions. Means \pm SD are from three independent experiments.

(B) Statistical analysis of normalized C-to-T editing frequencies shown in (A), setting the ones induced by hA3A-dCas12a-BE, hA3A_{W104A}-dCas12a-BE, hA3A_{W98Y/W104A}-dCas12a-BE, hA3A_{W104A/P134Y}-dCas12a-BE, or hA3A-dCas12a-BE-op as 100%. n = 93 samples from three independent experiments.

(C) Statistical analysis of C-to-T product purity yielded by indicated BEs shown in (A). n = 93 samples from three independent experiments.

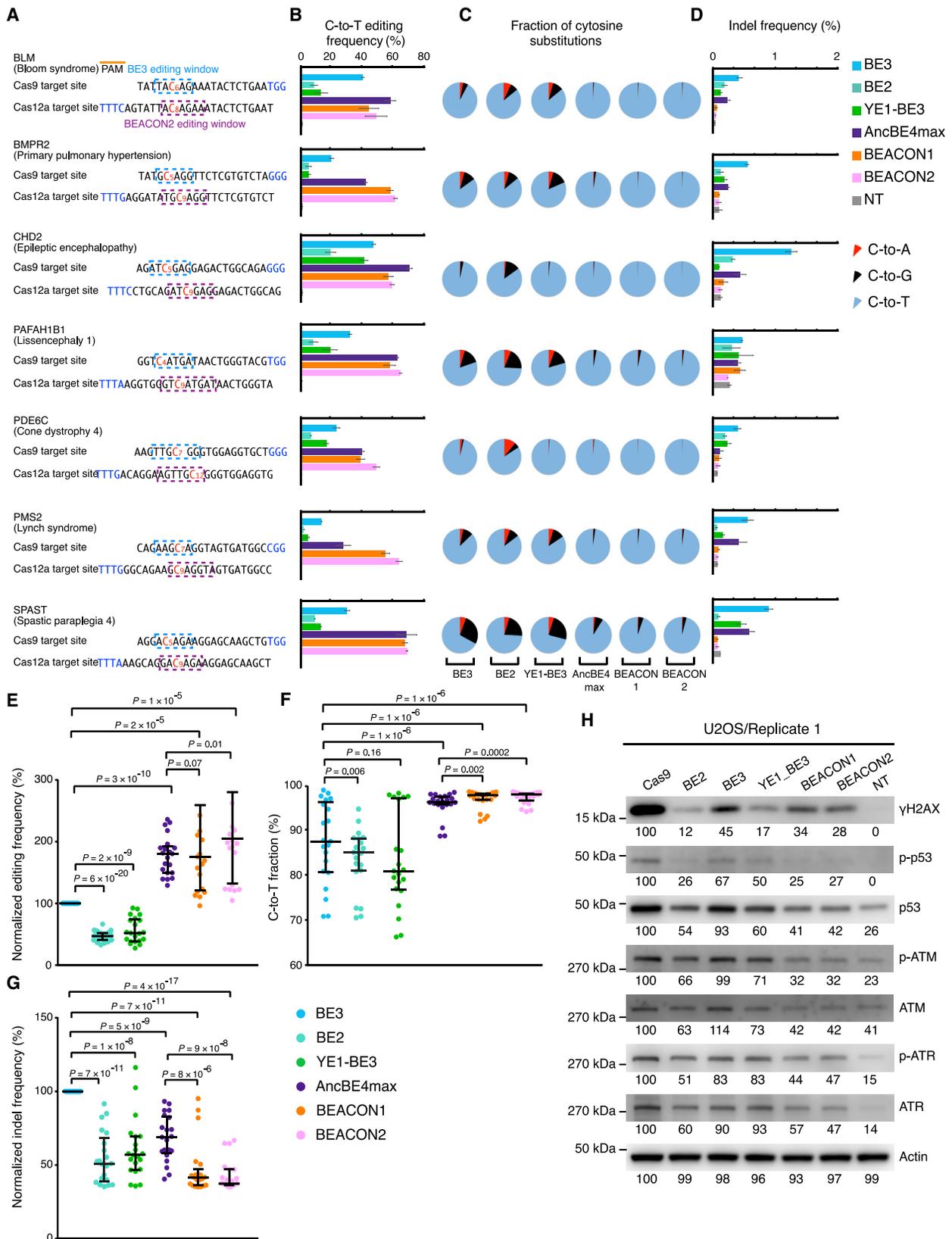
(D) The major editing windows and editing efficiencies of BEACON1 and BEACON2. Average C-to-T editing frequencies are shown, and positions are counted with the base proximal to the PAM setting as position 1.

(E) Statistical analysis of normalized C-to-T editing frequencies shown in (D) and Figure S3B, setting the ones induced by hA3A_{W104A}-dCas12a-BE-op as 100%. n = 36 samples from three independent experiments.

(F) Statistical analysis of normalized C-to-T editing frequencies shown in (D) and Figure S3B, setting the ones induced by hA3A_{W98Y/W104A}-dCas12a-BE-op as 100%. n = 36 samples from three independent experiments.

p value, one-tailed Student's t test (B, C, E, and F). The median and IQR are shown.

See also Figure S3.



(legend on next page)

et al., 2016; Lei et al., 2018; Nishida et al., 2016). DSBs can activate downstream DDR proteins, e.g., the phosphorylation of p53 and H2AX, to disturb cellular homeostasis and eventually cause cell death (Haapaniemi et al., 2018; Ihry et al., 2018), especially when simultaneous editing at multiple sites is pursued (Niu et al., 2017). In addition, it has been reported that DSBs also trigger large deletions or chromosomal rearrangements (Kosicki et al., 2018). As therapeutic applications of gene editing generally involve stem cells or primary cells, in which DDR pathways are active, the unavoidable DNA breaks induced by nCas9-derived BEs may impede their applications in clinics. Different from nCas9-based BEs, DSBs that are induced by dCas12a-conjugated BEs were barely detectable (Figures 5D and 5G); also, dCas12a-conjugated BEs activate lower levels of DDR proteins, including γ H2AX and p53 (Figure 5H), than nCas9-BEs, shedding new light on the potential of dCas12a-conjugated BEs in therapeutics.

The original dCas12a-based BE exhibited relatively low editing efficiency (Li et al., 2018), which hampers its application *in vivo*. In this study, we screened different APOBEC/AID members and found that the fusion of hA3A-dCas12a could induce efficient base editing (Figure 2). With a series of structure-guided protein engineering, codon optimization, and editing window narrowing (Figures 3 and 4), we further developed the BEACON system, which induced high editing efficiency in mammalian cells (Figure 5) and in mouse embryos (Figure 7). The construction of BEACON also has several advantages in terms of editing precision. First, given that Cas12a has a generally high targeting specificity (Kim et al., 2016; Kleinstiver et al., 2016; Yan et al., 2017), we examined the editing induced by BEACON *in vivo* at predicted genomic OT sites and found BEACON induced undetectable OT editing at these sites (Figure S6). Moreover, BEACONS with the Y130F or Y132D alternation in the hA3A moiety have narrowed editing windows (Figures 4D, 4E, 4F, and S3B), which enable precise base editing when multiple cytosines locate in a same target site (Figures 7C and 7E). Finally, we also compared the RNA OT mutations triggered by various BEs and found that BEACON2 induced only background levels of OT editing transcriptome wide (Figure 6). Of note, the incorporation of hA3A in BEACON may also induce efficient base editing in methylated DNA regions and/or GpC sites, as previously developed hA3A-nCas9-BE did (Wang et al., 2018).

