

Engineering RsDddA as mitochondrial base editor with wide target compatibility and enhanced activity

Kai Cheng,^{1,7} Cao Li,^{1,7} Jiachuan Jin,^{2,7} Xuezhen Qian,¹ Jiayin Guo,^{1,3} Limini Shen,¹ YiChen Dai,¹ Xue Zhang,⁴ Zhanwei Li,⁴ Yichun Guan,² Fei Zhou,⁵ Jin Tang,⁴ Jun Zhang,¹ Bin Shen,^{1,3,4,6} and Xin Lou⁴

¹State Key Laboratory of Reproductive Medicine, Women's Hospital of Nanjing Medical University, Nanjing Maternity and Child Health Care Hospital, Nanjing Medical University, Nanjing, Jiangsu, China; ²Center for Reproductive Medicine, The Third Affiliated Hospital of Zhengzhou University, Zhengzhou, Henan, China; ³Gusu School, Nanjing Medical University, Nanjing, Jiangsu, China; ⁴Research Institute of Intelligent Computing, Zhejiang Laboratory, Hangzhou, Zhejiang, China; ⁵Cambridge-Suda Genomic Resource Center, Jiangsu Key Laboratory of Neuropsychiatric Diseases Research, Medical College of Soochow University, Suzhou, China; ⁶Center for Global Health, School of Public Health, Nanjing Medical University, Nanjing, Jiangsu, China

Double-stranded DNA-specific cytidine deaminase (DddA) base editors hold great promise for applications in bio-medical research, medicine, and biotechnology. Strict sequence preference on spacing region presents a challenge for DddA editors to reach their full potential. To overcome this sequence-context constraint, we analyzed a protein dataset and identified a novel DddA_{tox} homolog from *Ruminococcus* sp. AF17-6 (RsDddA). We engineered RsDddA for mitochondrial base editing in a mammalian cell line and demonstrated RsDddA-derived cytosine base editors (RsDdCBE) offered a broadened NC sequence compatibility and exhibited robust editing efficiency. Moreover, our results suggest the average frequencies of mitochondrial genome-wide off-target editing arising from RsDdCBE are comparable to canonical DdCBE and its variants.

INTRODUCTION

Owing to the vital and diverse roles of the mitochondria in human cell, mutations in mitochondrial DNA (mtDNA) can lead to devastating multi-system disorders with prevalence greater than 1 in 5,000 in adults.^{1–3} Technologies that enable precise introduction of genetic variants within mtDNA are urgently needed both to deepen our understanding of mitochondrial disease and develop strategies to correct the mutations for therapeutic applications.

In the past few years, steps have been taken to materialize mitochondrial genome engineering.⁴ Programmable nucleases such as mitochondria-targeted zinc-finger nucleases (mtZFN) and mitochondria-targeted transcription activator-like effector nucleases (mito-TALEN) could effectively achieve directional shifting of mtDNA heteroplasmy by making double-strand breaks within mtDNA.^{5–7} Although the present genome editing tools, including base editors and prime editors, could install precise changes in target nuclear DNA, mtDNA has remained resistant to CRISPR-Cas-based technologies due to inefficient nucleic acid importation to mitochondria.⁸ Recently, an important new type of tool for mtDNA manipulation has emerged. Through har-

nessing an interbacterial toxin, Mok et al. demonstrated efficient C•G-to-T•A conversions in the mtDNA sequences.⁹ In this design, double-stranded DNA deaminase toxin A (BcDddA) derived from *Burkholderia cenocepacia* was split into two inactive halves to avoid toxicity in host cells, then each of the halves was fused with custom-designed transcription activator-like effector (TALE) arrays and a uracil glycosylase inhibitor (UGI) to make DddA-derived cytosine base editors (BcDdCBEs). Since being devised, DdCBEs have been successfully deployed for mitochondrial base editing in multiple experiment systems.^{10–14} Although DdCBEs are highly versatile, due to the strict sequence-context preference of BcDddA, the application of the initial DdCBE was predominantly restrained to tC targets. Efforts have been made to expand the targeting scope of canonical BcDdCBE. In a recent study, researchers adapted phage-assisted directed evolution and obtained BcDddA variants that offered a broadened HC (H = a, c, or t) sequence compatibility, while the experiment was unable to yield variants that can edit gC substrates with improved efficiency.¹⁵

In this study, we sought to overcome the sequence-context constraint by identifying unexploited DddAs from metagenomic datasets and

Received 3 April 2023; accepted 1 September 2023;
<https://doi.org/10.1016/j.omtn.2023.09.005>.

⁷These authors contributed equally

Correspondence: Jin Tang, Research Institute of Intelligent Computing, Zhejiang Laboratory, Hangzhou, Zhejiang, China.

E-mail: jin.tang@zhejianglab.com

Correspondence: Jun Zhang, State Key Laboratory of Reproductive Medicine, Women's Hospital of Nanjing Medical University, Nanjing Maternity and Child Health Care Hospital, Nanjing Medical University, Nanjing, Jiangsu, China.

E-mail: zhang_jun@njmu.edu.cn

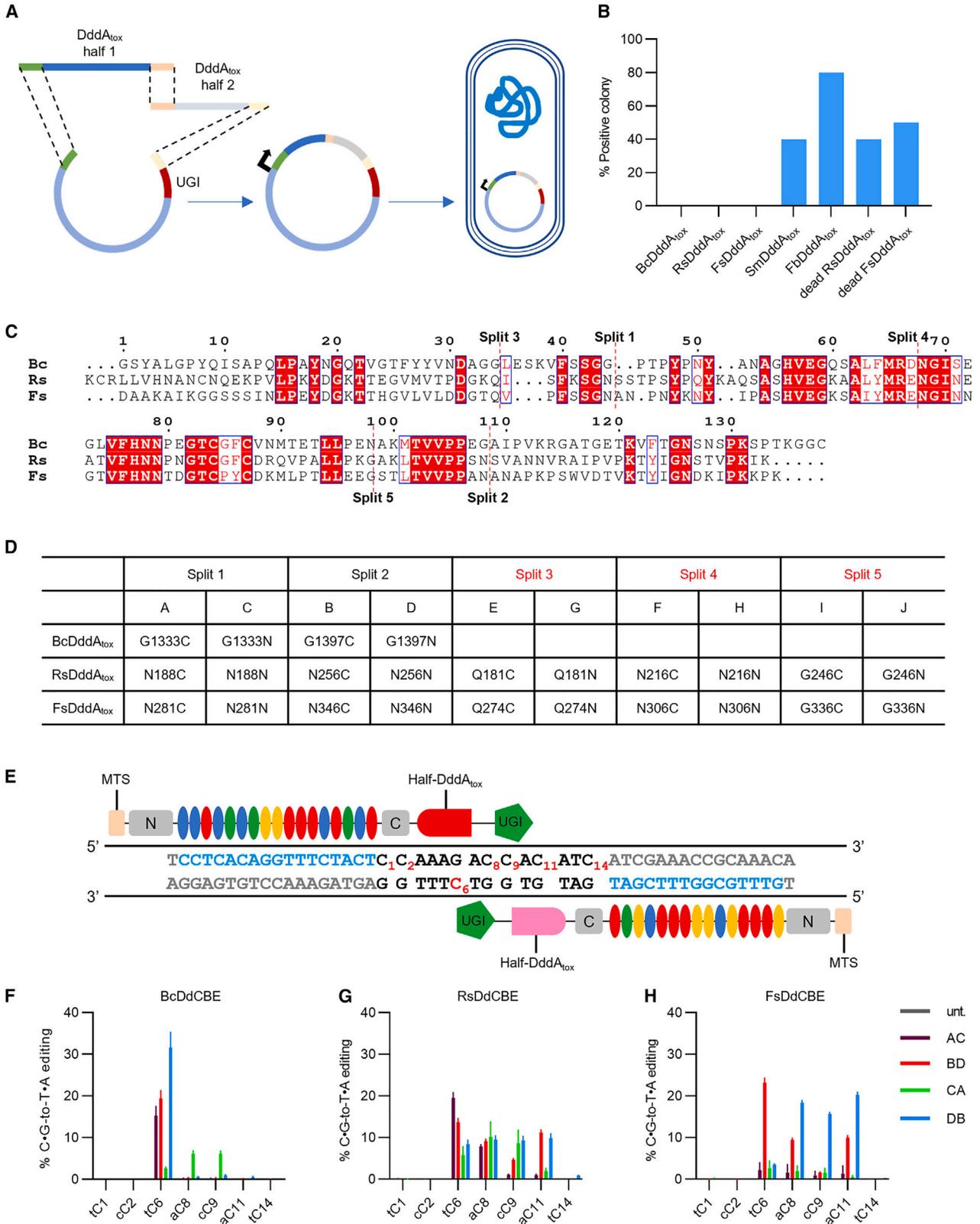
Correspondence: Bin Shen, State Key Laboratory of Reproductive Medicine, Women's Hospital of Nanjing Medical University, Nanjing Maternity and Child Health Care Hospital, Nanjing Medical University, Nanjing, Jiangsu, China.

