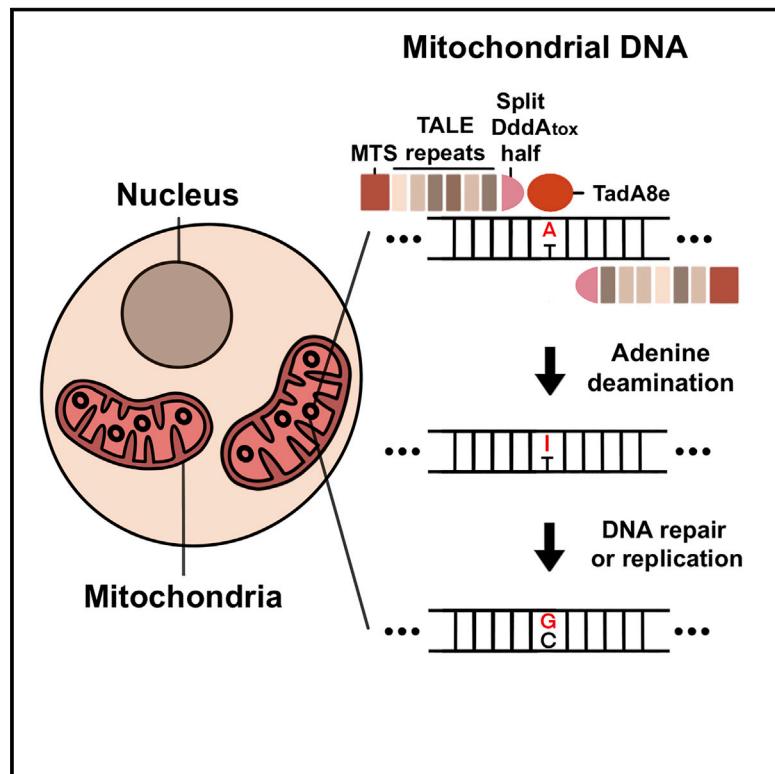


Targeted A-to-G base editing in human mitochondrial DNA with programmable deaminases

Graphical abstract



Authors

Sung-Ik Cho, Seonghyun Lee,
Young Geun Mok, ..., Ji Min Lee,
Eugene Chung, Jin-Soo Kim

Correspondence

jskim01@snu.ac.kr

In brief

Development of a programmable mitochondrial genome editing tool that enables efficient and accurate A-to-G conversions at diverse target sites in human mitochondrial genome.

Highlights

- TALE-linked adenine deaminases enable A-to-G editing in human mitochondria
- DddA variants function in *cis* or in *trans* with TadA8e to catalyze A-to-G conversions
- Human cells with homoplasmic mtDNA mutations can be obtained via drug selection



Resource

Targeted A-to-G base editing in human mitochondrial DNA with programmable deaminases

Sung-Ik Cho,^{1,2} Seonghyun Lee,¹ Young Geun Mok,¹ Kayeong Lim,¹ Jaesuk Lee,^{1,2} Ji Min Lee,^{1,2} Eugene Chung,^{1,2} and Jin-Soo Kim^{1,3,*}

¹Center for Genome Engineering, Institute for Basic Science, Daejeon 34126, Republic of Korea

²Department of Chemistry, Seoul National University, Seoul 08826, Republic of Korea

³Lead contact

*Correspondence: jskim01@snu.ac.kr

<https://doi.org/10.1016/j.cell.2022.03.039>

SUMMARY

Mitochondrial DNA (mtDNA) editing paves the way for disease modeling of mitochondrial genetic disorders in cell lines and animals and also for the treatment of these diseases in the future. Bacterial cytidine deaminase *DddA*-derived cytosine base editors (DdCBEs) enabling mtDNA editing, however, are largely limited to C-to-T conversions in the 5'-TC context (e.g., TC-to-TT conversions), suitable for generating merely 1/8 of all possible transition (purine-to-purine and pyrimidine-to-pyrimidine) mutations. Here, we present transcription-activator-like effector (TALE)-linked deaminases (TALEDs), composed of custom-designed TALE DNA-binding arrays, a catalytically impaired, full-length *DddA* variant or split *DddA* originated from *Burkholderia cenocepacia*, and an engineered deoxyadenosine deaminase derived from the *E. coli* *TadA* protein, which induce targeted A-to-G editing in human mitochondria. Custom-designed TALEDs were highly efficient in human cells, catalyzing A-to-G conversions at a total of 17 target sites in various mitochondrial genes with editing frequencies of up to 49%.

INTRODUCTION

Various mutations in human mtDNA are associated not only with maternally heritable genetic diseases, collectively affecting at least one in 5,000 individuals (Gorman et al., 2015; Silva-Pinheiro and Minczuk, 2022), but also with aging and age-related diseases including cancer and diabetes (Taylor and Turnbull, 2005). Among a total of 95 clinically confirmed pathogenic mtDNA mutations, listed in mitomap (www.mitomap.org), point mutations account for the vast majority (90/95 = 94.7%), which, in principle, can be corrected using base editors enabling single-nucleotide conversions. Targeted base editing in mtDNA is a promising approach for the treatment of these mitochondrial genetic diseases and also for the production of disease models in cell lines and animals but has been hampered by the lack of appropriate tools and methods. Engineered nucleases, including zinc-finger nucleases (ZFNs) and transcriptional-activator-like effector nucleases (TALENs), can be designed to cleave mutant mtDNA selectively and thereby to reduce mtDNA levels in a mixture of wild-type and mutant mtDNA molecules (a.k.a. heteroplasmy) (Bacman et al., 2018; Gammie et al., 2016; Gammie et al., 2014; Gammie et al., 2018b; Kazama et al., 2019; Reddy et al., 2015) but cannot be used for correcting homoplasmic mutations or inducing *de novo* mutations in mtDNA.

Clustered regularly interspaced short palindromic repeats (CRISPRs) RNA-guided base editors (Gaudelli et al., 2017; Komer et al., 2016; Li et al., 2018; Nishida et al., 2016), widely used for targeted single-nucleotide conversions in nuclear DNA, are not suitable for mtDNA editing because guide RNAs cannot be delivered to mitochondria (Gammie et al., 2018a). Recently developed DdCBEs, composed of the split interbacterial toxin *DddA*_{tox} derived from *Burkholderia cenocepacia*, custom-designed DNA-binding transcription-activator-like effector (TALE) arrays, and a uracil glycosylase inhibitor (UGI), allow researchers to edit mtDNA in cell lines (Mok et al., 2020), animals (Guo et al., 2021; Guo et al., 2022; Lee et al., 2021), and plants (Kang et al., 2021) but are largely limited to cytosine editing in the 5'-TC context. As a result, only 9 (=10%) out of 90 confirmed pathogenic point mutations can be corrected by currently available DdCBEs (Table S1). Base editors enabling targeted A-to-G conversions in human mtDNA could correct 39 (=43%) out of these 90 pathogenic mutations, including those causing Leber hereditary optic neuropathy (LHON), mitochondrial encephalomyopathy, lactic acidosis, stroke-like episodes (MELAS), and Leigh syndrome. Here, we present programmable TALE-linked deaminases, composed of custom-designed TALE proteins, a split or catalytically deficient *DddA* variant, and an engineered adenine deaminase that catalyzes the hydrolytic



deamination of adenine to yield inosine, which is paired with cytosine during replication, for targeted conversions of A:T base pairs to G:C pairs (that is, A-to-G conversions) in human mtDNA.

RESULTS

TALE-fused adenine deaminases inducing A-to-G editing in mtDNA

In an effort to create adenine base editors enabling A-to-G substitutions in organelle DNA, we first investigated whether fusion proteins containing a TALE protein custom designed to bind to the *ND1* or *ND4* gene in human mitochondria, a deoxyadenosine deaminase variant, termed TadA8e, engineered from *E. coli* TadA (Richter et al., 2020) (shown as adenine deaminase [AD] in the figures), and a mitochondrial targeting sequence (MTS) could catalyze A-to-G conversions in human embryonic kidney 293T (HEK 293T) cells (Figures 1 and S1). Targeted deep sequencing showed that the fusion proteins were poorly active, inducing A-to-G conversions in the immediate proximity of TALE-binding sites with frequencies that ranged from 0.7% to 1.2% (indicated by the arrows in Figure 1B), which were impractically low but clearly above noise levels caused by sequencing errors (~0.4%) (Figures 1B and S1). This result suggests that, although the TadA variant is known to operate on single-stranded DNA (ssDNA) rather than double-stranded DNA (dsDNA), it can catalyze adenine deamination in dsDNA, when fused to a TALE protein bound to a target DNA site, albeit much less efficiently than in ssDNA, reminiscent of zinc-finger-protein-fused cytidine deaminases poorly catalyzing C-to-T conversions in a reporter gene stably integrated into the human genome (Yang et al., 2016).

TALE deaminases enabling concurrent A-to-G and C-to-T editing

Encouraged by this result, we sought to enhance the editing efficiency of TALE-linked deaminases by fusing the TadA variant to pre-characterized DdCBEs containing DddA_{tox}, an enzymatic moiety responsible for cytosine deamination in the interbacterial toxin DddA. We reasoned that the addition of DddA_{tox}, which operates on dsDNA, might make DNA more accessible to the TadA variant. DdCBEs are composed of a pair of TALE arrays (shown as L [left TALE] or R [right TALE]), each fused to a catalytically deficient, split DddA_{tox} half (shown as 1397N and 1397C [N- and C-terminal DddA_{tox} half, respectively, split at G1397] in Figures 2 and S2), and UGI, which suppresses the removal of uracil, a product of cytosine deamination, in DNA by endogenous uracil glycosylase. Note that DddA_{tox} is split to avoid cytotoxicity of the full-length DddA_{tox} protein. We replaced UGI with the TadA variant in a pre-characterized *ND1*-specific DdCBE pair to create a chimeric subunit containing the TadA variant and the C-terminal DddA_{tox} half split at G1397 (L-1397C-AD or R-1397C-AD). Interestingly, when paired with the other TALE fusions containing 1397N and UGI (R-1397N-UGI or L-1397N-UGI, respectively), these chimeric subunits catalyzed A-to-G (= T-to-C in the other DNA strand) substitutions in the spacer region of 16 base pairs (bps) in length between the two TALE-binding sites with editing frequencies of up to 19% in HEK 293T cells (Figures 2B–2D).

Small insertions and deletions (indels) or A-to-C or T conversions were rarely induced at the target site. In addition, C-to-T (= G-to-A) conversions were also induced by these pairs with comparable efficiencies (up to 14%). In contrast, the two original DdCBE pairs (L-1397N-UGI + R-1397C-UGI and L-1397C-UGI + R-1397N-UGI) induced C-to-T conversions exclusively without A-to-G conversions (Figure 2B). This result showed that composite TALE deaminases created by fusing the TadA variant to a split DddA_{tox} half in DdCBEs could induce simultaneous A-to-G and C-to-T edits in human mtDNA, which would be useful for random mutagenesis or composite editing, as shown by CRISPR RNA-guided base editors for nuclear DNA editing (Zhang et al., 2020; Grünwald et al., 2020; Sakata et al., 2020; Li et al., 2020).