In summary, we developed a dCas12a-based BEACON system and successfully applied it for both *in vitro* and *in vivo* editing (Table S1). The BEACON system achieves efficient and specific base editing without generating DNA breaks or triggering DDR

cascades, which is essential for its broad applications in mammalian cells and in clinics.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
 - Lead Contact
 - Materials Availability
 - Data and Code Availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
 - Animals
 - Human cell lines and cell culture
- METHOD DETAILS
 - Plasmid construction
 - Transfection
 - Cell sorting and RNA extraction
 - *In vitro* transcription
 - Microinjection and embryos transfer
 - Western blot
 - Immunofluorescence
 - DNA library preparation and sequencing
- QUANTIFICATION AND STATISTICAL ANALYSIS
 - RADAR
 - Image analysis/Immunostaining quantification
 - Indel frequency calculation
 - Base substitution calculation
 - C-to-T fraction calculation
 - Molecular dynamics simulation
 - Statistical analysis

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at <https://doi.org/10.1016/j.celrep.2020.107723>.

ACKNOWLEDGMENTS

This work was supported by grants 2018YFC1004602 (J.C.), 2018YFA0801401 (J.C.), and 2019YFA0802804 (L.Y.) from National Key R&D Program of China and 31925011 (L.Y.), 91940306 (L.Y.), 31822016 (J.C.), 81872305 (J.C.), 31600654 (J.C.), and 31600619 (B.Y.) from National Natural Science Foundation of China. We thank Molecular and Cell Biology Core Facility and Molecular Imaging Core Facility, School of Life Science and Technology, ShanghaiTech University for providing experimental service.

Figure 5. Comparison of BEACONS and Cas9-BEs

(A) The target sites and editing window induced by Cas9-BEs and Cas12a-BEs.
 (B–D) C-to-T editing frequencies (B), fractions of cytosine substitutions (C), and indel frequencies (D) induced by indicated BEs.
 (E) Statistical analysis of normalized C-to-T editing frequencies shown in (B), setting the ones induced by BE3 as 100%.
 (F) Statistical analysis of C-to-T fractions of base-editing products shown in (C).
 (G) Statistical analysis of normalized indel frequencies shown in (D), setting the ones induced by BE3 as 100%.
 (H) Immunoblots of DDR proteins and their phosphorylated forms triggered by Cas9-mediated gene editing, Cas9-BEs, or BEACON-mediated base editing. The numbers represent individual protein contents relative to that triggered by Cas9 (setting as 100).
 Means \pm SD are from three independent experiments in (B) and (D). $n = 21$ samples from three independent experiments in (E)–(G). p value, one-tailed Student's t test. The median and IQR are shown.
 See also Figure S4.

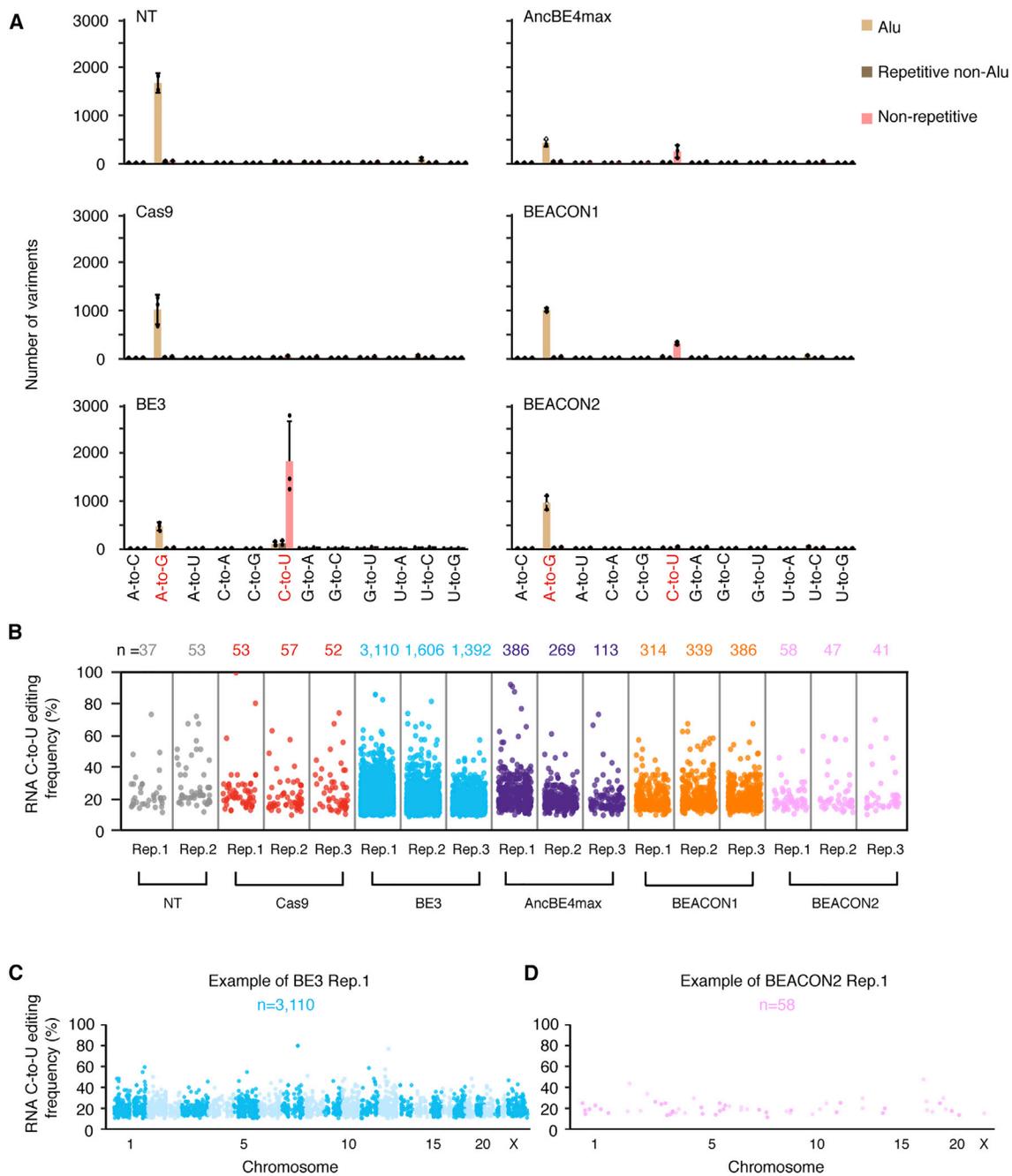


Figure 6. RNA Off-Target Editing Analysis by RADAR

(A) Histogram to show numbers of all 12 types of RNA editing in different defined regions from cells treated with Cas9, Cas9-BEs, and BEACONs. RNA editing was analyzed and visualized by RADAR (Figure S4E). Means \pm SD are from two (NT) or three (Cas9, BE3, AncBE4max, BEACON1, and BEACON2) independent experiments.

(B) Manhattan plot of RNA off-target editing (C-to-U) frequency shown in (A).