E-mail: binshen@njmu.edu.cn

Correspondence: Xin Lou, Research Institute of Intelligent Computing, Zhejiang Laboratory, Hangzhou, Zhejiang, China.

E-mail: xin.lou@zhejianglab.edu.cn





(legend on next page)

engineered them for mitochondrial base editing in mammalian cells. We found DddA from *Ruminococcus* sp. (RsDddA) displayed high virulence in *Escherichia coli* viability screen, and RsDddA-derived cytosine base editors (RsDdCBE) could mediate effectual C•G-to-T•A conversion. Sequence compatibility profiling showed RsDdCBE could also support high editing efficiencies at aC (up to 71.51%), cC (up to 42.61%), and gC (up to 86.10%) targets, and this indicated that RsDdCBE offers an NC sequence compatibility. Moreover, mtDNA-wide analysis showed off-target editing frequencies of RsDdCBE are similar to canonical BcDdCBE and its variants.

RESULTS

Identification of toxic DddA_{tox} homologs

To surpass the sequence-context constraint of BcDdCBE, we searched for previously uncharacterized DddA proteins from NCBI metagenomic datasets. We hypothesized that remotely homologous proteins of BcDddA might evolve to have different targeting contexts. To identify these proteins, an iterative search with hidden Markov model (HMM)-based algorithms was conducted with BcDddA protein sequence. A list of potential homologous proteins that have 40%–60% sequence similarity with BcDddA_{tox} was compiled and structure models for these proteins were generated with AlphaFold 2.¹⁶ Then we aligned the structure models of these proteins with structure of BcDddA_{tox} and chose four candidates from different clades with high structure similarity (measured by root-mean-square deviation [RMSD]) for the following experiments (Figure S1; Table S1). To validate the virulence of these potential cytidine deaminases, we deployed an *E. coli* colony formation assay. The sequence coding the potential toxin domain (DddA_{tox}) of candidates was synthesized as split halves, then the halves were incorporated into a prokaryotic open reading frame (ORF) by Gibson assembly and transferred into *E. coli* for expression (Figure 1A). Functional DddA_{tox} would reduce the viability of *E. coli* then lead to compromised colony formation. Among the candidates that demonstrated their cytidine deaminase activity in this screen, the DddA_{tox} from *Ruminococcus* sp. AF17-6 (RsDddA_{tox}) and *Falcatimonas* sp. MSJ-15 (FsDddA_{tox}) displayed comparable virulence with BcDddA_{tox}, while the DddA_{tox} from *Streptomyces massaporeus* (SmDddA_{tox}) and a Frankiaceae bacterium (FbDddA_{tox}) showed low virulence (Figure 1B). When a substitution was introduced to the conserved catalytic glutamic acid residue of RsDddA_{tox} (E206A) and FsDddA_{tox} (E296A), the defect in *E. coli* growth was relieved (Figure 1B), indicating that virulence was dependent on the cytidine deaminase activity. Then we focused on the two toxic DddA_{tox} homologs. Structure models of RsDddA_{tox} and FsDddA_{tox} were predicted with AlphaFold 2 and aligned with the structure of BcDddA_{tox}, and the result showed these two DddA_{tox} homologs potentially adopt a highly analogous fold with BcDddA_{tox}

(Figure S1B), even though they only shared 45.45%, and 44.35% sequence identity with BcDddA_{tox} respectively (Figure 1C). Compared with BcDddA_{tox}, the most prominent disparity is that the first β sheet of BcDddA_{tox} is absent in the two homologs (Figure S1B); this divergence may offer them with the virtue of different biochemical properties.