TALE deaminases catalyzing A-to-G but not C-to-T conversions

We next sought to create different types of base editors that would catalyze A-to-G conversions exclusively without C-to-T conversions in mtDNA. We reasoned that a composite TALE deaminase pair with no UGI in either subunit would avoid C-to-T substitutions while retaining A-to-G editing activity (Figure 3). Interestingly, two UGI-free TALE deaminase pairs (L-1397C-AD + R-1397N and L-1397N + R-1397C-AD) targeted to the *ND1* site induced A-to-G conversions with an editing frequency of 49% or 40% (Figure 3B), much more efficiently than UGI-containing TALE deaminase pairs, without causing C-to-T substitutions (<0.2%) or unwanted indels (<0.3%). We also constructed UGI-free TALE deaminases targeted to the *ND4* gene and found that these adenine base editors achieved A-to-G conversions with high frequencies of up to 34% without causing unwanted C-to-T edits or indels (Figure 3C). In addition, we found that UGI-free TALE deaminases in which DddA_{tox} was split at G1333 rather than G1397 were also highly efficient, catalyzing A-to-G edits with frequencies that ranged from 10% to 33% (Figure 3B). Adenines and thymines in the middle of the spacer region were edited more efficiently (31%~46%) than those positioned toward the edge (<10%) near the TALE-binding site (Figures 3D and 3E). The product purity, defined as the percentage of sequencing reads with the target A converted to G among those with the target A converted to C, G, or T, ranged from 93% to 99%. These results showed that the TadA-derived deoxyadenosine deaminase fused with a TALE protein and a split DddA_{tox} system could deaminate adenine in dsDNA, catalyzing A-to-G substitutions efficiently in human mitochondria.

Monomeric and dimeric TALE deaminases enabling A-to-G conversions

We next investigated whether the catalytically deficient, non-toxic, full-length DddA_{tox} variant with an active-site E1347A mutation, rather than the G1333 or G1397 split system, could still be used for boosting the A-to-G editing activity of TALE-AD fusion deaminases. Two types of TALE deaminases (hereinafter termed TALEDs) containing the E1347A DddA_{tox} variant were constructed: monomeric and dimeric TALEDs. Monomeric TALEDs (mTALEDs) were composed of a single TALE array fused to the TadA variant and the full-length E1347A DddA_{tox}, resulting in the TALE-AD-E1347A architecture, which was more efficient

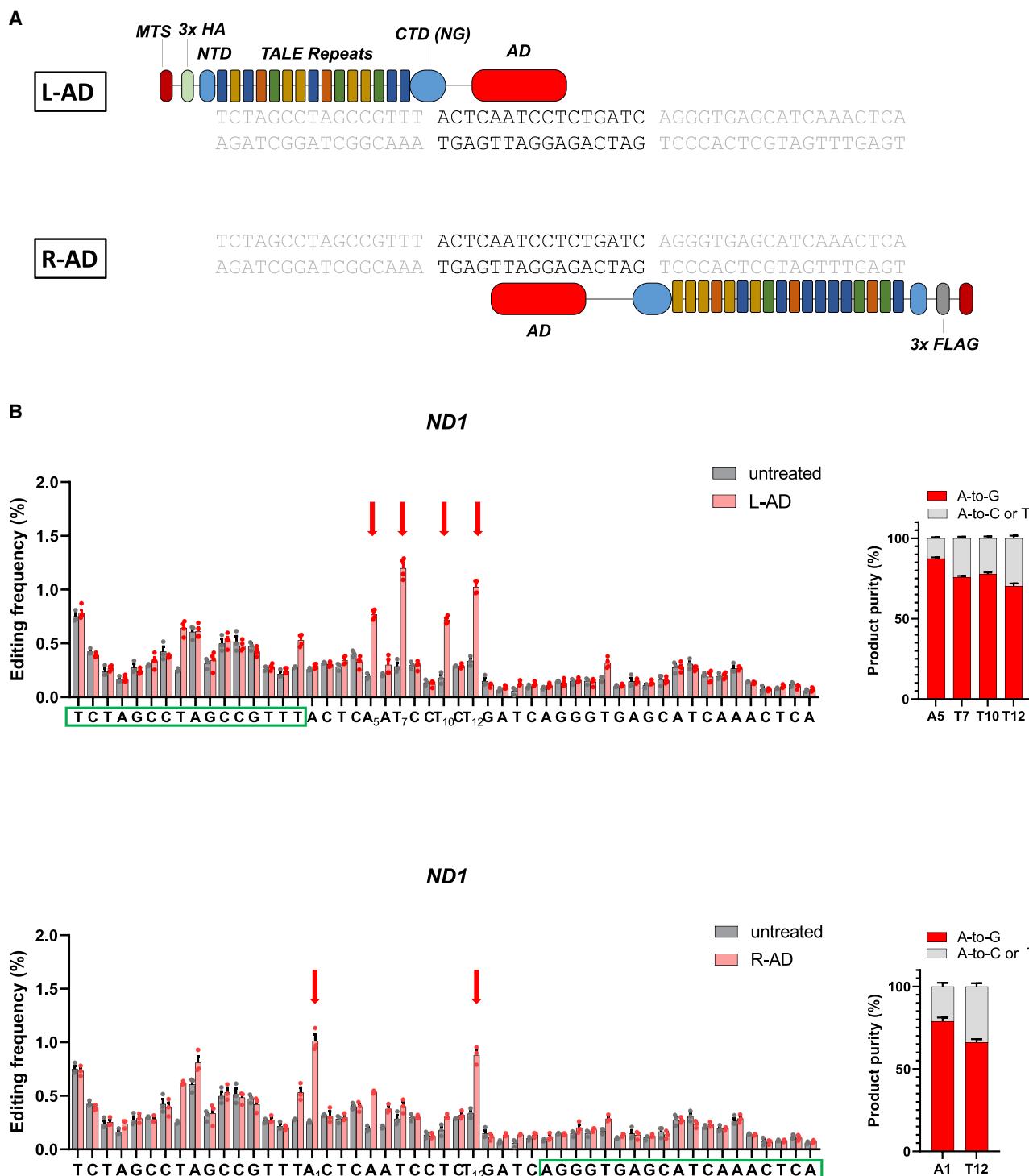


Figure 1. TALE-fused adenine deaminases inducing A-to-G editing in human mtDNA

(A) TadA-derived adenine deaminase (AD) fused to left and right TALE proteins (L-AD and R-AD, respectively) targeted to the *MT-ND1* gene. MTS, mitochondrial targeting sequence; NTD, N-terminal domain; CTD (NG), C-terminal domain with NG repeat variable di-residues.

(B) Base-editing frequencies and product purities of TALE-linked adenine deaminases at the *ND1* site. TALE-binding sequences are outlined with green boxes. Arrows indicate the positions of A-to-G conversions. Error bars represent SEM for three (untreated) and four (L-AD or R-AD) independent biological replicates. See also Figure S1.

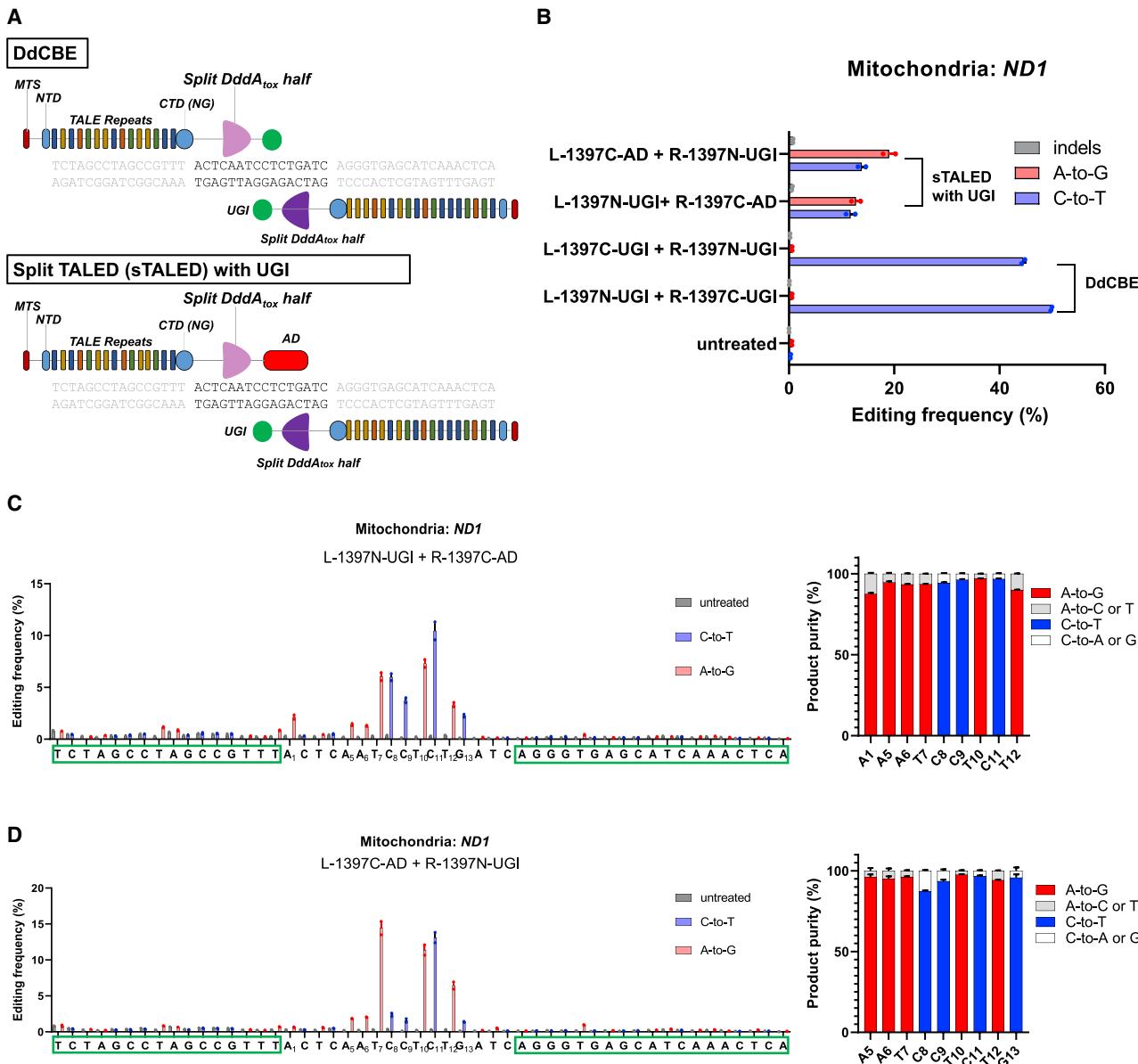


Figure 2. TALE deaminases enabling concurrent A-to-G and C-to-T editing

(A) Architectures of DdCBE and split TALE deaminase (sTALED) containing UGI.

(B) Editing frequencies of DdCBEs and sTALEDs containing UGI at the *ND1* site. Error bars represent SEM for two independent biological replicates.

(C and D) Editing frequencies at each base position and product purities of DdCBEs and sTALEDs containing UGI at the *ND1* site. TALE-binding sequences are outlined with green boxes. Error bars represent SEM for two independent biological replicates.