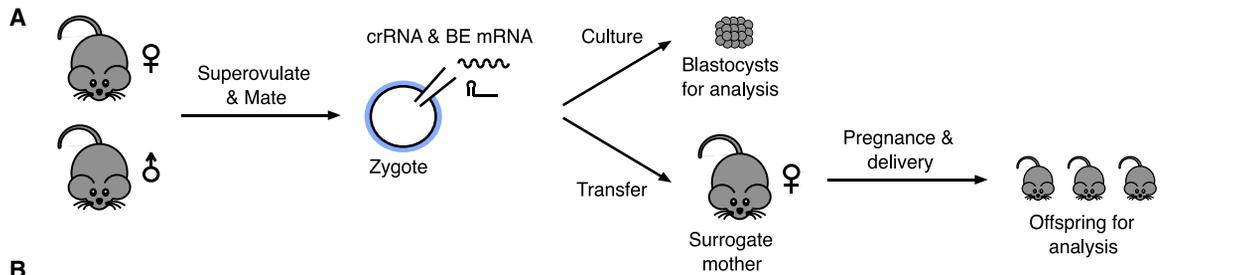
(C and D) The RNA off-target editing frequencies and sites induced by BE3 replicate 1 (C) or BEACON2 replicate 1 (D).

See also Figure S4.

AUTHOR CONTRIBUTIONS

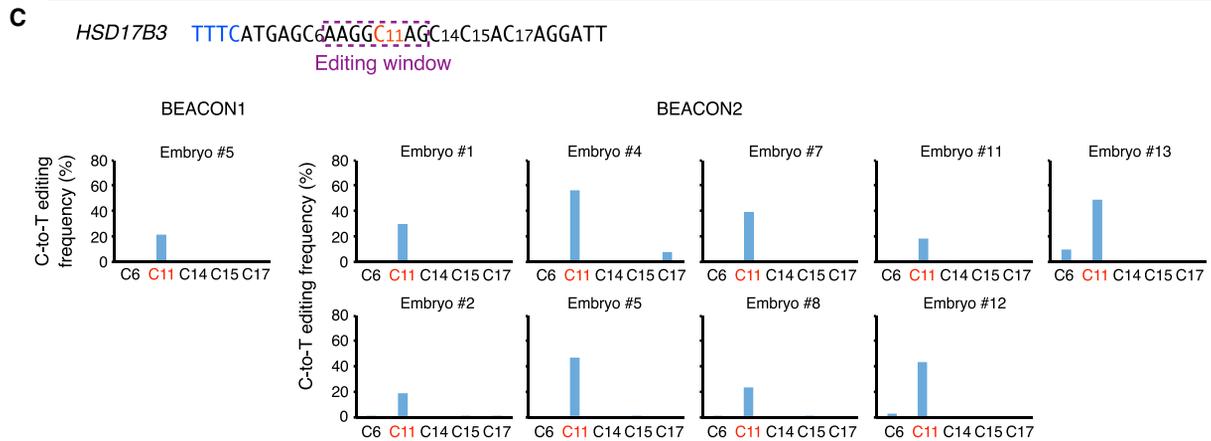
J.C., L.Y., X.H., and Z. Liu conceived, designed, and supervised the project. X.W., C.D., and W.Y. performed most experiments with the help of J. Li, Z.

Lu, W.Z., J. Liang, and J. Wu on cell culture, immunofluorescence, and plasmid construction. S.H. performed animal experiments, supervised by Z. Liu. B.Y. analyzed hA3A structure and performed molecular simulation. J. Wei prepared libraries for deep sequencing, supervised by L.Y., Y.W., and Y.-C.X. performed



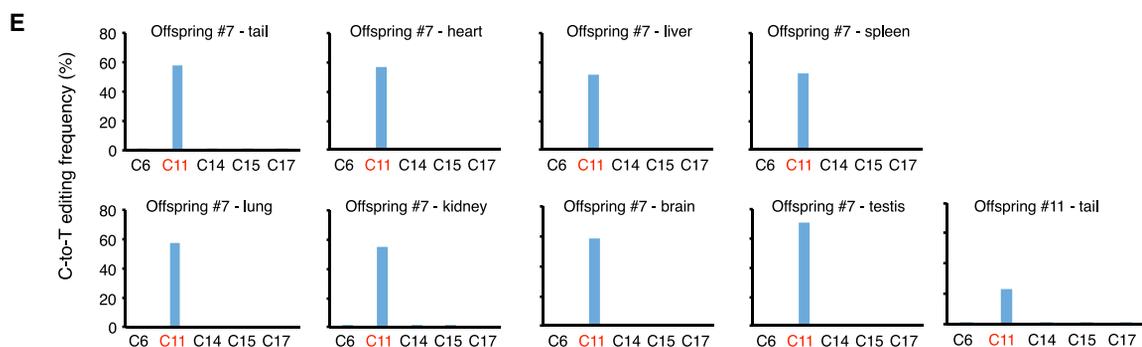
B

Target gene	Base editor	No. of examined blastocysts	Mutant ratio		
			No. of base-edited blastocysts/total blastocysts (%)	No. of indel-containing blastocysts/total blastocysts (%)	No. of proximal OT editing-containing blastocysts/total blastocysts (%)
<i>Hsd17b3</i>	rA1-dCas12a-BE	11	0/11 (0%)	0/11 (0%)	0/11 (0%)
<i>Hsd17b3</i>	BEACON1	8	1/8 (12.5%)	0/8 (0%)	0/8 (0%)
<i>Hsd17b3</i>	BEACON2	14	9/14 (64.3%)	0/14 (0%)	0/14 (0%)



D

Target gene	Base editor	No. of transferred embryos	No. of offspring	Mutant ratio		
				No. of base-edited offspring/total offspring (%)	No. of indel-containing offspring/total offspring (%)	No. of proximal OT editing-containing offspring/total offspring (%)
<i>Hsd17b3</i>	BEACON1	61	24	2/24 (8.3%)	0/24 (0%)	0/24 (0%)
<i>Hsd17b3</i>	rA1-dCas12a-BE	54	8	0/8 (0%)	0/8 (0%)	0/8 (0%)



(legend on next page)

bioinformatics analyses, supervised by L.Y. Z.Z. provided technical support. J.C., L.Y., and B.Y. wrote the paper with input from the other authors.

DECLARATION OF INTERESTS

J.C., L.Y., X.H., B.Y., X.W., C.D., and W.Y. have filed patent applications on aspects on this work. The authors declare no competing non-financial interests.

Received: July 12, 2019

Revised: November 28, 2019

Accepted: May 12, 2020

Published: June 2, 2020

REFERENCES

Biegging, K.T., Mello, S.S., and Attardi, L.D. (2014). Unravelling mechanisms of p53-mediated tumour suppression. *Nat. Rev. Cancer* *14*, 359–370.

Billon, P., Bryant, E.E., Joseph, S.A., Nambiar, T.S., Hayward, S.B., Rothstein, R., and Ciccia, A. (2017). CRISPR-Mediated Base Editing Enables Efficient Disruption of Eukaryotic Genes through Induction of STOP Codons. *Mol. Cell* *67*, 1068–1079.e1064.

Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* *30*, 2114–2120.

Ceccaldi, R., Rondinelli, B., and D'Andrea, A.D. (2016). Repair Pathway Choices and Consequences at the Double-Strand Break. *Trends Cell Biol.* *26*, 52–64.

Chakrabarti, A.M., Henser-Brownhill, T., Monserrat, J., Poetsch, A.R., Luscombe, N.M., and Scaffidi, P. (2019). Target-Specific Precision of CRISPR-Mediated Genome Editing. *Mol. Cell* *73*, 699–713.e696.

Chapman, J.R., Taylor, M.R., and Boulton, S.J. (2012). Playing the end game: DNA double-strand break repair pathway choice. *Mol. Cell* *47*, 497–510.

Chen, J., Miller, B.F., and Furano, A.V. (2014). Repair of naturally occurring mismatches can induce mutations in flanking DNA. *eLife* *3*, e02001.