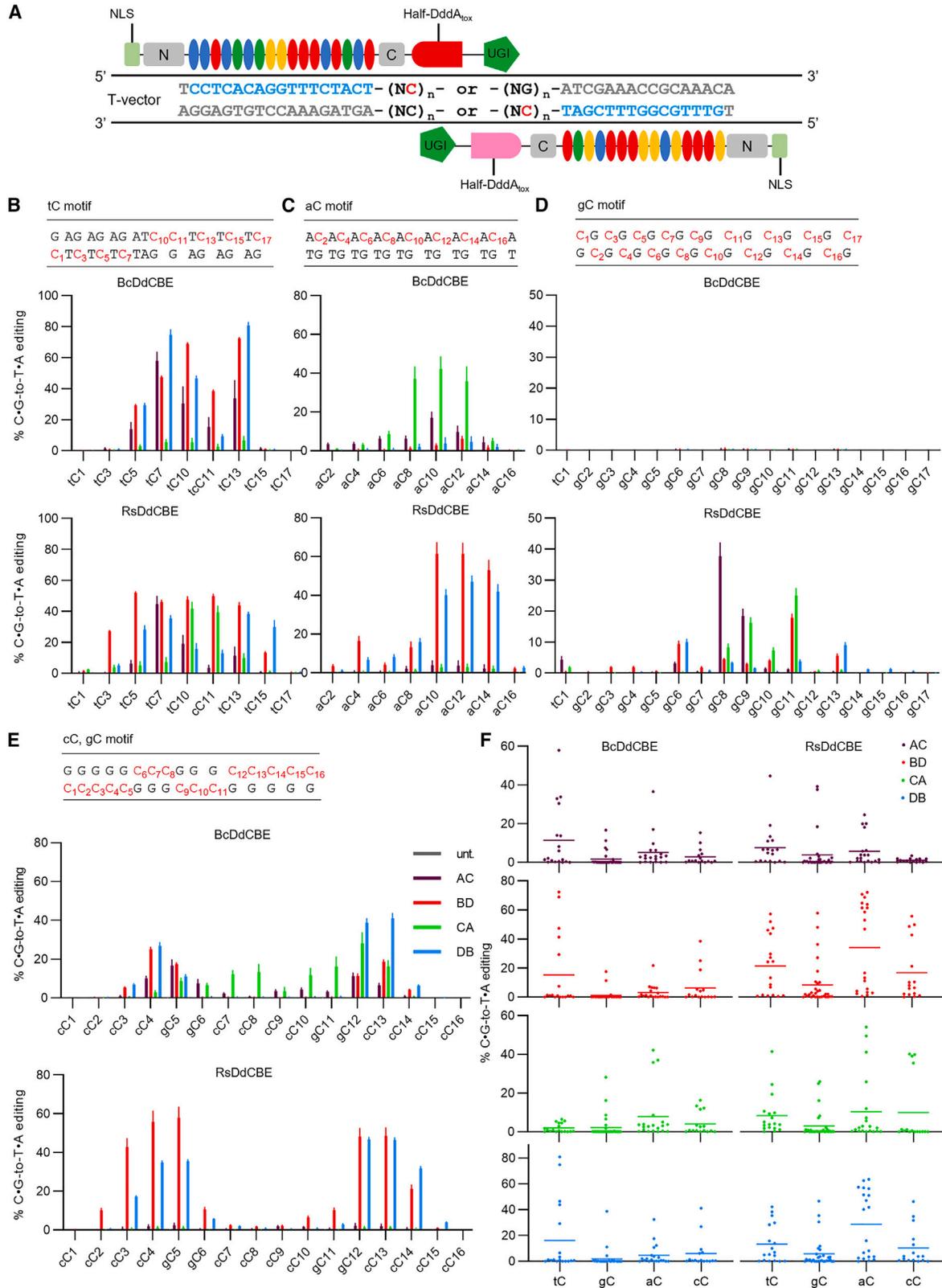
To examine the applicability of RsDddA_{tox} and FsDddA_{tox} as mitochondrial base editors, we split them into halves at five sites (Figure 1C). Split type 1 and 2, corresponding to the original G1333 or G1397 split of BcDddA_{tox}, yielded AC, CA, BD, and DB pairs. Split types 3–5 yielded EG, GE, FH, HF, IJ, and JI pairs (Figure 1D). Each pair was incorporated into the scaffold of our previously published DdCBE, which targeting human *MT-ND5* (Figure 1E), resulting in RsDdCBE and FsDdCBE pairs with different split type and orientation combinations. HEK293FT cells were transfected with these pairs and collected for targeted deep sequencing 96 h post transfection. Compared with the BcDdCBE, the sequencing results revealed that RsDdCBE and FsDdCBE with canonical split type displayed effectual C•G-to-T•A conversion in the spacing region (Figures 1F–1H), while RsDdCBE and FsDdCBE pairs derived from split type 3–5 yielded low editing efficiency (Figures S2A and S2B). Promisingly, besides editing the canonical tC target, certain pairs of RsDdCBE and FsDdCBE could also effectively edit aC and cC targets (Figures 1G and 1H). These results collectively suggest RsDdCBE and FsDdCBE are capable of performing genome editing in mtDNA with potentially wide target compatibility.

Profiling the target compatibility and editing window of RsDdCBE

To preliminarily assess the target compatibility of RsDdCBE and FsDdCBE, these editors with nuclear localization signals (NLSs) were co-transfected into HEK293FT cells with a plasmid bearing different combinations of tC, gC, aC, and cC targets within a 16- or 17-bp spacer (Figure 2A). At 96 h post transfection, plasmids were recovered and subjected to Sanger sequencing to preliminarily examine the C•G-to-T•A conversion. The results showed RsDdCBE could support editing on NC targets, while FsDdCBE displayed imperceptible activity at gC sites (Figures S3A–S3D), so we focused the following study on characterizing the property of RsDdCBE with targeted deep sequencing. At tC targets, RsDdCBEs yielded a slightly lower editing efficiencies (up to 51.28%) compared to BcDdCBE (up to 82.64%) with a wider editing window (Figure 2B). At aC sites, RsDdCBEs are capable to achieve appreciably higher editing efficiencies than BcDdCBE (Figure 2C). In contrast to BcDdCBEs, which resulted in negligible editing (<0.4%) at gC targets, RsDdCBEs could mediate effectual conversion on gC sites (up to 37.72%) (Figure 2D). At plasmids containing variant combinations

Figure 1. Identification and examination of new cytidine deaminases

(A) Schematic of *E. coli* colony formation assay. (B) Results of the *E. coli* clone formation assay. (C) Sequence alignment of DddA_{tox} proteins. Conserved amino acids are marked in red. The positions of split are marked by red dash lines. (D) Different split types of DddA_{tox}. (E) DdCBE architecture targeting human m. 13513G>A site. (F–H) The editing efficiencies of BcDdCBE, RsDdCBE, and FsDdCBE within the spacing region in HEK 293FT cells. AC, BD, CA, and DB indicate different split types and orientation combinations. Data are presented as mean \pm SD from three independent biological replicates. unt, untreated cell.



(legend on next page)

of NC targets, RsDdCBEs also performed better over BcDdCBEs and showed a relatively wide editing window (Figure 2E and S4). To better visualize the target compatibility of RsDdCBE, editing efficiencies of all sites within the spacing region were plotted (Figure 2F). The data show RsDdCBE has no obvious sequence-context preference, while BcDdCBE displays limited activity at gC, cC, and aC targets.

Next, we preliminarily characterized the editing window of RsDdCBE with plasmid editing assay the same as context profiling. Because continuous cC dinucleotide repeats can lead to errors of plasmid construction or sequencing, we generated a separate library of plasmids with target regions that contained 5–10 repeats of tC, aC, or gC dinucleotide (Figure S5). The targeted deep-sequencing results showed that, at tC targets, RsDdCBE displayed the highest efficiency at the right of center within the spacing region, and cytosines located as far as 2–6 bp from the center could be effectually edited (Figures 3 and S5). At aC targets, the highest editing efficiency arose at the right side of the spacing region and the editing window could span up to 12 bp (Figure 3B). At gC targets, the efficient editing window of DB split type shifts from left to right with spacer's extension, while the BD combination showed the opposite trend (Figure 3C). These results indicate that, consistent with the above context profiling, RsDdCBE potentially has a relatively wide editing window, and this feature makes it a more flexible base-editing tool.

Installing previously inaccessible pathogenic mtDNA mutations with RsDdCBE

To further validate the utility of RsDdCBE offered by its wide target compatibility, we designed RsDdCBE pairs to introduce disease-associated C•G-to-T•A conversions at nC positions within human mtDNA. The m.3890G>A mutation affects an extremely conserved amino acid and has been associated with Leber's hereditary optic neuropath.^{17,18} RsDdCBE could install this missense mutation in a tC context with editing efficiency up to 32.92%, comparable with BcDdCBE (Figure 4A). The m.3277G>A variant may lead to deficiencies in leucine tRNA metabolism and has been linked to hypertension.¹⁹ RsDdCBE could mediate conversion at this site in an aC context with editing efficiency up to 26.51%, whereas BcDdCBE had little activity at this site (Figure 4B). On another set of aC targets, RsDdCBE also displayed substantial C•G-to-T•A conversion capability (Figure S6). The m.9380G>A is a synonymous polymorphism of COX3 and has been associated with Alzheimer's disease risk.²⁰ RsDdCBE could effectively (up to 20.70%) convert target cytosine to thymine in a gC context (Figure 4C), while BcDdCBE, even the evolved V6 and V11 variants, showed inefficient editing capacity at the site. The m. 8393C>T mutation causes a proline-to-serine change on ATP8 protein and has been associated with brain pseudoatrophy and mental regression.²¹ RsDdCBE could edit the on-target cytosines

in a gC context with efficiency up to 39.41%, significantly higher than BcDdCBE (16.73%) and BcDdCBE-V6 (2.65%) (Figure 4D). It is interesting to note that, although BcDdCBE-V11 failed to achieve effective editing at gC targets in the original study,¹⁵ ATP8-BcDdCBE-V11 could mediate conversion at m. 8393C>A with an efficiency comparable to ATP8-RsDdCBE (Figure 4D). Similar to the canonical BcDdCBE and its variants, most of the RsDdCBE-generated alleles contained both on-target edits and bystander edits that may result in unintended changes (Figures 4A–4D).