See also Figure S2.

than the TALE-E1347A-AD architecture (Figure S3A), whereas dimeric TALEDs (dTaleds) were composed of two neighboring, Watson and Crick strand-specific TALE arrays in a tail-to-tail configuration, each fused either to the TadA variant or the E1347A DddA_{tox} variant (Figure 4A). Both mTALEDs and dTALEDs induced A-to-G edits at the *ND1* site with frequencies that ranged from 10% to 35% and with high product purity (Figures 4B and 4C). Unwanted indels and C-to-T edits were rarely induced by these TALEDs. This result showed that the cyti-

dine deaminase activity of DddA_{tox} is dispensable for enhancing the adenine deaminase activity of the TadA variant on dsDNA and that E1347A DddA_{tox} can be either *cis*-acting (mTALEDs) or *trans*-acting (dTaleds) with the TadA variant.

We designed mTALEDs, dTALEDs, and TALEDs with DddA_{tox} split at G1397 (split TALEDs or sTALEDs) targeted to 11 additional sites (a total of 12 sites) in human mtDNA to compare the A-to-G editing efficiencies of the three types of programmable deaminases with one another in HEK 293T cells (Figures 4D,

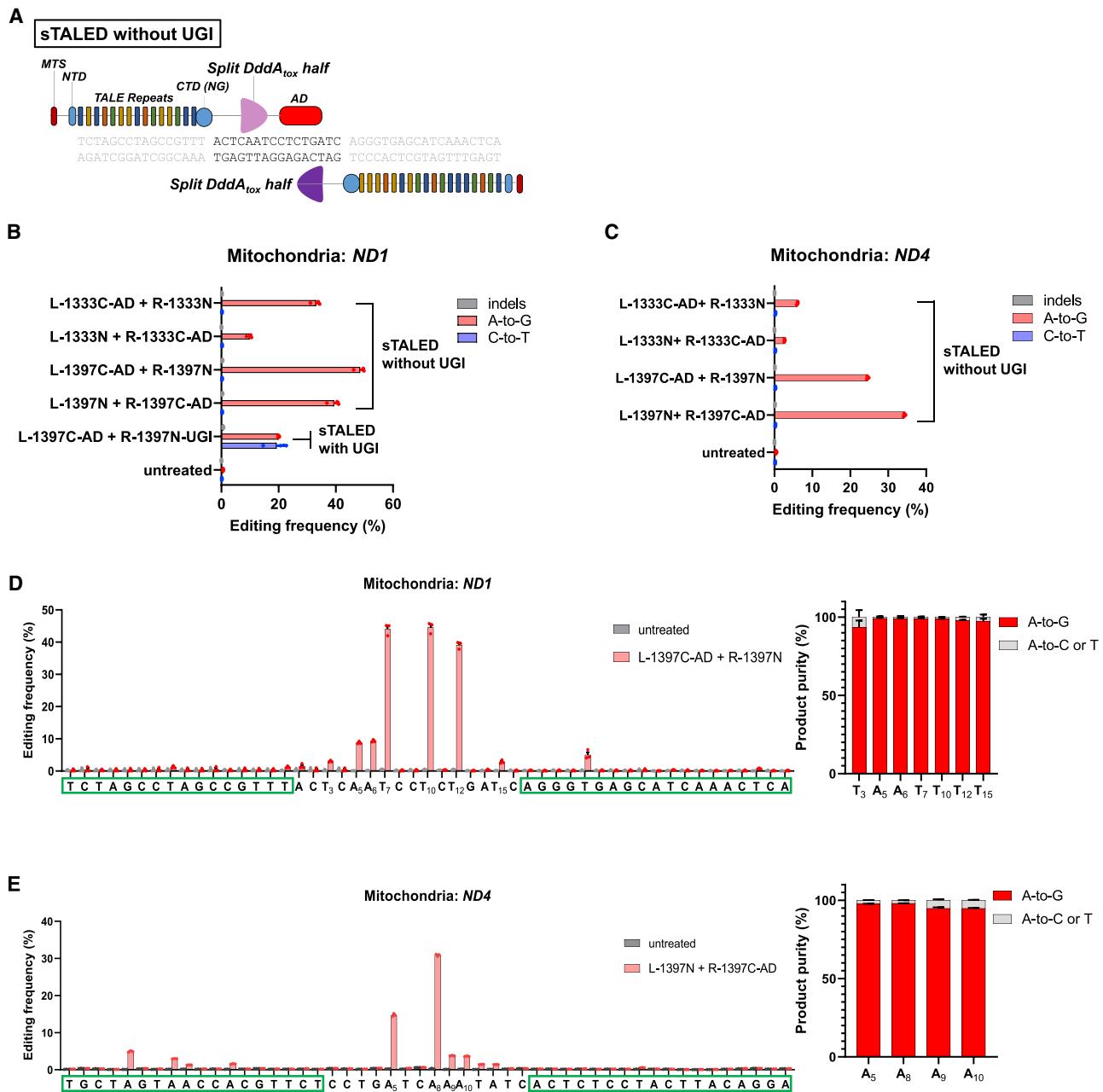


Figure 3. UGI-free TALE deaminases catalyzing A-to-G but not C-to-T conversions

(A) Architecture of split TALED without UGI targeted to the ND1 site.

(B and C) Editing frequencies of TALEDs at the ND1 and ND4 sites. Error bars represent SEM for three independent biological replicates.

(D and E) Editing frequencies at each base position and product purities of TALEDs at the ND1 and ND4 sites. Error bars represent SEM for three independent biological replicates. TALE-binding sequences are outlined with green boxes.

S3, and S4). sTALEDs were the most active with an average A-to-G editing frequency of $27\% \pm 3\%$ at a total of 12 sites including the ND1 site, on par with DdCBEs targeted to 5 sites with an average C-to-T editing frequency of $32\% \pm 4\%$ (Figure S3B). mTALEDs and dTALEDs were slightly less efficient than sTALEDs with an average editing frequency of $19\% \pm 4\%$. However, there were some exceptions: The CYTB-targeted

mTALED ($31\% \pm 2\%$) and the COX3.1-targeted dTALED ($28\% \pm 3\%$) were more efficient than the corresponding sTALEDs ($18\% \pm 1\%$ and $16\% \pm 1\%$, respectively) (Figure S4P).

We also found that TALEDs, unlike DdCBEs, could catalyze base editing at sites that did not contain a 5'-TC motif, which is recognized by DddA_{tox}. Thus, at the ND3, ND5.2, and ATP8 sites, which all lack a 5'-TC motif in the spacer region between

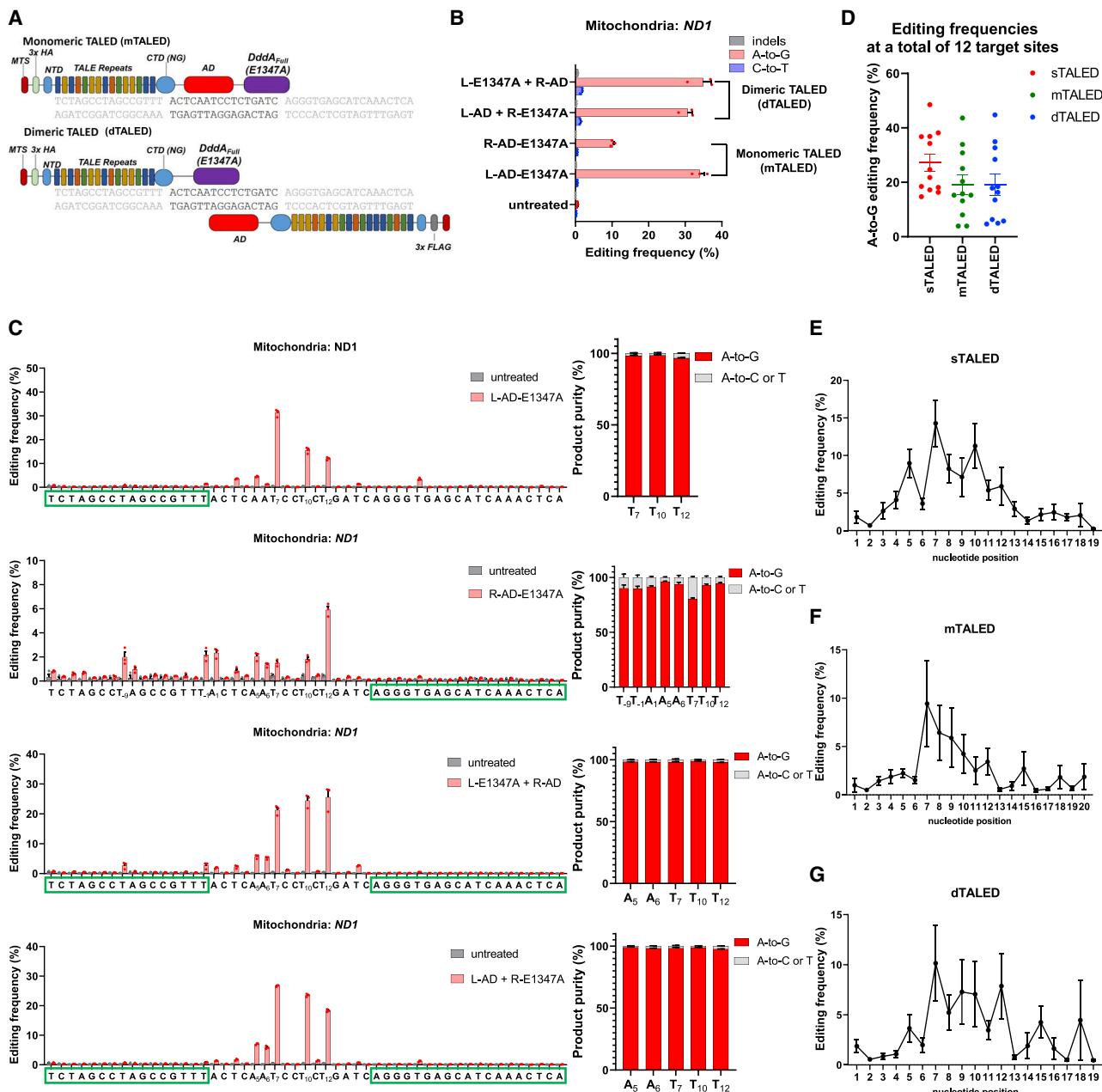


Figure 4. Monomeric and dimeric TALE deaminases enabling A-to-G conversions

(A) Architectures of monomeric and dimeric TALE deaminases (mTALEDs and dTALEDs, respectively) targeted to the *ND1* site.
(B and C) Editing frequencies of mTALEDs and dTALEDs at the *ND1* site. Error bars represent SEM for three independent biological replicates. TALE-binding sequences are shown in green boxes.
(D) Editing frequencies of mTALEDs, dTALEDs, and sTALEDs at 12 target sites. Lines and bars indicate mean and SEM, respectively.
(E–G) Editing frequencies obtained with total of 16 sites (sTALED) or 11 sites (dTALEDs and mTALEDs) at each nucleotide position downstream of a TadA8e-linked TALE-binding sequence.