Cimprich, K.A., and Cortez, D. (2008). ATR: an essential regulator of genome integrity. *Nat. Rev. Mol. Cell Biol.* *9*, 616–627.

Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., and Zhang, F. (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science* *339*, 819–823.

Eisenberg, E., and Levanon, E.Y. (2018). A-to-I RNA editing—immune protector and transcriptome diversifier. *Nat. Rev. Genet.* *19*, 473–490.

Gehrke, J.M., Cervantes, O., Clement, M.K., Wu, Y., Zeng, J., Bauer, D.E., Pinello, L., and Joung, J.K. (2018). An APOBEC3A-Cas9 base editor with minimized bystander and off-target activities. *Nat. Biotechnol.* *36*, 977–982.

Grünewald, J., Zhou, R., Garcia, S.P., Iyer, S., Lareau, C.A., Aryee, M.J., and Joung, J.K. (2019). Transcriptome-wide off-target RNA editing induced by CRISPR-guided DNA base. *Nature* *569*, 433–437.

Haapaniemi, E., Botla, S., Persson, J., Schmierer, B., and Taipale, J. (2018). CRISPR-Cas9 genome editing induces a p53-mediated DNA damage response. *Nat. Med.* *24*, 927–930.

Harris, R.S., and Liddament, M.T. (2004). Retroviral restriction by APOBEC proteins. *Nat. Rev. Immunol.* *4*, 868–877.

Hsu, P.D., Lander, E.S., and Zhang, F. (2014). Development and applications of CRISPR-Cas9 for genome engineering. *Cell* *157*, 1262–1278.

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. *Nat. Biotechnol.* *37*, 626–631.

Ihry, R.J., Worringer, K.A., Salick, M.R., Frias, E., Ho, D., Theriault, K., Kommneni, S., Chen, J., Sondey, M., Ye, C., et al. (2018). p53 inhibits CRISPR-Cas9 engineering in human pluripotent stem cells. *Nat. Med.* *24*, 939–946.

Jinek, M., East, A., Cheng, A., Lin, S., Ma, E., and Doudna, J. (2013). RNA-programmed genome editing in human cells. *eLife* *2*, e00471.

Kent, W.J. (2002). BLAT—the BLAST-like alignment tool. *Genome Res.* *12*, 656–664.

Kim, D., Kim, J., Hur, J.K., Been, K.W., Yoon, S.H., and Kim, J.S. (2016). Genome-wide analysis reveals specificities of Cpf1 endonucleases in human cells. *Nat. Biotechnol.* *34*, 863–868.

Kim, Y.B., Komor, A.C., Levy, J.M., Packer, M.S., Zhao, K.T., and Liu, D.R. (2017). Increasing the genome-targeting scope and precision of base editing with engineered Cas9-cytidine deaminase fusions. *Nat. Biotechnol.* *35*, 371–376.

Kim, D., Paggi, J.M., Park, C., Bennett, C., and Salzberg, S.L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.* *37*, 907–915.

Kleinstiver, B.P., Tsai, S.Q., Prew, M.S., Nguyen, N.T., Welch, M.M., Lopez, J.M., McCaw, Z.R., Aryee, M.J., and Joung, J.K. (2016). Genome-wide specificities of CRISPR-Cas Cpf1 nucleases in human cells. *Nat. Biotechnol.* *34*, 869–874.

Kleinstiver, B.P., Sousa, A.A., Walton, R.T., Tak, Y.E., Hsu, J.Y., Clement, K., Welch, M.M., Horng, J.E., Malagon-Lopez, J., Scarfò, I., et al. (2019). Engineered CRISPR-Cas12a variants with increased activities and improved targeting ranges for gene, epigenetic and base editing. *Nat. Biotechnol.* *37*, 276–282.

Knott, G.J., and Doudna, J.A. (2018). CRISPR-Cas guides the future of genetic engineering. *Science* *361*, 866–869.

Koblan, L.W., Doman, J.L., Wilson, C., Levy, J.M., Tay, T., Newby, G.A., Maiani, J.P., Raguram, A., and Liu, D.R. (2018). Improving cytidine and adenine base editors by expression optimization and ancestral reconstruction. *Nat. Biotechnol.* *36*, 843–846.

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.

Komor, A.C., Badran, A.H., and Liu, D.R. (2017). CRISPR-Based Technologies for the Manipulation of Eukaryotic Genomes. *Cell* *168*, 20–36.

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

Kuscu, C., Parlak, M., Tufan, T., Yang, J., Szlachta, K., Wei, X., Mammadov, R., and Adli, M. (2017). CRISPR-STOP: gene silencing through base-editing-induced nonsense mutations. *Nat. Methods* *14*, 710–712.

Lei, L., Chen, H., Xue, W., Yang, B., Hu, B., Wei, J., Wang, L., Cui, Y., Li, W., Wang, J., et al. (2018). APOBEC3 induces mutations during repair of CRISPR-Cas9-generated DNA breaks. *Nat. Struct. Mol. Biol.* *25*, 45–52.

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* *25*, 1754–1760.

Figure 7. *In Vivo* Editing Mediated by BEACONS in Mouse Embryos and Offspring

(A) Schematic diagram illustrating the procedures of *in vivo* base editing in mouse embryos.

(B) Genotyping of mouse embryos treated with indicated BEs.

(C) *HSD17B3* target sequence and C-to-T editing frequency induced by BEACON1 and BEACON2 in mouse embryos at blastocyst stage.

(D) Genotyping of mouse offspring treated with indicated BEs.

(E) C-to-T editing frequencies induced by BEACON1 in the tails of mouse offspring #7 and #11 and C-to-T editing frequencies induced by BEACON1 in different tissues of offspring #7.

See also Figures S5, S6, and S7.