Recently, Mi et al. identified DddA homolog from *Simiioa sunii* (Ddd_Ss) and demonstrated Ddd_Ss-derived cytosine base editors (DdCBE_Ss) could introduce mutations in the gC context,²² so we generated RsDdCBE and DdCBE_Ss with two split types (Figures S7A–S7C) and compared the editing efficiency of the editors on four different gC targets (m.3460G>A, m.3635G>A, m.12258C>T, and m.14250C>T). The sequencing results showed DdCBE_Ss could convert target cytosine with efficiency up to 6.6%, while RsDdCBE could achieve significantly higher conversion frequency (up to 86.1%) on all four targets (Figures S7D–S7G).

It has come to our notice that, on the m.3890G>A site, DB pairs (N256C split) of RsDdCBE achieved high on-target editing with minimized bystander editing (Figure 4A). To explore the mechanisms underlying the editing result on this site, we generated RsDdCBE pairs that targeted similar spacer sequences (NNGGGG at the center of the spacer) and observed comparable minimized bystander editing on the m.6150G>A site (Figure S8). This indicated that the precise editing relies upon both the sequence context and editor configuration, and further studies are required to realize more precise editing.

To further confirm the utility of RsDdCBE as a base editor, we assessed its editing activity in other two cell types: the mouse N2A cell line and human-derived HeLa cells. After RsDdCBE pairs targeting m.2820G>A or m.3177G>A were transfected into N2A cells, cytosine editing on the target site could be detected with efficiency up to 54.77% (Figure S9A). Similarly, in HeLa cells, RsDdCBE pairs targeting m.13513G>A, m.3890G>A, or m.8393C>T could edit the on-target cytosines with efficiency up to 20.95% (Figure S9B).

These results collectively indicate that RsDdCBE enables high levels of base editing at endogenous mitochondrial targets with nC contexts.

Mitochondrial and nuclear off-target activity of RsDdCBE

To profile mitochondrial off-target editing activities of RsDdCBE, we performed whole mtDNA sequencing with the depth of 3,000–8,000×. The average frequencies of mitochondrial genome-wide off-target C•G-to-T•A editing by m. 3277G>A-RsDdCBE

Figure 2. Editing efficiencies of RsDdCBE at NC targets

(A) Architecture of RsDdCBE for plasmid editing assay. Each spacing region contains NC repeats within the top strand or bottom strand. (B–E) Editing efficiencies of RsDdCBE and BcDdCBE on NC-target plasmids. Subscripted numbers refer to the positions of cytosines in the spacing region, counting the DNA nucleotide immediately after the binding site of Left-TALES as position 1. AC, BD, CA, and DB refer to different splits and orientations of DdCBE listed in Figure 1D (F) Scatterplot for all targets from (B–E). (B–E) Data are mean ± SD from three independent biological replicates. unt., HEK293FT cells transfected with target plasmids only.