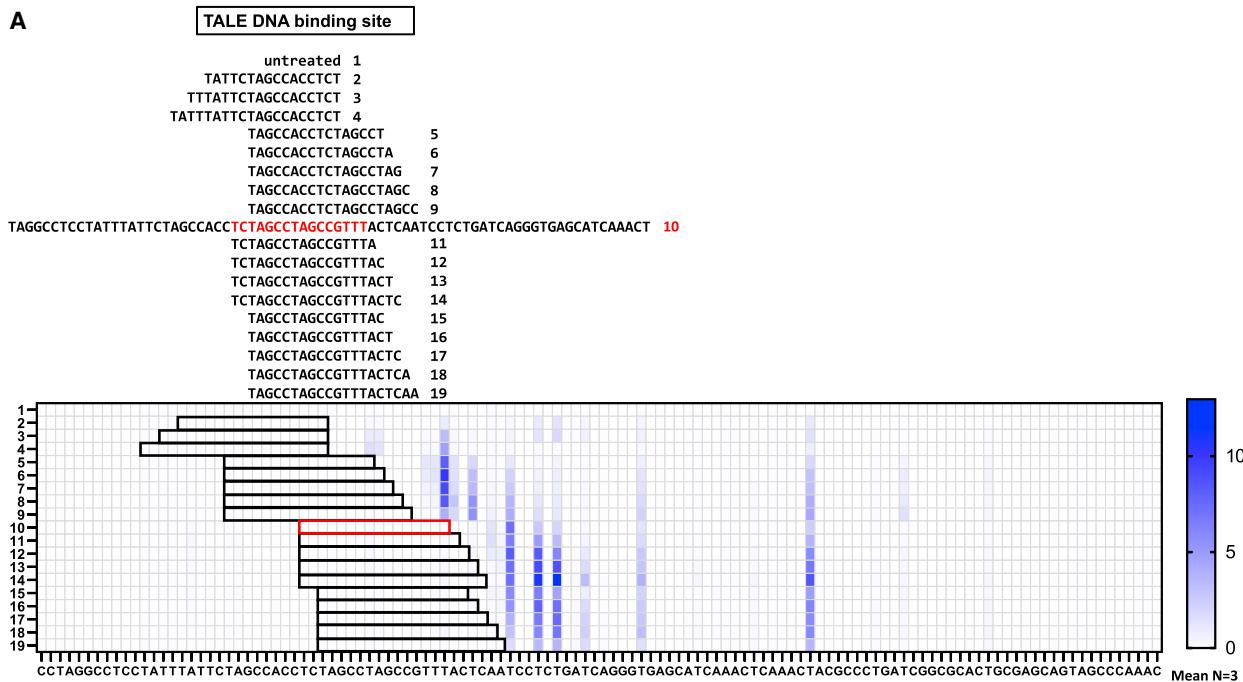
See also Figures S3 and S4.

the two TALE-binding sites, sTALEDs achieved A-to-G conversions with editing frequencies of $3.8\% \pm 0.5\%$, $13.4\% \pm 1.3\%$, and $19\% \pm 0.5\%$, respectively (Figures S4B, S4E, and S4H). This result suggests that, although the deaminase activity of DddA_{tox} in DdCBEs is largely limited to the 5'-TC context,

DddA_{tox} in TALEDs can make dsDNA accessible to the TadA variant without the requirement for the TC motif at a target site.

To define an editing window for sTALEDs, dTALEDs, and mTALEDs, we plotted adenine-editing frequencies at each nucleotide position downstream of a TadA8e-linked TALE-binding sequence.

A



B

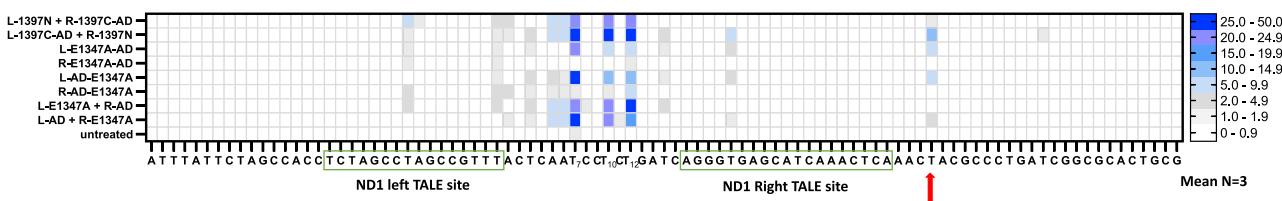


Figure 5. Shifting editing windows of TALE deaminases

(A) Top: monomeric TALE deaminases composed of a tiling array of overlapping TALE proteins targeted to the *ND1* site. The original left TALE-binding site is shown in red. Bottom: a heatmap showing A-to-G conversions by TALE deaminases. The original left TALE-binding site is outlined with a red box. An arrow indicates the position of a bystander edit (T39). Data are shown as means from $n = 3$ biologically independent samples.

(B) A heatmap showing A-to-G editing frequencies induced by sTALEDs, mTALEDs, and dTALEDs targeted to the *ND1* site. TALE-binding sites are shown in green boxes. Data are shown as means from $n = 3$ biologically independent samples.

Adenines immediately or farther downstream of the TadA8e-containing TALE-binding site were poorly edited, whereas those in the middle (positions 5 to 12 for sTALEDs or 7 to 12 for mTALEDs and dTALEDs in Figures 4E–G) were converted more efficiently. We also found that A and T (A in the opposite strand) in the editing window were equally editable by sTALEDs (Figure S4Q).

Shifting editing windows of TALE deaminases

We next designed a tiling array of TALE proteins, each fused to E1347A DddA_{tox} and the TadA variant, to construct a series of overlapping mTALEDs targeted to the *ND1* site and investigated whether editing windows could be shifted around the target site (Figure 5). As expected, mTALEDs showed A-to-G editing immediately 3' downstream of TALE-binding sites within a window of 10~20 bp in length. Some mTALEDs (e.g., #14 in Figure 5A) induced A-to-G conversion at nucleotide position 39 (indicated

by an arrow in Figure 5) \geq 35 bp downstream of the TALE-binding sites with editing frequencies of up to 9%.

We also compared the editing patterns of mTALEDs with those of dTALEDs and sTALEDs targeted to the *ND1* site. Interestingly, each TALED showed a different edit pattern. For example, the L-1397C-AD + R-1397N sTALED pair induced A-to-G conversions at three positions (T7, T10, and T12 in Figure 5B) comparably with >25% editing frequencies, whereas the L-AD-E1347A mTALED induced base edits only at T7 with a frequency of >25%. The L-E1347A + R-AD dTALED pair showed the highest editing frequency (>25%) at T12, whereas the L-AD + R-E1347A dTALED pair was most efficient at T7 with a frequency of $27\% \pm 0.2\%$. We also noted that a bystander edit at T39 (indicated by an arrow) observed with some mTALEDs and the L-1397C-AD + R-1397N sTALED pair was almost absent (<1.0%) with the L-E1347A + R-AD dTALED pair. Taken together, these results suggest that one can choose appropriate

TALEDs with desired edit patterns by testing sTALEDs, dTALEDs, and mTALEDs with overlapping TALE proteins.

Single-cell-derived clones containing mtDNA edits

We next investigated whether TALED-induced mtDNA edits were maintained in single-cell-derived clones. We isolated clonal populations of cells that had been transfected using the *ND4* and *ND1*-targeted split TALED or mTALED via limiting dilution (Figure S5). Targeted A-to-G edits were observed in 5 out of 10 clones (=50%) that had been treated using the *ND4*-specific sTALED with editing frequencies that ranged from 0.6% to 67% ($16\% \pm 13\%$, on average) and in 16 out of 20 clones (= 80%) that had been transfected using the *ND4* mTALED with editing frequencies of up to 76% ($23\% \pm 6\%$, on average) (Figures S5A and S5B). We also were able to obtain 5 *ND1*-edited, single-cell-expanded clones with editing frequencies of up to 84% ($38\% \pm 14\%$) using the *ND1*-targeted sTALEDs or mTALEDs (Figures S5C and S5D). The other TALED-transfected clones and several mock-treated clones showed editing frequencies of $\leq 0.3\%$, presumably caused by high-throughput sequencing errors or naturally occurring heteroplasmy. This result showed that TALED expression in cells could be tolerated, allowing isolation of single-cell-derived mutant clones, and also that mtDNA edits were not evenly distributed in a population of cells transfected with TALEDs.

Effects of TALED-mediated mtDNA editing on mitochondrial functions

We examined the effects of TALED-mediated mtDNA editing on cell viability, mtDNA copy numbers, and mitochondrial functions. We transfected each of five highly efficient sTALED pairs targeted to *ND1*, *ND4*, *ND5.2*, *COX3.1*, and *RNR2* into HEK 293T cells, monitored cell viability using an MTS assay for up to 15 days (Figure S6A), and measured mtDNA levels relative to genomic DNA levels by quantitative PCR for up to 16 days (Figure S6B). No significant changes in cell viability or mtDNA levels were observed in the resulting TALED-transfected cells, compared with control cells transfected with the plasmid encoding the inactive monomer R-1397N or L-1397N alone. We also measured oxygen consumption rates in transfected cells and found that A-to-G editing in these genes did not significantly alter rates of oxidative phosphorylation for up to 6 days (Figure S6C). These results suggest that TALED-mediated mtDNA editing is not cytotoxic and does not cause mtDNA instability or dysfunction.

mtDNA edits in the *RNR2* gene caused chloramphenicol resistance

To demonstrate that TALEDs can be used for mtDNA editing in cell lines other than HEK 293T and also that homoplasmic (>99%) edits can be obtained via drug selection, we transfected the *RNR2*-sTALEDs into human fibrosarcoma HT1080 and colon cancer HCT116 cell lines (Figure 6), which were then treated with chloramphenicol, an antibiotic inhibiting the peptidyl transferase activity of prokaryotic and mitochondrial ribosomes. Note that various mutations in the yeast or human *RNR2* gene encoding the mitochondrial 16S RNA can confer cells with chloramphenicol resistance (Figure 6A) (Kearsey and Craig, 1981; King and At-

tardi, 1988). Before chloramphenicol treatment (at day 3 post-transfection), the *RNR2* site was edited with frequencies of up to 17% in HT1080 cells (Figures 6C–6E) and 13% in HCT116 cells (Figures S6D–S6F). Editing frequencies at this site were increased to 92% or 87% at day 30 post-transfection in HT1080 or HCT116 cells, respectively, upon chloramphenicol treatment but were unchanged without the drug selection. Encouraged by this result, we isolated several clonal populations of HT1080 and HCT116 cells and found that these clones retained multiple A-to-G edits in the editing window with high frequencies of up to 99.9% (Figure S6G). As expected, the HCT116 clone with $\sim 99.9\%$ A-to-G mutations was resistant to chloramphenicol (Figure S6H). This result showed that TALED-induced mutations at the *RNR2* site caused chloramphenicol resistance and that homoplasmic mutations in mtDNA could be obtained by drug selection.