- 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.
- Mah, L.J., El-Osta, A., and Karagiannis, T.C. (2010). gammaH2AX: a sensitive molecular marker of DNA damage and repair. *Leukemia* *24*, 679–686.
- Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E., and Church, G.M. (2013). RNA-guided human genome engineering via Cas9. *Science* *339*, 823–826.
- McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., and DePristo, M.A. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* *20*, 1297–1303.
- Nishida, K., Arazoe, T., Yachie, N., Banno, S., Kakimoto, M., Tabata, M., Mochizuki, M., Miyabe, A., Araki, M., Hara, K.Y., et al. (2016). Targeted nucleotide editing using hybrid prokaryotic and vertebrate adaptive immune systems. *Science* *353*, aaf8729.
- Niu, D., Wei, H.J., Lin, L., George, H., Wang, T., Lee, I.H., Zhao, H.Y., Wang, Y., Kan, Y., Shrock, E., et al. (2017). Inactivation of porcine endogenous retrovirus in pigs using CRISPR-Cas9. *Science* *357*, 1303–1307.
- Ramaswami, G., Zhang, R., Piskol, R., Keegan, L.P., Deng, P., O'Connell, M.A., and Li, J.B. (2013). Identifying RNA editing sites using RNA sequencing data alone. *Nat. Methods* *10*, 128–132.
- Ranzau, B.L., and Komor, A.C. (2019). Genome, Epigenome, and Transcriptome Editing via Chemical Modification of Nucleobases in Living Cells. *Biochemistry* *58*, 330–335.
- 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.
- Rees, H.A., Wilson, C., Doman, J.L., and Liu, D.R. (2019). Analysis and minimization of cellular RNA editing by DNA adenine base. *Sci. Adv.* *5*, eaax5717.
- Roos, W.P., Thomas, A.D., and Kaina, B. (2016). DNA damage and the balance between survival and death in cancer biology. *Nat. Rev. Cancer* *16*, 20–33.
- Shi, K., Carpenter, M.A., Banerjee, S., Shaban, N.M., Kurahashi, K., Salamango, D.J., McCann, J.L., Starrett, G.J., Duffy, J.V., Demir, Ö., et al. (2017). Structural basis for targeted DNA cytosine deamination and mutagenesis by APOBEC3A and APOBEC3B. *Nat. Struct. Mol. Biol.* *24*, 131–139.
- Shiloh, Y., and Ziv, Y. (2013). The ATM protein kinase: regulating the cellular response to genotoxic stress, and more. *Nat. Rev. Mol. Cell Biol.* *14*, 197–210.
- Thuronyi, B.W., Koblan, L.W., Levy, J.M., Yeh, W.H., Zheng, C., Newby, G.A., Wilson, C., Bhaumik, M., Shubina-Oleinik, O., Holt, J.R., and Liu, D.R. (2019). Continuous evolution of base editors with expanded target compatibility and improved activity. *Nat. Biotechnol.* *37*, 1070–1079.
- van Overbeek, M., Capurso, D., Carter, M.M., Thompson, M.S., Frias, E., Russ, C., Reece-Hoyes, J.S., Nye, C., Gradia, S., Vidal, B., et al. (2016). DNA Repair Profiling Reveals Nonrandom Outcomes at Cas9-Mediated Breaks. *Mol. Cell* *63*, 633–646.
- Wang, X., Li, J., Wang, Y., Yang, B., Wei, J., Wu, J., Wang, R., Huang, X., Chen, J., and Yang, L. (2018). Efficient base editing in methylated regions with a human APOBEC3A-Cas9 fusion. *Nat. Biotechnol.* *36*, 946–949.
- Wang, Y., Gao, R., Wu, J., Xiong, Y.C., Wei, J., Zhang, S., Yang, B., Chen, J., and Yang, L. (2019). Comparison of cytosine base editors and development of the BEable-GPS database for targeting pathogenic SNVs. *Genome Biol.* *20*, 218.
- Yan, W.X., Mirzazadeh, R., Garnerone, S., Scott, D., Schneider, M.W., Kallas, T., Custodio, J., Wernersson, E., Li, Y., Gao, L., et al. (2017). BLISS is a versatile and quantitative method for genome-wide profiling of DNA double-strand breaks. *Nat. Commun.* *8*, 15058.
- Yang, B., Yang, L., and Chen, J. (2019). Development and Application of Base. *CRISPR J.* *2*, 91–104.
- Zetsche, B., Gootenberg, J.S., Abudayyeh, O.O., Slaymaker, I.M., Makarova, K.S., Essletzbichler, P., Volz, S.E., Joung, J., van der Oost, J., Regev, A., et al. (2015). Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell* *163*, 759–771.
- Zhou, B.B., and Elledge, S.J. (2000). The DNA damage response: putting checkpoints in perspective. *Nature* *408*, 433–439.
- 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.
- Zhu, S., Xiang, J.F., Chen, T., Chen, L.L., and Yang, L. (2013). Prediction of constitutive A-to-I editing sites from human transcriptomes in the absence of genomic sequences. *BMC Genomics* *14*, 206.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit monoclonal anti-phospho-ATM (Ser1981) (D6H9)	Cell Signaling Technology	Cat#5883; RRID: AB_10835213
Rabbit monoclonal anti-phospho-p53 (Ser15)	Cell Signaling Technology	Cat#9284; RRID: AB_331464
Rabbit monoclonal anti-phospho-ATR (Thr1989)	Cell Signaling Technology	Cat#58014; RRID: AB_2722679
Mouse monoclonal anti-ATM [2C1(1A1)]	Abcam	Cat#ab78; RRID: AB_306089
Mouse monoclonal anti-p53 (DO-1)	Santa Cruz	Cat#Sc-126; RRID: AB_628082
Rabbit monoclonal anti-ATR	Abcam	Cat#ab10312; RRID: AB_297050
Rabbit monoclonal anti-gamma H2AX (phosphor S139) [EP854(2)Y]	Abcam	Cat#ab81299; RRID: AB_1640564
Mouse monoclonal anti-beta actin	Absci	Cat#AB21800
Goat Anti-Rabbit IgG H&L (Alexa Fluor® 594)	Abcam	Cat#ab150084; RRID: AB_2734147
Chemicals, Peptides, and Recombinant Proteins		
Protease inhibitor cocktail (EDTA-free, 100X in DMSO)	Medchemexpress (MCE)	Cat#HY-K0010
Phosphatase inhibitor cocktail (2X Tubes, 100X)	Bimake	Cat#B15001
LIPOFECTAMINE LTX	Life, Invitrogen	Cat#15338100
QuickExtract™ DNA Extraction Solution	Epicenter	Cat#QE09050
Pierce ECL Western Blotting substrate	Thermo	Cat#32106
BisBenzimide H 33342 trihydrochloride	Sigma	Cat#B2261
CC/Mount (TM) tissue mounting medium	Sigma	Cat#C9368
Critical Commercial Assays		
MMESSAGE MMACHINE T7 ULTRA 1 KIT	Ambion	Cat#AM1345
RNeasy Mini Kit (50)	QIAGEN	Cat#74104
MEGASHORTSCRIPT T7 KIT 25 RXNS EACH	Ambion	Cat#AM1354
MEGACLEAR KIT 20 RXNS EACH	Ambion	Cat#AM1908
Clone Express®-II One step cloning Kit	Vazyme	Cat#C112-02
Plasmid DNA extraction Kit	TIANGEN	Cat#DP107-T
Next Ultra II FS DNA Library Prep Kit	NEB	Cat#E7805L
TruSeq Stranded Total RNA Library Prep Kit	Illumina	Cat#RS-122-2201
Deposited Data		
Deep-sequencing data	this paper	GSE145552
Deep-sequencing data	this paper	OEP000822
Experimental Models: Cell Lines		
Human: HEK293T	ATCC	Cat#CRL-11268
Human: U2OS	ATCC	Cat#HTB-96
Experimental Models: Organisms/Strains		
Mouse: C57BL/6JSlac	Shanghai SLAC Laboratory Animal Co. Ltd www.slaccas.com	N/A
Mouse: DBA2J	Shanghai SLAC Laboratory Animal Co. Ltd www.slaccas.com	N/A
Mouse: ICR	Shanghai SLAC Laboratory Animal Co.	N/A
Oligonucleotides		
Oligos used for plasmid construction, see Table S2	this paper	N/A
crRNA and sgRNA target sequences and PCR primers for amplifying genomic DNA, see Table S3	this paper	N/A