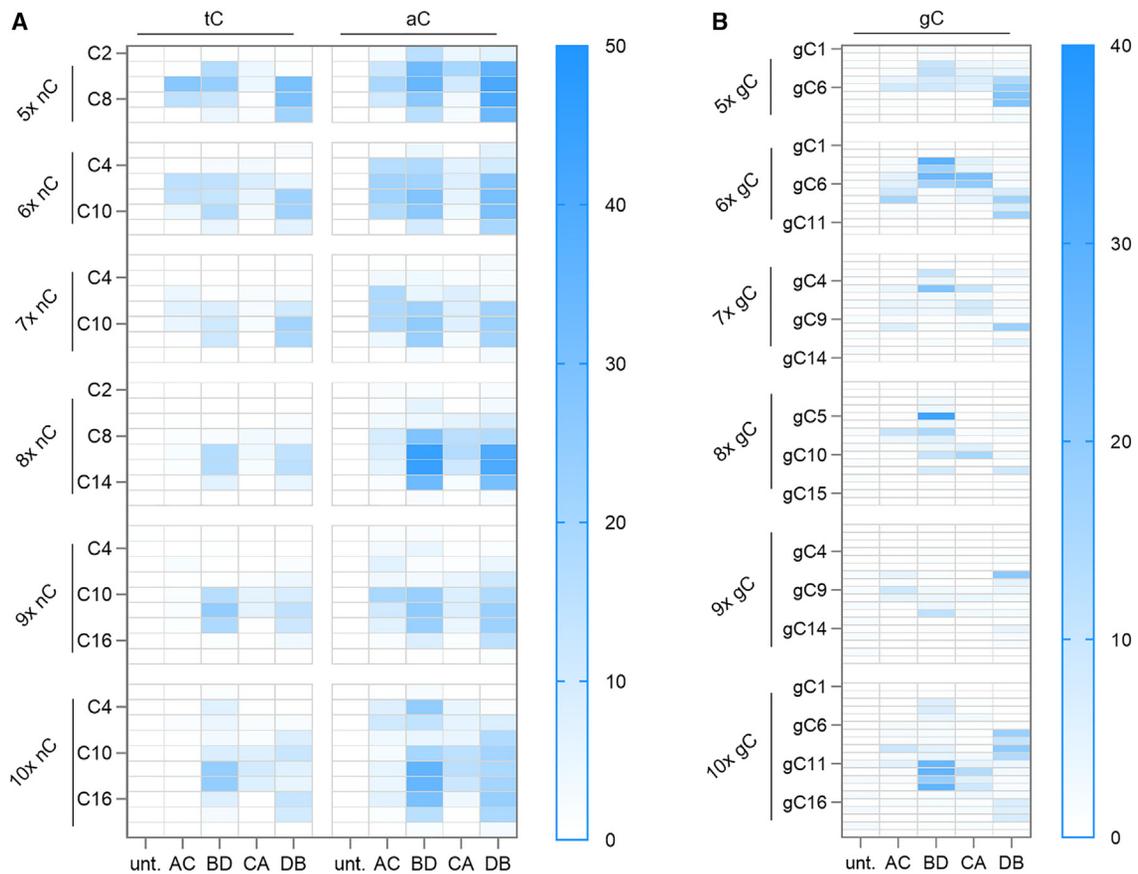


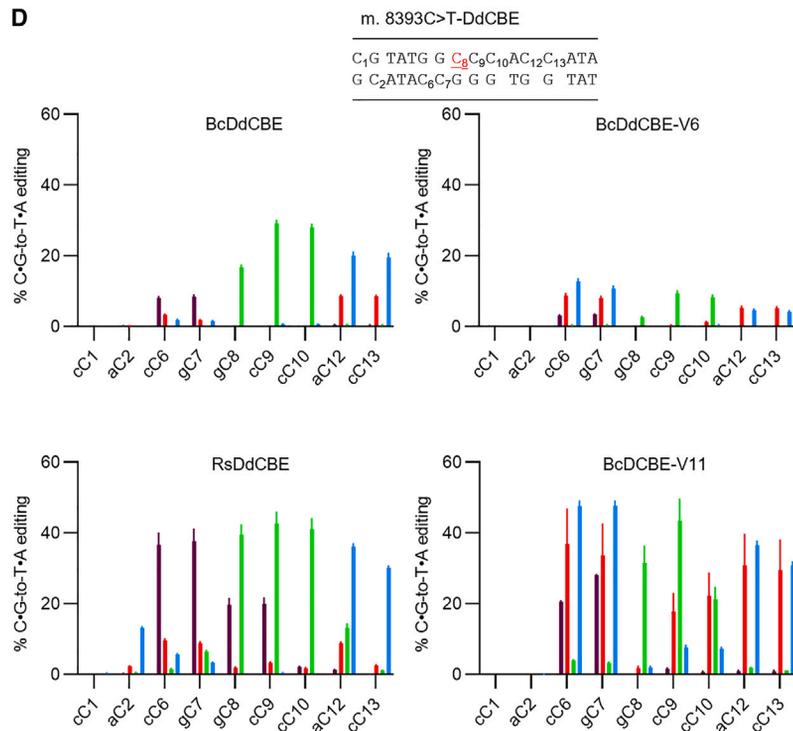
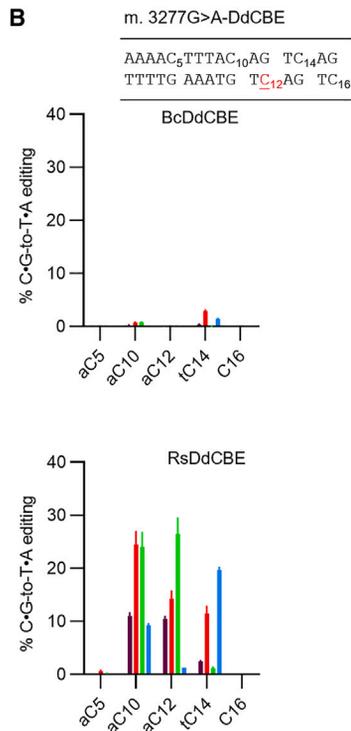
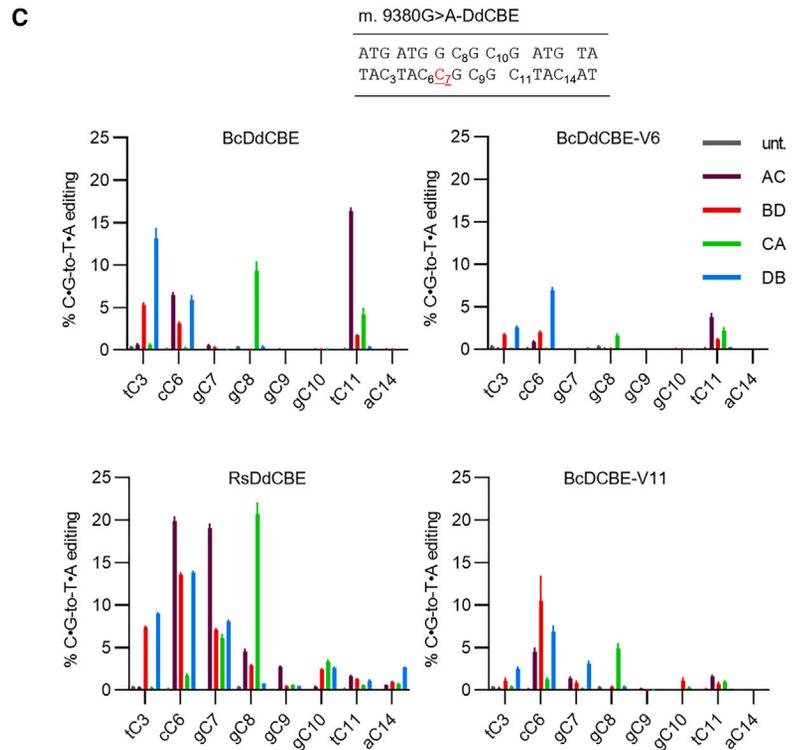
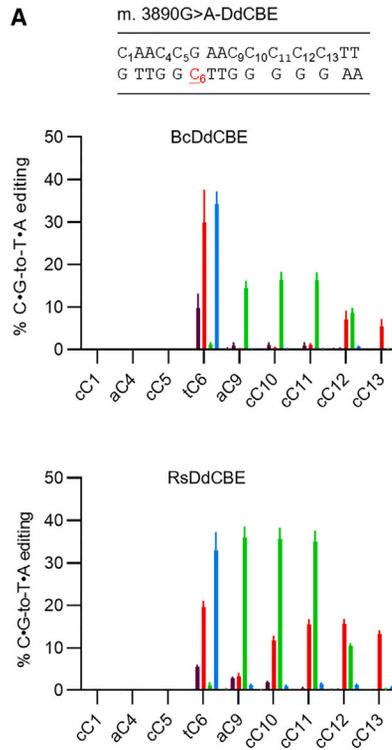
Figure 3. Characterizing the editing window of RsDdCBE

(A and B) Heatmap of editing efficiencies on \underline{tC} and \underline{aC} (A) or \underline{gC} (B) targets mediated by RsDdCBE with spacing regions ranging from 10 to 20 bp. Numbers refer to the positions of cytosines in the spacing region, counting the DNA nucleotide immediately after the binding site of TALE-L as position 1. AC, BD, CA, and DB refer to different splits and orientations of DdCBE listed in Figure 1D. Data are mean from three independent biological replicates.

(0.098%), m.3684C>T-RsDdCBE (0.045%), m. 4735C>T-RsDdCBE (0.059%), and m. 8393C>T-RsDdCBE (0.223%) were higher than those of the canonical BsDdCBEs (0.018%–0.068%) (Figures 5A and 5C). This increase could be ascribed to high activity and expanded targeting compatibility of RsDdCBE. The frequencies of off-target editing per on-target editing event of RsDdCBEs are comparable to or even significantly lower than that of canonical BcDdCBE and its variants (Figures 5B and 5D). We also analyzed the distribution pattern of the off-target single-nucleotide variants (SNVs) caused by RsDdCBE; 121, 30, 28, and 346 SNVs with more than 1% editing were detected from cells treated with m. 3277G>A-RsDdCBE, m. 3684C>A-RsDdCBE, m. 4735C>T-RsDdCBE, and m. 8393C>T-RsDdCBE, respectively. Assays on plasmids and endogenous sites of mtDNA showed RsDdCBE offers NC sequence compatibility, however off-target editing of RsDdCBEs mainly happened at \underline{tC} and \underline{aC} contexts (Figure 5E). Consistently, the alignment of sequences flanking the off-target cytosine for these four RsDdCBEs revealed a strong preference for \underline{tC} and \underline{aC} contexts (Figure 5F). Among the off-target sites, only a small number of these SNVs are shared by two RsDdCBEs and none of them are shared by the four RsDdCBEs (Figure 5G), indi-

cating that the TALE-binding-independent deaminase activity is limited. This is in line with the observation that most off-target events were centered on the on-target sites (Figure S10), which could be mediated by spontaneous assembly of DdCBE on the half-sites bound by one TALE array.