Off-target editing by TALEDs

We performed whole mitochondrial genome sequencing to profile off-target activities of various forms (split, monomeric, and dimeric) of TALEDs in comparison with those of DdCBEs. mtDNA samples isolated from TALED-untreated cells or cells transfected with TALE-free deaminase constructs were also analyzed. High-throughput sequencing revealed several naturally occurring single-nucleotide variations (SNVs) with a heteroplasmy fraction of $\geq 10\%$ across all DNA samples, including that isolated from TALED-untreated cells, which were excluded in our analyses of mtDNA genome-wide off-target editing frequencies. In line with previous reports, DdCBEs targeted to 6 human mitochondrial genes induced off-target mutations with average frequencies of mitochondrial genome-wide off-target editing that ranged from 0.010% to 0.018% ($0.013\% \pm 0.001\%$), 2–4-fold higher than that observed with the untreated control (0.005%) (Figure 7A). sTALEDs targeted to the same 6 sites and one additional site also showed off-target editing with frequencies that ranged from 0.013% to 0.025% ($0.019\% \pm 0.002\%$). mTALEDs and dTALEDs were more specific than sTALEDs, with average off-target editing frequencies of $0.009\% \pm 0.001\%$ and $0.008\% \pm 0.001\%$, respectively. As expected, DdCBE off-target edits were largely confined to C-to-T conversions, whereas TALED off-target edits were largely limited to A-to-G conversions. Interestingly, TALE-free split DddA_{tox} fused to the TadA variant (1397N + 1397C-AD) showed a 24-fold higher average off-target editing frequency (0.12%) compared with the untreated control, suggesting that the spontaneous assembly of DddA_{tox} can cause off-target A-to-G edits by the TadA deoxyadenosine deaminase. Notably, however, the TadA variant fused to the catalytically deficient, full-length E1347A DddA_{tox} variant (AD-E1347A), with an average mitochondrial genome-wide editing frequency of 0.004%, did not exhibit off-target editing compared with the untreated control, whereas separate expression of the TALE-free E1347A DddA_{tox} variant and the TALE-free TadA deoxyadenosine deaminase (AD + E1347A) led to off-target effects with an average off-target editing frequency of 0.015%. This result suggests that off-target edits caused by the TALE-free deaminases (1397N + 1397C-AD and AD + E1347A) can be largely avoided by using TALE deaminase fusion proteins.

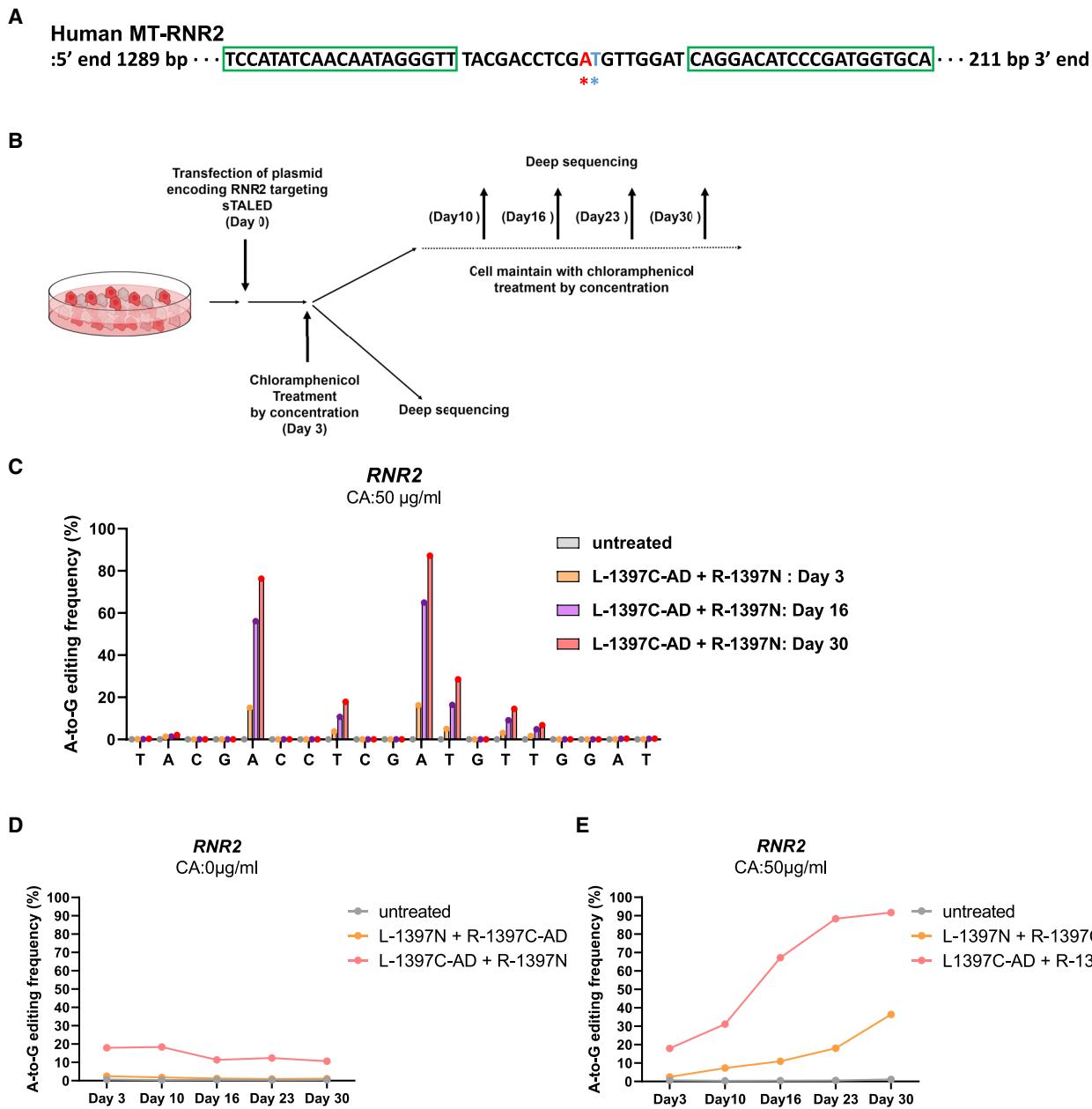


Figure 6. TALED-mediated RNR2 mutations causing chloramphenicol resistance

(A) The RNR2 site targeted by TALEDs. TALE-binding sequences are outlined with green boxes. Two nucleotides shown in red and blue indicate the positions at which mutations responsible for chloramphenicol resistance in yeast and human cells, respectively.

(B) Experimental scheme through which chloramphenicol resistance was induced by TALEDs.

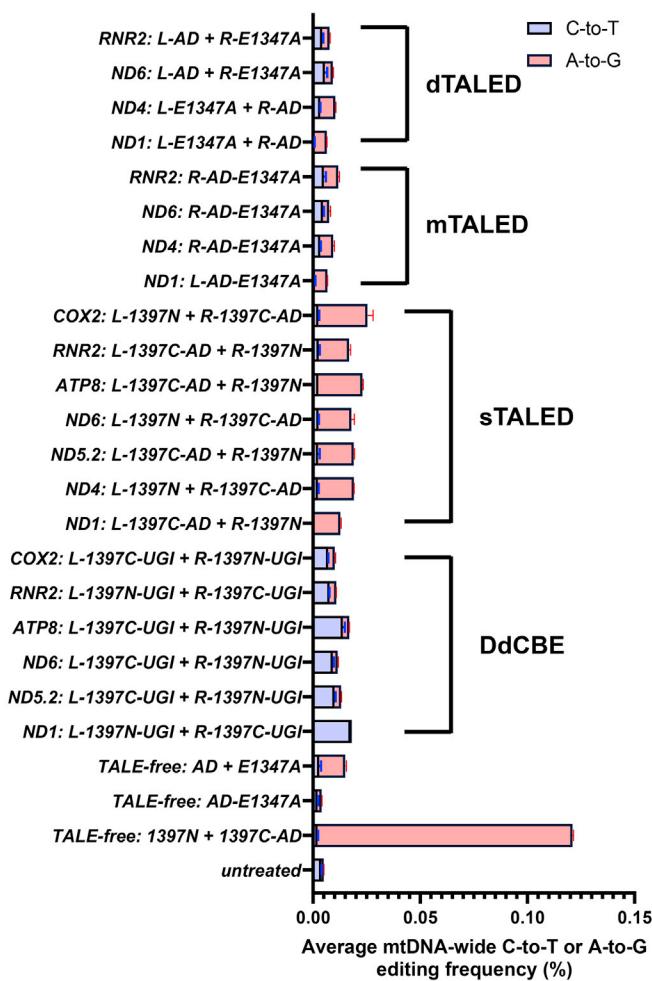
(C–E) Editing frequencies at day 3, 10, 16, 23, and 30 post-transfection.

See also Figure S6.

We next examined the positions of off-target edits induced by DdCBEs and TALEDs composed of the same TALE arrays (Figure 7B). The *ND1*-specific DdCBE caused C-to-T off-target edits at 41 sites in human mtDNA with frequencies of $\geq 1.0\%$, whereas sTALED, composed of the same TALE arrays, induced A-to-G edits at 32 sites. Notably, DdCBE off-target sites did not overlap with sTALED off-target sites, suggesting that off-target

effects caused by these deaminases are independent of the TALE arrays. In fact, no consensus sequences homologous to the two TALE-binding sites were observed at the DdCBE and TALED off-target sites (Figure S7A). mTALED and dTALED containing E1347A DddA_{tox} targeted to the same site largely avoided these off-target edits. Thus, *ND1*-mTALED induced merely 10 A-to-G off-target edits with frequencies of $\geq 1.0\%$, whereas

A



B

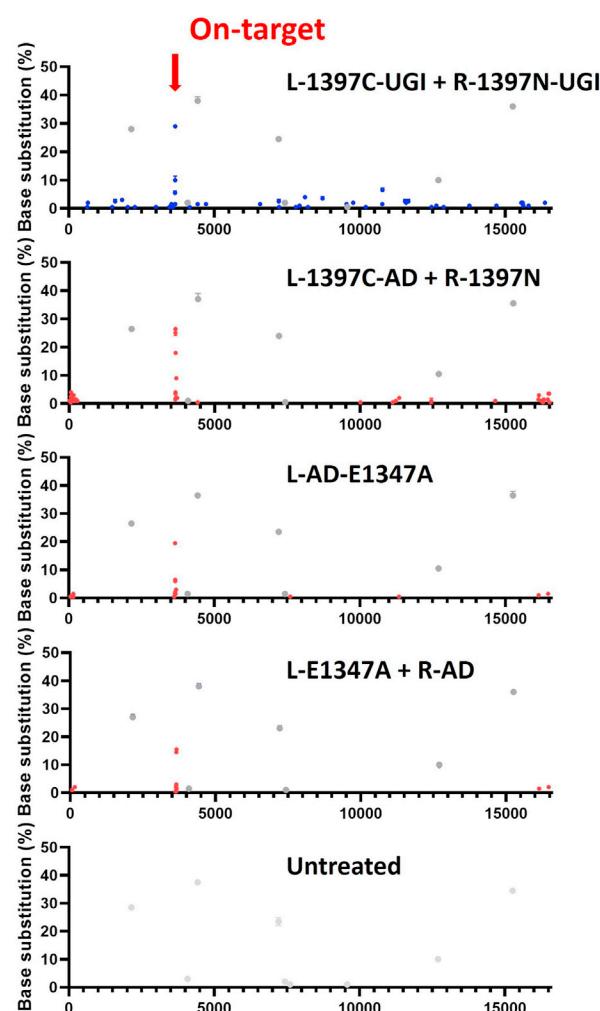


Figure 7. Mitochondrial genome-wide off-target effects

(A) Average percentage of mitochondrial genome-wide off-target editing for DdCBEs and TALEDs and for TALE-free controls containing a mitochondrial targeting sequence. Data are shown as means from $n = 2$ biologically independent samples.