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Recombinant DNA		
pCMV-hA3B-dCas12a-BE	this paper	N/A
pCMV-hAID-dCas12a-BE	this paper	N/A
pCMV-hA3A-dCas12a-BE	this paper	N/A
pCMV-hA3A-dCas12a-BE-op	this paper	N/A
hA3A _{W98A} -dCas12a-BE	this paper	N/A
hA3A _{W98Y} -dCas12a-BE	this paper	N/A
hA3A _{W104A} -dCas12a-BE	this paper	N/A
hA3A _{W104Y} -dCas12a-BE	this paper	N/A
hA3A _{P134A} -dCas12a-BE	this paper	N/A
hA3A _{P134Y} -dCas12a-BE	this paper	N/A
hA3A _{C101S} -dCas12a-BE	this paper	N/A
hA3A _{C106S} -dCas12a-BE	this paper	N/A
hA3A _{W98Y/W104A} -dCas12a-BE	this paper	N/A
hA3A _{W98Y/P134Y} -dCas12a-BE	this paper	N/A
hA3A _{W104A/P134Y} -dCas12a-BE	this paper	N/A
hA3A _{W98Y/W104A/P134Y} -dCas12a-BE	this paper	N/A
hA3A-dCas12a-BE-op	this paper	N/A
hA3A _{W104A} -dCas12a-BE-op	this paper	N/A
hA3A _{W104A/Y130F} -dCas12a-BE-op	this paper	N/A
hA3A _{W104A/Y132D} -dCas12a-BE-op	this paper	N/A
hA3A _{W98Y/W104A} -dCas12a-BE-op	this paper	N/A
hA3A _{W98Y/W104A/Y130F} -dCas12a-BE-op	this paper	N/A
hA3A _{W98Y/W104A/Y132D} -dCas12a-BE-op	this paper	N/A
hA3A _{W104A/P134Y} -dCas12a-BE-op	this paper	N/A
hA3A _{W104A/P134Y/Y130F} -dCas12a-BE-op	this paper	N/A
hA3A _{W104A/P134Y/Y132D} -dCas12a-BE-op	this paper	N/A
hA3A _{W98Y/W104A/Y130F/Y132D} -dCas12a-BE-op	this paper	N/A
pCMV-BE3	Komor et al., 2016	Addgene Plasmid #73021
pCMV-AncBE4max	Koblan et al., 2018	Addgene Plasmid #112094
pBK-YE1-BE3	Kim et al., 2017	Addgene Plasmid #85174
Software and Algorithms		
FastQC v0.11.4	https://www.bioinformatics.babraham.ac.uk/projects/fastqc/	https://www.bioinformatics.babraham.ac.uk/projects/fastqc/
Trimmomatic v0.36	Bolger et al., 2014	http://www.usadellab.org/cms/?page=trimmomatic
BWA v0.7.9	Li and Durbin, 2009	http://bio-bwa.sourceforge.net/
HISAT2 v2.1.0	Kim et al., 2019	https://ccb.jhu.edu/software/hisat2/index.shtml
Picard v2.7.1	https://broadinstitute.github.io/picard/	https://broadinstitute.github.io/picard/
GATK v4.1.2.0	McKenna et al., 2010	https://gatk.broadinstitute.org
BLAT v364	Kent, 2002	https://genome.ucsc.edu/cgi-bin/hgBlat
R v3.5.1	https://www.r-project.org	https://www.r-project.org
RADAR	This paper	https://github.com/YangLab/RADAR
HPB	Zhu et al., 2013	N/A
Zeiss LSM800	https://www.zeiss.com/microscopy/us/products/confocal-microscopes.html	https://www.zeiss.com/microscopy/us/products/confocal-microscopes.html

RESOURCE AVAILABILITY

Lead Contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Jia Chen (chenjia@shanghaitech.edu.cn).

Materials Availability

This study did not generate new unique reagents.

Data and Code Availability

The original DNA deep-sequencing and RNA sequencing data from this study have been deposited in the NCBI Gene Expression Omnibus (GEO: GSE145552) and the National Omics Data Encyclopedia (accession number OEP000822).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Animals

All mice experiment procedures were approved by the Animal Care and Use Committee of the Institute of Neuroscience, Chinese Academy of Sciences, Shanghai, China. Mice were maintained in standard cages in an Assessment and Accreditation of Laboratory Animal Care credited specific pathogen free facility under a 12 hr dark-light cycle. Sample sizes were estimated based on experiments in similar studies, and the experiments were not randomized or blinded. Female B6D2F1 (C57BL/6 × DBA2J) mice (4-week-old) were superovulated and mated with male B6D2F1 mice (4-week-old). The injected zygotes were cultured in KSOM medium at 37°C (in 5% CO₂) and transferred to oviducts of pseudopregnant females (ICR mice, 10-week-old).

Human cell lines and cell culture

HEK293T and U2OS cells from ATCC were maintained in DMEM (10566, GIBCO/Thermo Fisher Scientific) + 10% FBS (16000-044, GIBCO/Thermo Fisher Scientific) and regularly tested to exclude mycoplasma contamination.

METHOD DETAILS

Plasmid construction

Primer sets (hA3A_dCpf1_PCR_F/hA3A_dCpf1_PCR_R) were used to amplify the fragment Human_APOBEC3A with template pUC57-Human_APOBEC3A (synthesized by Genscript). Then the fragment Human_APOBEC3A was cloned into the SacI and SmaI linearized dCpf1-BE with plasmid recombination kit Clone Express® (Vazyme, C112-02) to generate the hA3A-dCas12a-BE expression vector pCMV-SV40NLS-hAPOBEC3A-XTEN-dLbCpf1(D832A/E1006A/D1125A)-SV40NLS-SGGS-UGI-SV40NLS. hA3B-dCas12a-BE and hAID-dCas12a-BE expression vectors were constructed with the same strategy.

Two primer sets (hA3A_dCpf1_PCR_F/hA3A_dCpf1_W98A_PCR_R) (hA3A_dCpf1_W98A_PCR_F/hA3A_dCpf1_PCR_R) were used to amplify the W98A-containing fragment hA3A-W98A. Then the fragments were cloned into the Apal and SmaI linearized hA3A-dCas12a-BE expression vector to generate the hA3A_{W98A}-dCas12a-BE expression vector pCMV-SV40NLS-hAPOBEC3A_W98A-XTEN-dLbCpf1(D832A/E1006A/D1125A)-SV40NLS-SGGS-UGI-SV40NLS. hA3A_{W98Y}-dCas12a-BE, hA3A_{W104A}-dCas12a-BE, hA3A_{W104Y}-dCas12a-BE, hA3A_{P134A}-dCas12a-BE, hA3A_{P134Y}-dCas12a-BE, hA3A_{C101S}-dCas12a-BE, hA3A_{C106S}-dCas12a-BE, hA3A_{W98Y/W104A}-dCas12a-BE, hA3A_{W98Y/P134Y}-dCas12a-BE, hA3A_{W104A/P134Y}-dCas12a-BE and hA3A_{W98Y/W104A/P134Y}-dCas12a-BE expression vectors were constructed with the same strategy.

pCMV-hA3A-dCas12a-BE-op was commercially synthesized. Two primer sets (hA3A_dCpf1_PCR_F-op/hA3A_dCpf1_W104A_PCR_R-op) (hA3A_dCpf1_W104A_PCR_F-op/hA3A_dCpf1_PCR_R-op) were used to amplify the W104A containing fragment hA3A-W104A-op. Then the fragments were cloned into the NotI and SmaI linearized hA3A-dCas12a-BE-op expression vector to generate the hA3A_{W104A}-dCas12a-BE-op expression vector pCMV-SV40NLS-hAPOBEC3A_W104A-XTEN-dLbCpf1(D832A/E1006A/D1125A)-SV40NLS-SGGS-UGI-SV40NLS-op. hA3A_{W104A/Y130F}-dCas12a-BE-op, hA3A_{W104A/Y132D}-dCas12a-BE-op, hA3A_{W98Y/W104A}-dCas12a-BE-op, hA3A_{W98Y/W104A/Y130F}-dCas12a-BE-op, hA3A_{W98Y/W104A/Y132D}-dCas12a-BE-op, hA3A_{W104A/P134Y}-dCas12a-BE-op, hA3A_{W104A/P134Y/Y130F}-dCas12a-BE-op, hA3A_{W104A/P134Y/Y132D}-dCas12a-BE-op and hA3A_{W98Y/W104A/Y130F/Y132D}-dCas12a-BE-op expression vectors were constructed with the same strategy.