It has been reported that the mitochondrial base editor induces extensive off-target editing in the nuclear genome and such off-target effects could be alleviated through incorporating nuclear export signal (NES) sequences into the DdCBE architectures.^{23,24} To assess the nuclear off-target activity of RsDdCBE, we transfected three pairs of mitochondrial targeting RsDdCBE (m.8393C>T-RsDdCBE, m.9380G>A-RsDdCBE, and m.13513G>A-RsDdCBE) that also have potential target sequence(s) present in the nuclear genome (Chr2:88124468-88124514, Chr1:569905-569954, and Chr5:134260949-134260995) into HEK293 cells and analyzed the C•G-to-T•A conversion on respective nuclear sequence(s) (Figure S11). Sequencing results showed RsDdCBE catalyzed the cytosine conversion in nucleus with the frequency range from 0.18% to 37.25% (Figure S9). To examine whether inhibiting the import of RsDdCBE into the nucleus could ameliorate



(legend on next page)

the nuclear off-target activity, an NES was fused to the C terminus of the TALE array. NES-RsDdCBE displayed greatly reduced nuclear off-target activity to 0.057%–9.18% (Figure S11).

These results collectively indicate that, although mtDNA off-target editing increases for RsDdCBE, which could be attributed to its expanded targeting scope, the ratios of off-target to on-target editing for RsDdCBE are comparable to or even lower than that of canonical BsDdCBE and its variants.

DISCUSSION

Canonical DdCBE and its variants enable introduction of precise change within mtDNA,^{9,15} but targets are limited to 5'-HC contexts; this restraint prevents this technology from reaching its potential. To address this challenge, we tapped uncharacterized DddA homologs to identify proteins with wide target compatibility. We found DdCBE derived from *Ruminococcus* sp. DddA offers an NC sequence compatibility for C•G-to-T•A conversion and displays high editing efficiency at previously inaccessible (especially gC) targets. With a similar strategy, Mi et al. recently identified a DddA homolog from *S. sunii* (Ddd_Ss) and engineered it into cytosine base editors (DdCBE_Ss), which are also able to introduce mutations in gC context.²² Both studies highlighted the diversity of interbacterial DNA deaminase toxin systems and the potential to enrich the genome editing toolbox by mining of metagenomic sequences.

The models of RsDdCBE and BcDdCBE generated with AlphaFold 2 showed that the most prominent disparity on these two structures is that the one β sheet is absent in RsDdCBE, and this divergence may contribute to their different target compatibility. The co-crystal structure of DdCBEs bound to target DNA would help to reveal how the target compatibility was dictated and lay the groundwork for rational design of DdCBEs with desired features. Our data from plasmid editing assays showed RsDdCBE has variant editing windows for targets with different context. The pattern discovered in the current study could direct the design of DdCBE to achieve high on-target editing and minimize unwanted bystander editing. With the compatibility in NC contexts, RsDdCBE mediates bystander editing at slightly higher frequencies compared to canonical DdCBE and its variants. To minimize bystander base editing and off-target activity, it would be valuable to develop context-specific RsDddA variants and further improve the design of the editors. Recently, researchers demonstrated A•T-to-G•C conversion in mitochondrial DNA with a DdCBE-centered design,²⁵ and it is worthwhile to examine whether the NC compatibility of RsDdCBE could be transferred to an adenine base editor.

MATERIALS AND METHODS

Computational identification of potential DddA_{tox}

By conducting homology search using the amino acid sequences of BcDddA_{tox} (6U08_A *Burkholderia cenocepacia*), a total of 142 homologs were retrieved from the NCBI database. The histidine-valine-glutamic acid (HVE) motif was reported to be essential for deaminase activity, and the substitution of glutamic acid with alanine leads to a catalytically inactive BcDddA_{tox}.⁹ After excluding 20 homologs without HVE motif, the phylogenetic tree of the remaining 122 homologs was constructed by MEGA.¹¹ Four representative candidates from different genera, RGG70441.1 (*Ruminococcus* sp. AF17-6), WP_216577045.1 (*Falcatimonas* sp. MSJ-15), WP_189594293.1 (*Streptomyces massasporeus*), and MBV9870847.1 (a Frankiaceae bacterium), were selected for DNA synthesis. The amino acid sequences of the four homologs are listed in Table S1.

Plasmid construction

The sequence coding the potential toxin domain of candidates was synthesized as split halves, then the halves were amplified and cloned into a prokaryotic ORF by Gibson assembly using a ClonExpress MultiS one-step cloning kit (Vazyme). The primers are listed in Table S2.

All DdCBE vectors were assembled using a DdCBE assembly kit developed by our lab, which can be obtained from Addgene (kit #1000000212). Plasmids containing two split types of RsDddA have been deposited to Addgene (Addgene: 205463, 205464, 205465, and 205466). Briefly, the expression backbone and repeat variable di-residues (RVD) plasmids were digested with Bsa I (NEB) and ligated with T4 DNA ligase (NEB) simultaneously in single tubes using the following program: 37° C, 10 min; 10 cycles of 10 min at 37° C and 10 min at 16° C; 5 min at 50° C; and 5 min at 80° C. The TALE-binding sequences and windows information of sites are listed in Table S3.

The pEGFP-N1 vector was used as a PCR template to construct plasmids bearing TALE-binding sites and different combination of nC targets within spacers. The primers are listed in Table S2.

The information for all plasmids generated in this study is listed in Table S4.

Mammalian cell culture and transfection

HEK293FT cells were cultured in DMEM supplemented with 10% FBS (Gemini) and 1% penicillin/streptomycin (Thermo) and maintained at 37° C with 5% CO₂. The 150,000 cells were co-transfected with 400 ng of left and right DdCBE each using the Lonza 4D-Nucleofector according to the manufacturer's manual. The nucleofected cells were treated with 1.5 μ g/mL puromycin 24 h post nucleofection and collected at day 4 for DNA extraction.

Figure 4. RsDdCBE achieve high-efficiency editing at previously inaccessible pathogenic mtDNA loci

(A–D) Mitochondrial base editing efficiencies on m.3890G>A (ND1), m.3277G>A (TRNL1), m.9380G>A (COX3), and m.8393C>T (ATP8) mediated by RsDdCBE, BcDdCBE, BcDdCBE variant 6, or BcDdCBE variant 11. Subscripted numbers refer to the positions of cytosines in the spacing region, counting the DNA nucleotide immediately after the binding site of TALE-L as position 1. All tests were carried out in HEK293FT cells. AC, BD, CA, and DB refer to different splits and orientations of DdCBE listed in Figure 1D. Untreated cells were used as controls. Data are mean \pm SD from three independent biological replicates.