(B) Mitochondrial genome-wide plots showing on-target and off-target sites. Blue and red dots indicate C-to-T and A-to-G substitutions, respectively, with $\geq 1\%$ frequencies. Gray dots indicate naturally occurring SNVs. The x axis represents nucleotide positions in human mtDNA. Data are shown as means from $n = 2$ biologically independent samples.

See also Figure S7 and Table S5.

ND1-dTALED induced 8 A-to-G off-target edits. Interestingly, most of these off-target edits were found in the displacement-loop (D-loop), a highly polymorphic, non-coding region of 1.1 kbp in length in human mtDNA (Table S5).

Finally, we investigated whether TALEDs containing an MTS rather than a nuclear localization signal (NLS) could, nevertheless, induce off-target DNA editing in the nuclear genome. We identified several potential off-target sites in the nuclear genome, which differ by one or two nucleotides from their respective on-target sites in mtDNA, and measured A-to-G editing frequencies at these sites using high-throughput sequencing. None of 10 sTALED pairs detectably induced A-to-G edits at these sites in

the nuclear genome (Figure S7B), suggesting that TALEDs containing an MTS do not have access to nuclear DNA.

DISCUSSION

In this study, we generated three different types of base editors termed TALEDs that enable A-to-G base editing in human mtDNA. We found that a full-length, non-toxic E1347A DddA_{tox} variant functioned in *cis* (mTALEDs) or in *trans* (dTALEDs) with the deoxyadenosine deaminase derived from TadA fused to a custom-designed TALE DNA-binding protein to allow the deaminase to catalyze A-to-G conversions in double-stranded

mtDNA. Both mTALEDs and dTALEDs were reasonably specific, merely inducing several off-target edits in the mtDNA D-loop, which appeared to be TALE-independent, with low frequencies. sTALEDs containing split DddA_{tox} rather than the full-length variant were more efficient in general than mTALEDs or dTALEDs, composed of the same TALE arrays, but were less specific. mTALED-encoding genes (3.1~3.5 kbp in length), unlike those encoding sTALEDs or dTALEDs or DdCBEs (>5.3 kbp), can be readily packaged in a viral vector with a small cargo size such as adeno-associated virus (AAV), a key advantage in *in vivo* gene delivery. Furthermore, mTALEDs with one TALE protein are more flexible than sTALEDs and dTALEDs with two TALE proteins, whose binding sites typically have thymine at both the 5' and 3' ends, although thymine at the 3' end is not essential: As a result, mTALEDs can be designed to bind to certain sites, whereas sTALEDs and dTALEDs cannot.

We admit that we cannot explain exactly how DddA_{tox} can make dsDNA accessible to the ssDNA-specific TadA variant. We speculate that DddA_{tox} may transiently unwind dsDNA, exposing a few nucleotides, upon binding to dsDNA or catalyzing cytosine deamination. It is also of note that TadA8e is a highly active enzyme, capturing transient DNA melting induced by Cas9 (Lapinaite et al., 2020). It will be interesting to investigate whether DddA_{tox} can be used as a generalized platform to make dsDNA accessible to other TadA variants with slower kinetics and also to other ssDNA-specific DNA-modifying enzymes to expand the scope of genome and epigenome editing.

In the short term, TALEDs will be useful for generating mtDNA mutations in cell lines and animals to create disease models, an essential step in drug development. Out of 90 confirmed pathogenic mtDNA point mutations listed in mitomap, 42 (=47%) of them can be produced in human cell lines or model organisms using TALEDs enabling A-to-G conversions, whereas only 9 (=10%) of the mutations can be induced using DdCBEs catalyzing TC-to-TT conversions. In the long term, TALEDs with improved efficiency and specificity could pave the way for correction of disease-causing mtDNA mutations in embryos, fetuses, newborns, or adult patients, heralding a new era of mitochondrial gene therapy. In addition, as shown recently with DdCBEs (Kang et al., 2021; Li et al., 2021b; Nakazato et al., 2021), chloroplast DNA encoding hundreds of genes in plants, many of which are essential for photosynthesis, can be edited with plant-compatible TALEDs, opening a new chapter in plant genetics and biotechnology.

Limitations of the study

Both TALEDs and DdCBEs are limited by bystander editing: in addition to the intended target nucleotide, adjacent nucleotides in an editing window can also be converted. Here, we showed that the use of overlapping TALE proteins in the construction of a tiling array of TALEDs can partially address this problem. Engineering of the TadA deoxyadenosine deaminase or DddA_{tox} may reduce or eliminate the bystander editing activity of TALEDs. TadA8e is known to catalyze adenine deamination in RNA (Lapinaite et al., 2020). As a result, CRISPR RNA-guided adenine base editors containing TadA8e, the Cas9 nickase, UGI, and an NLS induce off-target RNA editing in mammalian cells. Although TALEDs contain an MTS rather than an NLS, it

is possible that they still induce off-target editing of cellular RNAs encoded in the genome as well as mitochondrial RNAs. TadA variants with reduced RNA editing activity (Grünwald et al., 2019; Li et al., 2021a; Rees et al., 2019; Zhou et al., 2019) can be used to avoid this potential problem.

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
 - Cell lines
- METHOD DETAILS
 - Plasmid construction
 - Cell culture and transfection
 - Measurement of oxygen consumption rates
 - Cell viability assays
 - Chloramphenicol selection
 - Purification of genomic DNA from cultured mammalian cells
 - mtDNA copy number measurement
 - Growth rate determination
 - Targeted deep sequencing
 - Whole mitochondrial genome sequencing
 - Analysis of mitochondrial genome-wide off-target editing
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.cell.2022.03.039>.

ACKNOWLEDGMENTS

This work was supported by the Institute for Basic Science (IBS-R021-D1 to J.-S.K.).

AUTHOR CONTRIBUTIONS

J.-S.K. supervised the research. S.-I.C. and J.-S.K. designed the study. S.-I.C., S.L., Y.G.M., K.L., J.L., J.M.L., and E.C. performed the experiments. S.-I.C. and J.-S.K. wrote the manuscript.

DECLARATION OF INTERESTS

S.-I.C., S.L., Y.G.M., K.L., J.M.L., E.C., and J.-S.K. have submitted a provisional patent application based on results reported in this paper. J.-S.K. is a founder of and shareholder in ToolGen. The other authors declare no competing interests.

Received: December 17, 2021

Revised: February 18, 2022

Accepted: March 24, 2022

Published: April 25, 2022

REFERENCES

- 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. <https://doi.org/10.1038/s41591-018-0166-8>.
- 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. <https://doi.org/10.1002/emmm.201303672>.
- Gammage, P.A., Gaude, E., Van Haute, L., Rebelo-Guiomar, P., Jackson, C.B., Rorbach, J., Pekalski, M.L., Robinson, A.J., Charpentier, M., Conordet, J.P., et al. (2016). Near-complete elimination of mutant mtDNA by iterative or dynamic dose-controlled treatment with mtZFNs. *Nucleic Acids Res.* 44, 7804–7816. <https://doi.org/10.1093/nar/gkw676>.
- Gammage, P.A., Moraes, C.T., and Minczuk, M. (2018a). Mitochondrial genome engineering: the revolution may not be CRISPR-ized. *Trends Genet.* 34, 101–110. <https://doi.org/10.1016/j.tig.2017.11.001>.
- Gammage, P.A., Visconti, C., Simard, M.L., Costa, A.S.H., Gaude, E., Powell, C.A., Van Haute, L., McCann, B.J., Rebelo-Guiomar, P., Cerutti, R., et al. (2018b). Genome editing in mitochondria corrects a pathogenic mtDNA mutation in vivo. *Nat. Med.* 24, 1691–1695. <https://doi.org/10.1038/s41591-018-0165-9>.
- Gaudelli, N.M., Komor, A.C., Rees, H.A., Packer, M.S., Badran, A.H., Bryson, D.I., and Liu, D.R. (2017). Programmable base editing of A*T to G*C in genomic DNA without DNA cleavage. *Nature* 551, 464–471. <https://doi.org/10.1038/nature24644>.
- 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. <https://doi.org/10.1002/ana.24362>.
- Grünewald, J., Zhou, R., Iyer, S., Lareau, C.A., Garcia, S.P., Aryee, M.J., and Joung, J.K. (2019). CRISPR DNA base editors with reduced RNA off-target and self-editing activities. *Nat. Biotechnol.* 37, 1041–1048. <https://doi.org/10.1038/s41587-019-0236-6>.
- Grünewald, J., Zhou, R., Lareau, C.A., Garcia, S.P., Iyer, S., Miller, B.R., Langner, L.M., Hsu, J.Y., Aryee, M.J., and Joung, J.K. (2020). A dual-deaminase CRISPR base editor enables concurrent adenine and cytosine editing. *Nat. Biotechnol.* 38, 861–864. <https://doi.org/10.1038/s41587-020-0535-y>.
- 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. <https://doi.org/10.1038/s41421-021-00307-9>.
- 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. Therapy Nucleic Acids* 27, P73–P80. <https://doi.org/10.1016/j.mtn.2021.11.016>.
- 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. <https://doi.org/10.1038/s41477-021-00943-9>.
- Kazama, T., Okuno, M., Watari, Y., Yanase, S., Koizuka, C., Tsuruta, Y., Sugaya, H., Toyoda, A., Itoh, T., Tsutsumi, N., et al. (2019). Curing cytoplasmic male sterility via TALEN-mediated mitochondrial genome editing. *Nat. Plants* 5, 722–730. <https://doi.org/10.1038/s41477-019-0459-z>.
- Kearsey, S.E., and Craig, I.W. (1981). Altered ribosomal RNA genes in mitochondria from mammalian cells with chloramphenicol resistance. *Nature* 290, 607–608. <https://doi.org/10.1038/290607a0>.
- Kim, Y., Kweon, J., Kim, A., Chon, J.K., Yoo, J.Y., Kim, H.J., Kim, S., Lee, C., Jeong, E., Chung, E., et al. (2013). A library of TAL effector nucleases spanning the human genome. *Nat. Biotechnol.* 31, 251–258. <https://doi.org/10.1038/nbt.2517>.
- King, M.P., and Attardi, G. (1988). Injection of mitochondria into human cells leads to a rapid replacement of the endogenous mitochondrial DNA. *Cell* 52, 811–819. [https://doi.org/10.1016/0092-8674\(88\)90423-0](https://doi.org/10.1016/0092-8674(88)90423-0).
- Komor, A.C., Kim, Y.B., Packer, M.S., Zuris, J.A., and Liu, D.R. (2016). Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* 533, 420–424. <https://doi.org/10.1038/nature17946>.
- Lapinaite, A., Knott, G.J., Palumbo, C.M., Lin-Shiao, E., Richter, M.F., Zhao, K.T., Beal, P.A., Liu, D.R., and Doudna, J.A. (2020). DNA capture by a CRISPR-Cas9-guided adenine base editor. *Science* 369, 566–571. <https://doi.org/10.1126/science.abb1390>.
- 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. <https://doi.org/10.1038/s41467-021-21464-1>.
- Li, C., Zhang, R., Meng, X., Chen, S., Zong, Y., Lu, C., Qiu, J.L., Chen, Y.H., Li, J., and Gao, C. (2020). Targeted, random mutagenesis of plant genes with dual cytosine and adenine base editors. *Nat. Biotechnol.* 38, 875–882. <https://doi.org/10.1038/s41587-019-0393-7>.
- Li, J., Yu, W., Huang, S., Wu, S., Li, L., Zhou, J., Cao, Y., Huang, X., and Qiao, Y. (2021a). Structure-guided engineering of adenine base editor with minimized RNA off-targeting activity. *Nat. Commun.* 12, 2287. <https://doi.org/10.1038/s41467-021-22519-z>.
- Li, R., Char, S.N., Liu, B., Liu, H., Li, X., and Yang, B. (2021b). High-efficiency plastome base editing in rice with TAL cytosine deaminase. *Mol. Plant* 14, 1412–1414. <https://doi.org/10.1016/j.molp.2021.07.007>.
- Li, X., Wang, Y., Liu, Y., Yang, B., Wang, X., Wei, J., Lu, Z., Zhang, Y., Wu, J., Huang, X., et al. (2018). Base editing with a Cpf1-cytidine deaminase fusion. *Nat. Biotechnol.* 36, 324–327. <https://doi.org/10.1038/nbt.4102>.
- Lim, K., Cho, S.I., and Kim, J.S. (2022). Nuclear and mitochondrial DNA editing in human cells with zinc finger deaminases. *Nat. Commun.* 13, 366. <https://doi.org/10.1038/s41467-022-27962-0>.
- Mok, B.Y., de Moraes, M.H., Zeng, J., Bosch, D.E., Kotrys, A.V., Raguram, A., Hsu, F., Radley, 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. <https://doi.org/10.1038/s41586-020-2477-4>.
- Nakazato, I., Okuno, M., Yamamoto, H., Tamura, Y., Itoh, T., Shikanai, T., Takanashi, H., Tsutsumi, N., and Arimura, S.I. (2021). Targeted base editing in the plastid genome of *Arabidopsis thaliana*. *Nat. Plants* 7, 906–913. <https://doi.org/10.1038/s41477-021-00954-6>.
- 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, <https://doi.org/10.1126/science.aaf8729>.
- Reddy, P., Ocampo, A., Suzuki, K., Luo, J., Bacman, S.R., Williams, S.L., Sugawara, A., Okamura, D., Tsunekawa, Y., Wu, J., et al. (2015). Selective elimination of mitochondrial mutations in the germline by genome editing. *Cell* 161, 459–469. <https://doi.org/10.1016/j.cell.2015.03.051>.
- Rees, H.A., Wilson, C., Doman, J.L., and Liu, D.R. (2019). Analysis and minimization of cellular RNA editing by DNA adenine base editors. *Sci. Adv.* 5, eaax5717. <https://doi.org/10.1126/sciadv.aax5717>.
- Richter, M.F., Zhao, K.T., Eton, E., Lapinaite, A., Newby, G.A., Thuronyi, B.W., Wilson, C., Koblan, L.W., Zeng, J., Bauer, D.E., et al. (2020). Phage-assisted evolution of an adenine base editor with improved Cas domain compatibility and activity. *Nat. Biotechnol.* 38, 883–891. <https://doi.org/10.1038/s41587-020-0453-z>.
- Sakata, R.C., Ishiguro, S., Mori, H., Tanaka, M., Tatsuno, K., Ueda, H., Yamamoto, S., Seki, M., Masuyama, N., Nishida, K., et al. (2020). Base editors for simultaneous introduction of C-to-T and A-to-G mutations. *Nat. Biotechnol.* 38, 865–869. <https://doi.org/10.1038/s41587-020-0509-0>.
- Silva-Pinheiro, P., and Minczuk, M. (2022). The potential of mitochondrial genome engineering. *Nat. Rev. Genet.* 23, 199–214. <https://doi.org/10.1038/s41576-021-00432-x>.
- Taylor, R.W., and Turnbull, D.M. (2005). Mitochondrial DNA mutations in human disease. *Nat. Rev. Genet.* 6, 389–402. <https://doi.org/10.1038/nrg1606>.