Primer sets (T7-LBCpf1-For/T7-LBCpf1-REV) were used to construct the pLb-Cpf1-pGL3-T7-crRNA expression vector with template pLb-Cpf1-pGL3-U6-sgRNA. Then the fragment was recombined by plasmid recombination kit Clone Express® (Vazyme, C112-02).

Oligonucleotides hDYRK1A_cpf1_FOR/hDYRK1A_cpf1_REV were annealed and ligated into BsaI linearized pLb-Cpf1-pGL3-U6-sgRNA to generate crDYRK1A expression vector pcrDYRK1A. Oligonucleotides hDYRK1A_cpfsp_FOR/hDYRK1A_cpfsp_REV were annealed and ligated into BsaI linearized pGL3-U6-sgRNA-PGK-puromycin to generate sgDYRK1A expression vector psgDYRK1A. Other crRNA and sgRNA expression vectors were constructed by the same way.

The sequences of the oligos used for plasmid construction are listed in [Table S2](#) and the amino acid sequences of BEs are listed in [Table S3](#).

Transfection

For base editing in genomic DNA, HEK293T and U2OS cells were seeded in a 24-well plate at a density of 1.6×10^5 per well and transfected with 200 μ L serum-free Opti-MEM that contained 5.04 μ L LIPOFECTAMINE LTX (Life, Invitrogen), 1.68 μ L LIPOFECTAMINE plus (Life, Invitrogen), 1 μ g rA1-dCas12a-BE expression vector (or hA3A-dCas12a-BE, hA3B-dCas12a-BE, hAID-dCas12a-BE, hA3A_{W98A}-dCas12a-BE, hA3A_{W98Y}-dCas12a-BE, hA3A_{W104A}-dCas12a-BE, hA3A_{W104Y}-dCas12a-BE, hA3A_{P134A}-dCas12a-BE, hA3A_{P134Y}-dCas12a-BE, hA3A_{C101S}-dCas12a-BE, hA3A_{C106S}-dCas12a-BE, hA3A_{W98Y/W104A}-dCas12a-BE, hA3A_{W98Y/P134Y}-dCas12a-BE, hA3A_{W104A/P134Y}-dCas12a-BE, hA3A_{W98Y/W104A/P134Y}-dCas12a-BE, hA3A_{W104A}-dCas12a-BE-op, hA3A_{W104A/Y130F}-dCas12a-BE-op, hA3A_{W104A/Y132D}-dCas12a-BE-op, hA3A_{W98Y/W104A}-dCas12a-BE-op, hA3A_{W98Y/W104A/Y130F}-dCas12a-BE-op, hA3A_{W98Y/W104A/Y132D}-dCas12a-BE-op, hA3A_{W104A/P134Y}-dCas12a-BE-op, hA3A_{W104A/P134Y/Y130F}-dCas12a-BE-op, hA3A_{W104A/P134Y/Y132D}-dCas12a-BE-op and hA3A_{W98Y/W104A/Y130F/Y132D}-dCas12a-BE-op, BE3 (addgene, #73021), BE2 (addgene, #73020), YE1-BE3 (addgene, #85174) and AncBE4max (addgene, #112094) expression vector) and 0.68 μ g crRNA or sgRNA expression vector. After 72hr, the genomic DNA was extracted from the cells with QuickExtract DNA Extraction Solution (QE09050, Epicenter) or the cells were lysed in 2 \times SDS loading buffer for western blot.

Cell sorting and RNA extraction

After 40hr after transfection, Cells in the first 15% of the fluorescence intensity were sorted by FACSAriaIII. Total RNAs of sorted cells were extracted by using the RNeasy Mini Kit (QIAGEN #74104).

In vitro transcription

rA1-dCas12a-BE, hA3A_{W104A/Y132D}-dCpf1-BE-op, hA3A_{W98Y/W104A/Y130F}-dCpf1-BE-op, BE3 (addgene, #73021) and AncBE4max (addgene, #112094) vector including T7 promoter was linearized by BbsI (NEB, R3539L) and *in vitro* transcribed using T7 Ultra Kit (Ambion, AM1345). mRNA was purified by Mini Kit (QIAGEN, 74104). crRNA oligos were annealed into LbCpf1-pGL3-T7-crRNA expression vectors with T7 promoter. Then crRNA was amplified and *in vitro* transcribed by MEGAscript Kit (Ambion, AM1354). The sgRNAs were purified by MEGAclear Kit (Ambion, AM1908) according to the manufacturer's protocols. sgRNA expression vectors were constructed by the same way. Primers used for transcription *in vitro* were listed in [Table S4](#).

Microinjection and embryos transfer

Female B6D2F1 (C57BL/6 \times DBA2J) mice (4-week-old) were superovulated and mated with male B6D2F1 mice. One-cell-stage embryos were collected to inject with crRNA/sgRNA (50 ng/ μ l) and rA1-dCas12a-BE, BEACON1, BEACON2, BE3 and AncBE4max (100 ng/ μ l) into the cytoplasm of zygotes in a droplet of M2 medium containing 5 μ g/ml cytochalasin B (CB) using a piezo (Primetech) microinjector. The injected zygotes were cultured in KSOM mediums at 37°C under 5% of CO₂ in air and transferred to oviducts of pseudopregnant females at 0.5 day post copulation.

Western blot

Transfected cells were lysed in NP40 lysis buffer (50 mM Tris pH 8.0, 150mM NaCl, 1% NP-40, 0.1% SDS, 1mM PMSF, protease inhibitors, and phosphatase inhibitor) for 30 min on ice, then incubated at 97 °C for 15 min, separated by SDS-PAGE (Genscript) in sample loading buffer and proteins were transferred to nitrocellulose membranes (Thermo Fisher Scientific). After blocking with TBST (25 mM Tris pH 8.0, 150 mM NaCl, and 0.1% Tween 20) containing 5% (w/v) nonfat dry milk and 1% BSA for 2h, the membrane was reacted overnight with indicated primary antibody. After extensive washing, the membranes were reacted with HRP-conjugated secondary antibodies for 1h. Reactive bands were developed in ECL (Thermo Fisher Scientific) and detected with Amersham Imager 680.

Immunofluorescence

Transfected U2OS cells were fixed in 4% PFA for 1 hour at 4°C. After washing in PBS, cells were blocked with PBS containing 0.3% Triton X-100 and 2.5% BSA for 1 hour. Then, cells were incubated overnight at 4°C with primary antibodies (anti-gamma H2AX, 1:300) diluted in PBS containing 0.1% Tween-20. Signals were developed with Alexa Fluor secondary antibodies (Abcam, 1:500) at room temperature for 1 hour after washing in PBST (0.1% Tween-20). Before mounting, cells were counterstained with DAPI. Pictures were collected on a confocal microscope (Zeiss LSM800).

DNA library preparation and sequencing

Target genomic sites were PCR amplified by high-fidelity DNA polymerase PrimeSTAR HS (Clonetechn) with primers flanking each examined crRNA or sgRNA target site. The PCR primers used to amplify target genomic sequences were listed in [Table S4](#). Indexed DNA libraries were prepared by using the TruSeq ChIP Sample Preparation Kit (Illumina) with some minor modifications. Briefly, the PCR products were fragmented by Covaris S220 and then amplified by using the TruSeq ChIP Sample Preparation Kit (Illumina). After being quantitated with Qubit High-Sensitivity DNA kit (Life, Invitrogen), PCR products with different tags were pooled together for

deep sequencing by the Illumina NextSeq 500 (2 × 150) or HiSeq X Ten (2 × 150) at CAS-MPG Partner Institute for Computational Biology Omics Core, Shanghai, China. Raw read qualities were evaluated by FastQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). For paired-end sequencing, only R1 reads were used. Adaptor sequences and read sequences on both ends with Phred quality score lower than 28 were trimmed. Trimmed reads were then mapped with the BWA-MEM algorithm (BWA v0.7.9a) to target sequences. After being piled up with samtools (v0.1.18), indels and base substitutions were further calculated.