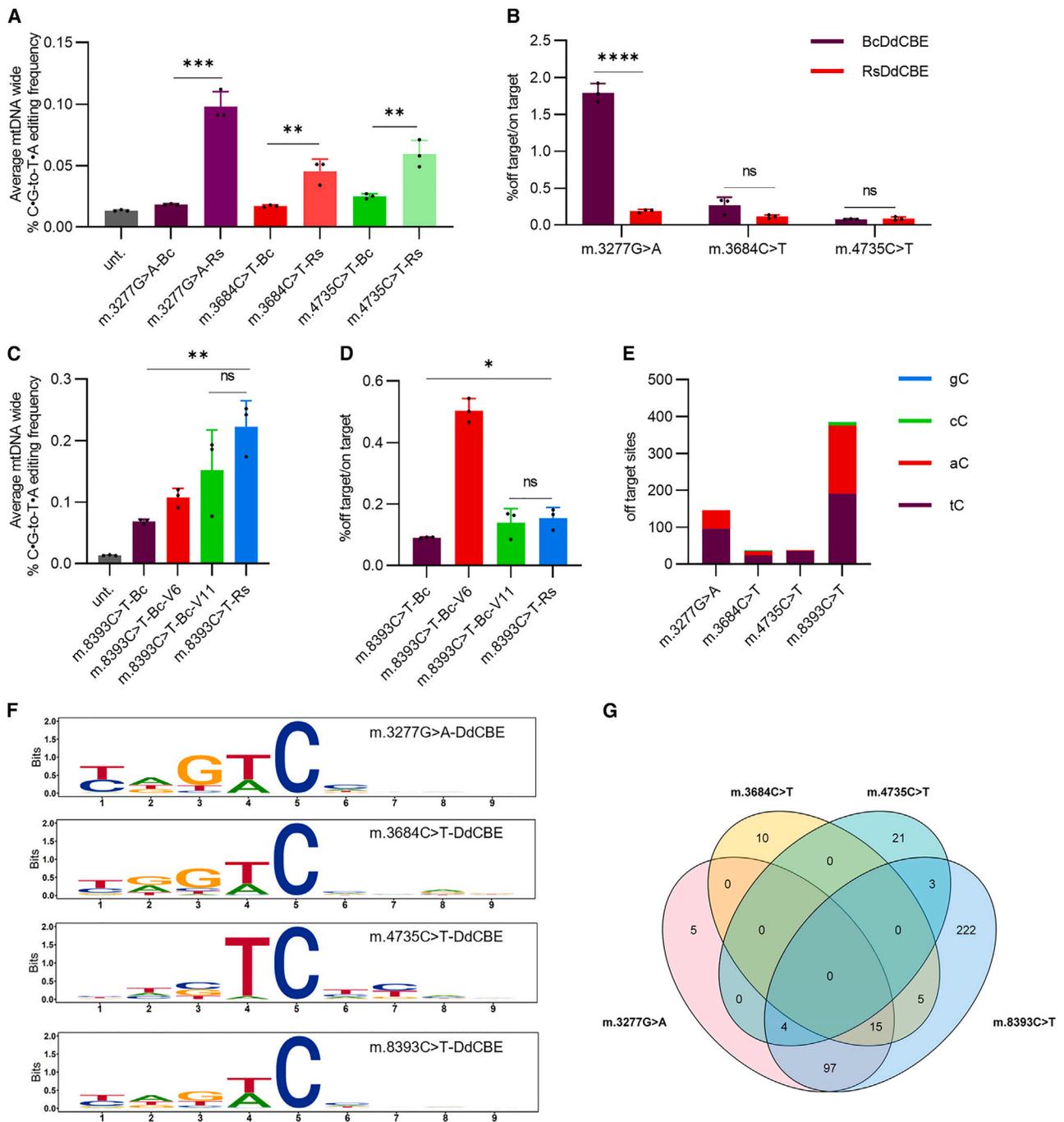


Figure 5. Mitochondrial genome-wide off-target analysis of RsDdCBE

(A) mtDNA genome-wide C•G-to-T•A editing frequency mediated by RsDdCBE and canonical BcDdCBE targeting m.3277G>A, m.3684C>T or m.4735C>T site. (B) The frequencies of off-target editing per on-target editing event of RsDdCBE and BsDdCBE targeting m.3277G>T, m.3684C>T or m.4735C>T site. (C) mtDNA genome-wide C•G-to-T•A editing frequency mediated by RsDdCBE and canonical BsDdCBE and BcDdCBE variants targeting mt.8393C site. (D) The frequencies of off-target editing per on-target editing event of RsDdCBE and canonical BsDdCBE and BcDdCBE variants targeting mt.8393C>T site. (E) Distribution of off-target editing of RsDdCBE pairs on different sequence contexts. (F) Sequence logos generated from motif analysis of off-target sites of RsDdCBEs targeting m.3277G>A, m.3684C>T, m.4735C>T and m.8393C>T. The off-target sites showing editing frequency higher than 1% were included. (G) Overlaps between off-target sites produced by RsDdCBEs targeting m.3277G>A, m.3684C>T, m.4735C>T and m.8393C>T. (A–D) Data are mean \pm SD from three independent biological replicates. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$; NS, not significant ($p > 0.05$) by Student's unpaired two-tailed t test.

To explore the editing window of DdCBE, HEK293FT cells were co-transfected with 100 ng of plasmid bearing nC targets and 800 ng of NLS-left-DdCBE and -right-DdCBE using ExFect Transfection Reagent (Vazyme). The cells were treated with 1.5 $\mu\text{g}/\text{mL}$ puromycin 24 h post transfection and collected at day 4 for DNA extraction.

DNA extraction

HEK293FT cells were lysed with QuickExtract DNA Extraction Solution and incubated at 65°C for 40 min, followed by 98°C for 2 min. The targeted loci were amplified by PCR with specific primers. The PCR products were sequenced by Sanger sequencing. The PCR primers are listed in [Table S1](#).

Deep sequencing and data analysis

Genomic regions of interest were first amplified with barcoded primers (first-round PCR [PCR1]) using Phanta Max Super-Fidelity DNA Polymerase (Vazyme). The PCR1 products were pooled with equal moles and purified for the second-round PCR (PCR2). PCR2 was performed using index primers (Vazyme) and purified by DNA Clean Beads for sequencing using the Illumina NovaSeq platform. Primers for PCR1 are listed in [Table S1](#).

For data analysis, the trimmed reads were aligned to the mitochondrial genome using bowtie2 with default parameters. The alignment results were converted to bam format by SAMtools. SAMtools mpileup was used to detect C-to-T or G-to-A conversion. Sites with conversion rate $\geq 1\%$ in any untreated sample were identified as SNPs and excluded from further analysis.

Whole-mtDNA sequencing and data analysis

Two overlapping fragments around 8.5 kb each were amplified by long-range PCR and purified by gel extraction. The two fragments were pooled with equal amounts and subjected to library preparation using TruePrep DNA Library Prep Kit V2 for Illumina (Vazyme) according to the manufacturer's instruction. The libraries were purified using DNA Clean Beads by $0.5\times/0.15\times$ double size selection. Libraries were pooled and sequenced by the Illumina NovaSeq platform. Primers used in long-range PCR are listed in [Table S1](#). For data analysis, the trimmed reads were aligned to the human mitochondrial genome (NC_012920.1) using bowtie2 with default parameters. The alignment results were converted to bam format by SAMtools. SAMtools mpileup was used to detect C-to-T or G-to-A conversion.

The following sites were excluded before analysis: (1) the sites with C•G-to-T•A variation over 1% in any untreated sample, (2) the sites with C/G-to-T/A variation over 90% in any sample, (3) sites within the DdCBE spacing region. The average off-target editing frequency was then calculated independently for each biological replicate of each treatment condition as the sum of events of C/G-to-T/A conversion divided by the total coverage of these sites.

DATA AND CODE AVAILABILITY

The high-throughput sequencing data generated in this study have been deposited to the NCBI Sequence Read Archive database (accession code, PRJNA907170). The shared URL for review is <http://www.ncbi.nlm.nih.gov/bioproject/907170>.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.omtn.2023.09.005>.

ACKNOWLEDGMENTS

This work was supported by the National Key R&D Program of China (2021YFC2700600 to B.S.), the Funds for Creative Research Groups of China (82221005 to B.S.), the National Natural Science Foundation of China (31970796 to B.S., 31970765 to X.L.), the Natural Science Foundation of the Jiangsu Higher Education Institutions of China (21KJA180005 to F.Z.).