- Yang, L., Briggs, A.W., Chew, W.L., Mali, P., Guell, M., Aach, J., Goodman, D.B., Cox, D., Kan, Y., Lesha, E., et al. (2016). Engineering and optimising deaminase fusions for genome editing. *Nat. Commun.* 7, 13330. <https://doi.org/10.1038/ncomms13330>.
- Zhang, X., Zhu, B., Chen, L., Xie, L., Yu, W., Wang, Y., Li, L., Yin, S., Yang, L., Hu, H., et al. (2020). Dual base editor catalyzes both cytosine and adenine base conversions in human cells. *Nat. Biotechnol.* 38, 856–860. <https://doi.org/10.1038/s41587-020-0527-y>.
- Zhou, C., Sun, Y., Yan, R., Liu, Y., Zuo, E., Gu, C., Han, L., Wei, Y., Hu, X., Zeng, R., et al. (2019). Off-target RNA mutation induced by DNA base editing and its elimination by mutagenesis. *Nature* 571, 275–278. <https://doi.org/10.1038/s41586-019-1314-0>.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and virus strains		
NEB 5-alpha Electrocompetent E. coli	New England Biolabs	Cat#C2989K
Chemicals, peptides, and recombinant proteins		
Lipofectamine 2000	Thermo Fisher Scientific	Cat#11668019
EcoRI	New England BioLabs	Cat#R0101S
DpnI	New England BioLabs	Cat#R0176S
Bsal-HFv2	New England BioLabs	Cat#R3733S
T4 DNA Ligase	New England BioLabs	Cat#M0202S
XbaI	New England BioLabs	Cat#R0145S
Esp3I	New England BioLabs	Cat#R0734S
BamHI	New England BioLabs	Cat#R0136S
BsmBI-v2	New England BioLabs	Cat#R0739S
XhoI	New England BioLabs	Cat#R0146L
XcmI	New England BioLabs	Cat#R0533S
Proteinase K	QIAGEN	Cat#19133
NEBuilder HiFi DNA Assembly Master Mix	New England BioLabs	Cat#E2621L
Q5® High-Fidelity DNA Polymerase	New England BioLabs	Cat#M0491L
PrimeSTAR® GXL DNA Polymerase	TAKARA	Cat#R051A
Fetal bovine serum (FBS)	WELGENE	Cat#S101-01
DMEM	WELGENE	Cat#LM001-05
RPMI 1640	WELGENE	Cat#LM011-01
Trypsin-EDTA 1X	WELGENE	Cat#LS015-01
Antibiotic Antimycotic Solution	WELGENE	Cat#LS203-01D
Dimethyl Sulfoxide	SIGMA	Cat#D2650
KANAMYCIN MONOSULFATE	Duchefa Biochemie	Cat#K0126.0025
AMPICILLIN SODIUM	Duchefa Biochemie	Cat#A0104.0025
MICRO AGAR	Duchefa Biochemie	Cat#M1002.0500
LB BROTH HIGH SALT	Duchefa Biochemie	Cat#L1704.0500
LB AGAR HIGH SALT	Duchefa Biochemie	Cat#L1706.0500
Certified™ Molecular Biology Agarose	BIO-RAD	Cat#19133
Dulbecco's Phosphate-Buffered Saline (D-PBS) (1X), liquid	WELGENE	Cat#LB 001-02
Critical commercial assays		
Exprep™ Plasmid SV	GeneALL	Cat#101-102
MG Gel Extraction kit	MGmed	Cat#M0494S
Mitochondria Isolation Kit for Cultured Cells	Thermo	Cat#89874
Nextera™ DNA CD Indexes	Illumina	Cat#20018708
ILMN DNA LP (M)Tag	Illumina	Cat#20018704
NucleoBond Xtra Midi Plus EF, Midi kit for endotoxin-free plasmid DNA	MACHEERY-NAGEL	Cat# 740420.50
DNeasy Blood & Tissue Kit	Qiagen	Cat#69506
MG PCR Product Purification SV	MGmed	Cat#MK12020
KAPA SYBR® FAST qPCR Kit	KAPABIOSYSTEMS	Cat#KK4602
CellTiter 96® Aqueous One Solution	Promega	Cat#G3582
MiniSeq Mid Output Kit	Illumina	Cat#FC-420-1004

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
PhiX Control v3 kit	Illumina	Cat#FC-110-3002.
Seahorse FluxPak	Agilent	Cat#102416-100
XF Cell Mito Stress Test Kit	Agilent	Cat#103015-100
Deposited data		
Targeted amplicon sequencing data	This study	PRJNA789761
Experimental models: Cell lines		
Human: HEK293T	ATCC	Cat#CRL-11268
Human: HCT116	Korean Cell Line Bank	Cat#10247
Human: HT1080	Korean Cell Line Bank	Cat#10121
Oligonucleotides		
Primers used for deep sequencing preparation, see Table S4	This manuscript	N/A
Recombinant DNA		
p3s-ND5-Left-G1397-C	Mok et al. (2020)	Addgene Cat#169212
p3s-ND5-Right-G1397-N	Mok et al. (2020)	Addgene Cat#169213
TALED_Left-ND1-1397C-AD	This study	Addgene Cat#183892
TALED_Left-ND1-1397C-UGI	This study	Addgene Cat#183893
TALED_Left-ND1-AD-E1347A	This study	Addgene Cat#183894
TALED_Left-ND1-E1347A	This study	Addgene Cat#183895
TALED_Left-RNR2-1397C-AD	This study	Addgene Cat#183896
TALED_Right-RNR2-1397N	This study	Addgene Cat#183897
TALED_Right-ND1-1397N	This study	Addgene Cat#183898
TALED_Right-ND1-1397N-UGI	This study	Addgene Cat#183899
TALED_Right-ND1-AD	This study	Addgene Cat#183900
Software and algorithms		
GraphPad Prism 9	GraphPad	https://www.graphpad.com/
Base editing frequencies and indel calculation algorithms	This study	https://github.com/ibs-cge/maund

RESOURCE AVAILABILITY

Lead contact

Please direct requests for resources and reagents to Lead Contact: Jin-Soo Kim (J.-S.K. jskim01@snu.ac.kr).

Materials availability

Plasmids generated in this study have been deposited to Addgene (see [key resources table](#) for details).

Data and code availability

- The sequencing data generated during this study are available at the NCBI Sequence Read Archive database under PRJNA789761.
- The code used for data processing and analysis is available at <https://github.com/ibs-cge/maund>.
- Any additional information required to reanalyze the data reported in this work paper is available from the [Lead Contact](#) upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Cell lines

HEK 293T cells were purchased from ATCC (ATCC CRL-11268). HCT116 cells and HT1080 cells were purchased from the Korean Cell Line Bank (KCLB No.10247 and KCLB No. 10121). HEK 293T cells were maintained in Dulbecco's Modified Eagle Medium

(DMEM; Welgene) supplemented with 10% fetal bovine serum (Welgene) and 1% antibiotic-antimycotic solution (Welgene). HCT116 cells and HT1080 cells were cultured in RPMI1640 medium containing L-glutamine (300 mg/L), 25 mM HEPES, and 25 mM NaHCO₃ (Welgene); this medium was also supplemented with 10% fetal bovine serum (Welgene) and 1% antibiotic-antimycotic solution (Welgene). Cells were maintained at 37°C with 5% CO₂. HEK 293T cells were passaged once every 2-3 days, and HCT116 and HT1080 cells were passaged once every 3-4 days. All cell lines were passaged before reaching 80% confluence.