QUANTIFICATION AND STATISTICAL ANALYSIS

RADAR

We developed a computational pipeline RADAR (RNA-editing Analysis-pipeline to Decode All twelve-types of RNA-editing events, <https://github.com/YangLab/RADAR>) to detect and visualize all possible twelve-types of RNA editing events from RNA-seq datasets. RADAR consists of three steps (Figure S4E).

STEP 1: RNA-seq read mapping. After quality control by FastQC (version 0.11.4, parameters: default), RNA-seq reads are trimmed by Trimmomatic (version 0.36, parameters: TruSeq3-PE-2.fa:2:30:10 TRAILING:25 MINLEN:30) (Bolger et al., 2014) to remove read sequences with low quality, and then mapped to ribosomal DNA (rDNA) sequences by BWA-MEM algorithm (version 0.7.9a, parameters: default) to remove reads mapped to redundant rDNAs. To capture more mismatches for RNA editing candidates, a two-round unique mapping strategy is then applied to align high-quality RNA-seq reads to human hg38 reference genome, sequentially by HISAT2 (version 2.1.0, parameters: `-rna-strandness RF-no-mixed-secondary-no-temp-splicesite-known-splicesite-infile-no-soft-clip-score-min L,-16,0-mp 7,7-rfg 0,7-rdg 0,7-max-seeds 20 -k 10-dta`) (Kim et al., 2019) with up to two mismatches and by BWA-MEM (version, parameters: default). Unique mapped reads by HISAT2 and BWA-MEM with up to six mismatches are selected and combined for subsequent analysis. After marking duplicate reads identified by Picard (version 2.7.1, parameters: CREATE_INDEX = true VALIDATION_STRINGENCY = SILENT) in the BAM file, uniquely-mapped reads that span exon-exon junctions are split into segments by the GATK (version 4.1.2.0) (McKenna et al., 2010) command, SplitNCigarReads (parameters: default). Base quality scores of all uniquely-mapped reads are recalibrated by two GATK (version 4.1.2.0) commands, BaseRecalibrator (parameters: default) and ApplyBQSR (parameters: default).

STEP 2: RNA editing calling. RNA variants are determined from the BAM file with uniquely-mapped reads by the GATK (version 4.1.2.0) command HaplotypeCaller (parameters: `-minimum-mapping-quality 0-stand_call_conf 0-dont-use-soft-clipped-bases true`). After filtering out RNA variants overlapped with single nucleotide polymorphisms (SNPs) from dbSNP version 151 (<https://www.ncbi.nlm.nih.gov/SNP/>), the 1000 Genomes Project (<https://www.internationalgenome.org/>) and the University of Washington Exome Sequencing Project (<https://evs.gs.washington.edu/EVS/>), low-quality RNA variants with mapped read numbers < 2, hits per billion-mapped-bases (HPB; Zhu et al., 2013) < 3 or editing ratio < 0.05 were further removed.

RNA variants for potential RNA editing events are classified into three groups according to their genomic locations as previously described (Ramaswami et al., 2013), including in *Alu*, non-*Alu* repetitive or non-repetitive regions. Different to those in *Alu* regions, RNA variants in non-*Alu* repetitive and non-repetitive regions are further filtered with a series of stringent cutoffs to remove false positive. In brief, RNA variants in simple repeats, in mononucleotide microsatellites ≥ 5 bp or within 4 bp of splice junctions are removed. In addition, BLAST-like alignment tool (BLAT, version 364, parameters: `-repMatch = 2253 -stepSize = 5`) (Kent, 2002) is used to remove RNA variants within highly similar regions. Finally, RNA variants within bidirectional transcription regions are also removed.

All twelve types of RNA editing events in *Alu*, non-*Alu* repetitive and non-repetitive regions are eventually determined according to the strands of overlapped genes (Human: hg38 knownGene.txt updated at 2015/6/28).

STEP 3: RNA editing visualizing. All possible RNA-editing events from each given RNA-seq dataset are listed in an Excel file. Numbers of all twelve-types of RNA editing events are plotted by histograms according to their genomic locations in *Alu*, repetitive non-*Alu* and non-repetitive regions. Manhattan plots are further used to illustrate RNA editing ratios of selected types of RNA-editing events, such as C-to-U or A-to-G.

Image analysis/Immunostaining quantification

The intensities of western blot bands were determined with ImageJ.

Indel frequency calculation

For dCas12a-BEs, indels were estimated in the aligned regions spanning from upstream 3 nucleotides to the downstream 48 nucleotides both according to PAM sites (55bp). For dCas9- and nCas9- BEs, indels were estimated in the aligned regions spanning from upstream 8 nucleotides to the target sites to downstream 19 nucleotides to PAM sites (50 bp). Indel frequencies were subsequently calculated through dividing the counts of reads containing at least one inserted and/or deleted nucleotides by the counts of all the mapped reads in the same region. Counts of indel-containing reads and total mapped reads are listed in Table S5.

Base substitution calculation

Base substitutions were selected at each position of the examined crRNA or sgRNA target sites that mapped with at least 1,000 independent reads, and obvious base substitutions were only observed at the targeted base editing sites. Counts of reads for each base and total reads are listed in Table S6. Base substitution frequencies were calculated by dividing base substitution reads by total reads.

C-to-T fraction calculation

C-to-T fractions were calculated by dividing T reads with the sum of non-C reads (A, G, T) at indicated editing sites.

Molecular dynamics simulation

Figures depicting the interaction between hA3A and ssDNA were prepared using the Pymol program (The PyMOL Molecular Graphics System, Version 2.1 Schrödinger, LLC.) All mutant hA3A-DNA complex structures were modified from the structure of hA3A in complex with ssDNA (PDBid:5KEG) and then energy optimized using Protein Preparation Wizard from Schrodinger Suite 2019-1 (<https://www.schrodinger.com>) following default settings. The simulation systems were solvated with SPC water and neutralized, containing Cl⁻ and Na⁺ ions at a concentration of 0.15 M to mimic physiological ionic strength. 200 ns molecular dynamics simulations were then performed with Desmond from Schrodinger Suite 2019-1 (<https://www.schrodinger.com>), using OPLS3 force field. During the simulation, a Desmond implemented multi-stage MD simulation protocol was employed with Temperature T and pressure P kept constant at 310 K and 1 atm, respectively. The RMSD and RMSF of protein and DNA molecule as well as the protein-ligand contacts diagram were calculated with Simulation Interactions Diagram from Schrodinger Suite 2019-1 (<https://www.schrodinger.com>). The binding energy between hA3A variants and ssDNA along the simulation trajectory was calculated with third-party script `thermal_mmgsbsa.py` from Schrodinger, with a `step_size` of 2.

Statistical analysis

For all figures, error bars show mean \pm SD. Statistics and graphs were prepared using Prism software version 8. Statistical significance of differences in C-to-T editing frequency, indel frequency and C-to-T fraction were determined using Student's t test (one-tailed). For all tests, p values of < 0.05 were considered statistically significant.