AUTHOR CONTRIBUTIONS

B.S., X.L., and J.Z. conceived the project and designed the experiments. K.C. and C.L. performed all plasmid construction, sequencing, and cell culture work with the help of X.Q., J. G., L.S., and Y.D. J.J. performed all bioinformatics analyses with the help of Y.G. and Z.L. J.T. performed the structure prediction of DddA homology using AphaFold2. B.S., X.L., and K.C. wrote the manuscript with input from all authors.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

- Gorman, G.S., Chinnery, P.F., DiMauro, S., Hirano, M., Koga, Y., McFarland, R., Suomalainen, A., Thorburn, D.R., Zeviani, M., and Turnbull, D.M. (2016). Mitochondrial diseases. *Nat. Rev. Dis. Prim.* 2, 16080.
- Gorman, G.S., Schaefer, A.M., Ng, Y., Gomez, N., Blakely, E.L., Alston, C.L., Feeney, C., Horvath, R., Yu-Wai-Man, P., Chinnery, P.F., et al. (2015). Prevalence of nuclear and mitochondrial DNA mutations related to adult mitochondrial disease. *Ann. Neurol.* 77, 753–759.
- Ng, Y.S., and Turnbull, D.M. (2016). Mitochondrial disease: genetics and management. *J. Neurol.* 263, 179–191.
- Silva-Pinheiro, P., and Minczuk, M. (2022). The potential of mitochondrial genome engineering. *Nat. Rev. Genet.* 23, 199–214.
- Bacman, S.R., Kauppila, J.H.K., Pereira, C.V., Nissanka, N., Miranda, M., Pinto, M., Williams, S.L., Larsson, N.G., Stewart, J.B., and Moraes, C.T. (2018). MitoTALEN reduces mutant mtDNA load and restores tRNA(Ala) levels in a mouse model of heteroplasmic mtDNA mutation. *Nat. Med.* 24, 1696–1700.
- Gammage, P.A., Rorbach, J., Vincent, A.I., Rebar, E.J., and Minczuk, M. (2014). Mitochondrially targeted ZFNs for selective degradation of pathogenic mitochondrial genomes bearing large-scale deletions or point mutations. *EMBO Mol. Med.* 6, 458–466.
- Jackson, C.B., Turnbull, D.M., Minczuk, M., and Gammage, P.A. (2020). Therapeutic Manipulation of mtDNA Heteroplasmy: A Shifting Perspective. *Trends Mol. Med.* 26, 698–709.
- Gammage, P.A., Moraes, C.T., and Minczuk, M. (2018). Mitochondrial Genome Engineering: The Revolution May Not Be CRISPR-ized. *Trends Genet.* 34, 101–110.

9. Mok, B.Y., de Moraes, M.H., Zeng, J., Bosch, D.E., Kotrys, A.V., Raguram, A., Hsu, F., Radey, M.C., Peterson, S.B., Mootha, V.K., et al. (2020). A bacterial cytidine deaminase toxin enables CRISPR-free mitochondrial base editing. *Nature* 583, 631–637.
10. Guo, J., Chen, X., Liu, Z., Sun, H., Zhou, Y., Dai, Y., Ma, Y., He, L., Qian, X., Wang, J., et al. (2022). DdCBE mediates efficient and inheritable modifications in mouse mitochondrial genome. *Mol. Ther. Nucleic Acids* 27, 73–80.
11. Guo, J., Zhang, X., Chen, X., Sun, H., Dai, Y., Wang, J., Qian, X., Tan, L., Lou, X., and Shen, B. (2021). Precision modeling of mitochondrial diseases in zebrafish via DdCBE-mediated mtDNA base editing. *Cell Discov.* 7, 78.
12. Lee, H., Lee, S., Baek, G., Kim, A., Kang, B.C., Seo, H., and Kim, J.S. (2021). Mitochondrial DNA editing in mice with DddA-TALE fusion deaminases. *Nat. Commun.* 12, 1190.
13. Wei, Y., Xu, C., Feng, H., Xu, K., Li, Z., Hu, J., Zhou, L., Wei, Y., Zuo, Z., Zuo, E., et al. (2022). Human cleaving embryos enable efficient mitochondrial base-editing with DdCBE. *Cell Discov.* 8, 7.
14. Kang, B.C., Bae, S.J., Lee, S., Lee, J.S., Kim, A., Lee, H., Baek, G., Seo, H., Kim, J., and Kim, J.S. (2021). Chloroplast and mitochondrial DNA editing in plants. *Nat. Plants* 7, 899–905.
15. Mok, B.Y., Kotrys, A.V., Raguram, A., Huang, T.P., Mootha, V.K., and Liu, D.R. (2022). CRISPR-free base editors with enhanced activity and expanded targeting scope in mitochondrial and nuclear DNA. *Nat. Biotechnol.* 40, 1378–1387.
16. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Židek, A., Potapenko, A., et al. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature* 596, 583–589.
17. Caporali, L., Ghelli, A.M., Iommarini, L., Maresca, A., Valentino, M.L., La Morgia, C., Liguori, R., Zanna, C., Barboni, P., De Nardo, V., et al. (2013). Cybrid studies establish the causal link between the mtDNA m.3890G>A/MT-ND1 mutation and optic atrophy with bilateral brainstem lesions. *Biochim. Biophys. Acta* 1832, 445–452.
18. Vacchiano, V., Caporali, L., La Morgia, C., Carbonelli, M., Amore, G., Bartolomei, I., Cascavilla, M.L., Barboni, P., Lamperti, C., Catania, A., et al. (2021). The m.3890G>A/MT-ND1 mtDNA rare pathogenic variant: Expanding clinical and MRI phenotypes. *Mitochondrion* 60, 142–149.
19. Zhu, H.Y., Wang, S.W., Liu, L., Chen, R., Wang, L., Gong, X.L., and Zhang, M.L. (2009). Genetic variants in mitochondrial tRNA genes are associated with essential hypertension in a Chinese Han population. *Clin. Chim. Acta* 410, 64–69.
20. Ridge, P.G., Maxwell, T.J., Corcoran, C.D., Norton, M.C., Tschanz, J.T., O'Brien, E., Kerber, R.A., Cawthon, R.M., Munger, R.G., and Kawwe, J.S.K. (2012). Mitochondrial genomic analysis of late onset Alzheimer's disease reveals protective haplogroups H6A1A/H6A1B: the Cache County Study on Memory in Aging. *PLoS One* 7, e45134.
21. Galimberti, C.A., Diegoli, M., Sartori, I., Uggetti, C., Brega, A., Tartara, A., and Arbustini, E. (2006). Brain pseudotrophy and mental regression on valproate and a mitochondrial DNA mutation. *Neurology* 67, 1715–1717.
22. Mi, L., Shi, M., Li, Y.X., Xie, G., Rao, X., Wu, D., Cheng, A., Niu, M., Xu, F., Yu, Y., et al. (2023). DddA homolog search and engineering expand sequence compatibility of mitochondrial base editing. *Nat. Commun.* 14, 874.
23. Lee, S., Lee, H., Baek, G., Namgung, E., Park, J.M., Kim, S., Hong, S., and Kim, J.S. (2022). Enhanced mitochondrial DNA editing in mice using nuclear-exported TALE-linked deaminases and nucleases. *Genome Biol.* 23, 211.
24. Lei, Z., Meng, H., Liu, L., Zhao, H., Rao, X., Yan, Y., Wu, H., Liu, M., He, A., and Yi, C. (2022). Mitochondrial base editor induces substantial nuclear off-target mutations. *Nature* 606, 804–811.
25. Cho, S.I., Lee, S., Mok, Y.G., Lim, K., Lee, J., Lee, J.M., Chung, E., and Kim, J.S. (2022). Targeted A-to-G base editing in human mitochondrial DNA with programmable deaminases. *Cell* 185, 1764–1776.e12.