METHOD DETAILS

Plasmid construction

In preparation for constructing plasmids encoding TALEDs that target specific sites, we first created master vectors for the sTALED, mTALED, and dTALED systems. These vectors were generated from plasmids p3s-ND5-Left-G1397-C (Addgene, #169212) and p3s-ND5-Right-G1397-N (Addgene, #169213), which encode DdCBE pairs targeted to the mouse *ND5* gene; the new master vectors encode the required components for the different TALED constructs and can be modified by inserting a custom-designed TALE array sequence, similar to what was previously done in the construction of a TALEN library (Kim et al., 2013). In all cases, p3s-ND5-Left-G1397-C and p3s-ND5-Right-G1397-N were digested with XcmI (NEB) and EcoRI (NEB) to remove the sequences encoding DdCBE, leaving other sequences for features such as the MTS and tag; inserts including a stuffer containing a Bsal enzyme (NEB) site and sequences encoding TALED components, synthesized by IDT, were inserted into the digested vector using a HiFi DNA assembly kit (NEB). In summary, the compositions of the master vectors are as follows: sTALED master vector (p3s-stuffer-DddA_{tox} half (1333C or 1397C)-AD and p3s-stuffer-DddA_{tox} half (1333N or 1397N)); mTALED master vector (p3s-stuffer-E1347A DddA_{tox} full-AD); and dTALED master vector (p3s-stuffer-E1347A DddA_{tox} full-AD and p3s-stuffer-AD). As in the previous study (Kim et al., 2013), the desired TALED construct could be produced by digesting the master vector with Bsal to cleave the site in the stuffer and assembling 6 TALE arrays in that position through the Golden Gate system. See also [Tables S2](#) and [S3](#).

Cell culture and transfection

HEK 293T cells were maintained in DMEM (Welgene) supplemented with 10% fetal bovine serum (Welgene) and 1% antibiotic-antimycotic solution (Welgene). HCT116 cells and HT1080 cells were cultured in RPMI1640 medium containing L-glutamine (300 mg/L), 25 mM HEPES, and 25 mM NaHCO₃ (Welgene); this medium was also supplemented with 10% fetal bovine serum (Welgene) and 1% antibiotic-antimycotic solution (Welgene). Cell lines were maintained at 37°C with 5% CO₂ and were passaged depending on the doubling time of the particular cell line, before cells reached 80% confluence.

Prior to transfection, cells (7.5×10^4) were seeded into 48-well plates. After 24 h, the cells were transfected with plasmids (500 ng) using Lipofectamine 2000 (1.5 µL, Invitrogen). When only one construct was required, such as in the case of mTALED, the total amount of transfected plasmid was 500 ng; when two constructs were required, in the case of dTALED, sTALED, and DdCBE, the total amount of plasmid was 1000 ng (500 ng each). After 96 h, the transfected cells were harvested and lysed by incubation at 55 °C for 1 h, and then at 95 °C for 10 min, in 100 µL of cell lysis buffer (50 mM Tris-HCl (Sigma-Aldrich), 1 mM EDTA (Sigma-Aldrich), 0.005% sodium dodecyl sulfate (Sigma-Aldrich), pH 8.0 at 25 °C) supplemented with 5 µL of Proteinase K (Qiagen).

Measurement of oxygen consumption rates

Oxygen consumption rates were measured using Seahorse XFe96 Analyzer (Agilent) and following the manufacturer's protocol. 100 µL of cells (1×10^6 cells/ml) were seeded into Seahorse XF96 cell culture microplates (Agilent) 16 h before measurements were taken. Analysis was performed in Seahorse XF DMEM pH 7.4 supplemented with 25 mM glucose and 1 mM sodium pyruvate (Agilent). The XF cell mito stress test protocol was applied using 1.5 µM oligomycin, 0.5 µM FCCP, and 0.5 µM rotenone + antimycin A.

Cell viability assays

Cell viability assays were performed using CellTiter 96® Aqueous One Solution (Promega) every 3 or 4 days over a 15-day time period. The MTS compound is bio-reduced into formazan, a colored substance, such that its concentration is proportional to the number of living cells. The viability is measured by recording the absorbance at 490nm. The protocol supplied by Promega was followed for the experiments.

Chloramphenicol selection

To introduce a mutation causing chloramphenicol resistance in the *RNR2* gene in mtDNA, a plasmid encoding the sTALED with the highest efficiency at the site was transfected into HT1080 cells and HCT116 cells (Day 0) ([Figure S6](#)). Three days after transfection, the cells were harvested. Some were passaged, and some were lysed by incubation at 55 °C for 1 h, and then at 95 °C for 10 min, in 100 µL of cell lysis buffer (50 mM Tris-HCl (Sigma-Aldrich), 1 mM EDTA (Sigma-Aldrich), 0.005% sodium dodecyl sulfate (Sigma-Aldrich), pH 8.0 at 25 °C) supplemented with 5 µL of Proteinase K (Qiagen) in preparation for deep sequencing. From day 3, cells were maintained in medium containing chloramphenicol. Cells were passaged every 3-4 days. An aliquot was obtained every 6-7 days for deep sequencing analysis ([Figure 6B](#)).

Purification of genomic DNA from cultured mammalian cells

After removal of the culture medium, cells were washed with 1X Dulbecco's phosphate-buffered saline (PBS; Welgene), trypsinized, and collected by centrifugation (500g, 4 min, 4° C). The medium containing trypsin was removed and cells were washed with PBS. After removal of the PBS, genomic DNA was isolated using a DNeasy Blood & Tissue Kit (Qiagen) according to manufacturer's instructions.

mtDNA copy number measurement

A method for obtaining the copy number of mtDNA has previously been described (Mok et al., 2020). Quantitative PCR (qPCR) was conducted on a C1000/CFX96 qPCR machine (Bio-rad) using KAPA SYBR® FAST qPCR Kit (Kapa Biosystems) every 3 or 4 days over a 16-day time period. 5 ng of purified DNA was used as template for qPCR in a 20 µL reaction volume. The protocol used was 40 cycles of amplification (10 s at 95°C, 20 s at 60 °C) after an initial heating step of 3 min at 95 °C. The abundance of mtDNA was determined by measuring the ratio of amplified mtDNA gene to genomic DNA (β -actin gene). The ratio was determined using the formula below.

$$\text{Ratio} = E_{\text{mtDNA}} \frac{(C_{q(\text{untreated})} - C_{q(\text{treated})})}{E_{\beta-\text{actin}}} / E_{\text{COX3.1}} \frac{(C_{q(\text{untreated})} - C_{q(\text{treated})})}{E_{\text{RNR2}}} = 2.0066, E_{\beta-\text{actin}} = 2.0038.$$

E is the efficiency of the qPCR reaction; $E_{ND1} = 2.0291$, $E_{ND4} = 1.9926$, $E_{ND5.2} = 1.9901$, $E_{COX3.1} = 2.0211$, $E_{RNR2} = 2.0066$, $E_{\beta-\text{actin}} = 2.0038$.

Growth rate determination

To measure the growth rate of chloramphenicol-resistant single cell-derived clonal populations of cells and wild-type cells in the presence of chloramphenicol, cells were grown in media at different chloramphenicol concentrations. Cells were counted each time the cultures were passaged, which was done every 2-4 days with a split ratio of 1:10.

Targeted deep sequencing

Nested PCR was used to produce libraries for next generation sequencing (NGS). The region of interest was first amplified by PCR using PrimeSTAR® GXL polymerase (Takara). To generate NGS libraries, amplicons were amplified again using TruSeq DNA-RNA CD index-containing primers to label each fragment with adapter and index sequences. Final PCR products were purified using a PCR purification kit (MGmed) and sequenced using a MiniSeq sequencer (Illumina) with a GenerateFASTQ workflow. Base editing frequencies from targeted deep sequencing data were calculated with source code (<https://github.com/ibs-cge/maund>). See also Table S4.

Whole mitochondrial genome sequencing

Because more cells were required for analysis of the whole mitochondrial genome than for analysis of the editing efficiency at defined sites, three samples, transfected under the same conditions as described above, were collected as a single sample (at 3 times the scale). For whole mitochondrial genome sequencing, four steps were required: mtDNA extraction from isolated mitochondria, PCR amplification, NGS library generation, and NGS. First, 3×10^5 HEK 293T cells were trypsinized and collected by centrifugation (500g, 4 min, 4° C) 96 h after transfection. Then, cells were washed with ice-cold PBS (Welgene), and collected again by centrifugation. The supernatant was removed, and the mitochondria were isolated from cultured cells using the reagent-based method of the Mitochondria Isolation Kit for Cultured Cells (Thermo Fisher) according to the manufacturer's protocol. mtDNA was extracted from isolated mitochondria with a DNeasy Blood & Tissue Kit (Qiagen). Second, extracted mtDNA was amplified by PCR using PrimeSTAR® GXL polymerase (Takara). To eliminate primer bias, PCR was performed using two sets of partially overlapping primers (Table S4). About half of the mtDNA was amplified by each set of primers. Then, the PCR products were purified using a PCR purification kit (MGmed). Third, to generate a NGS library from the purified PCR products, we used an Illumina DNA Prep kit with Nextera™ DNA CD Indexes (Illumina). Finally, the libraries were pooled and loaded onto a MiniSeq sequencer (Illumina).

Analysis of mitochondrial genome-wide off-target editing

To analyze NGS data from whole mitochondrial genome sequencing, we referred to a method previously used for analysis of the off-target effects in mitochondrial genome (Lim et al., 2022; Mok et al., 2020). First, we aligned the Fastq files to the GRCh38.p13 (release v102) reference genome using BWA (v.0.7.17), and generated BAM files with SAMtools (v.1.9) by fixing read pairing information and flags. Then, we used the REDItoolDenovo.py script from REDItools (v.1.2.1) to identify the positions with conversion rates $\geq 0.1\%$ among all cytosine, guanines, thymines, and adenines in the mitochondrial genome. We excluded positions with conversion rates $\geq 10\%$ in both treated and untreated samples, regarding these as SNVs in the cell lines. We also excluded the on-target sites for each construct. We considered the remaining positions to be off-target sites and counted the number of edited C/G or A/T nucleotides with an editing frequency $\geq 0.1\%$. We averaged the conversion rates at each base position in the off-target sites to calculate the average C/G to T/A or A/T to G/C editing frequency for all bases in the mitochondrial genome (Figure 7A) as shown below.

$$\text{Average mtDNA - wide C - to - T editing frequency} = \frac{\text{Sum of C/G to T/A off - target}(\%)}{\text{All bases in the mitochondrial genome}}$$

Or

$$\text{Average mtDNA - wide A - to - G editing frequency} = \frac{\text{Sum of A/T to G/C off - target}(\%)}{\text{All bases in the mitochondrial genome}}$$

Mitochondrial genome-wide graphs (Figure 7B) were created by plotting the conversion rates at on-target and off-target sites with an editing frequency $\geq 1\%$ across the entire mitochondrial genome.

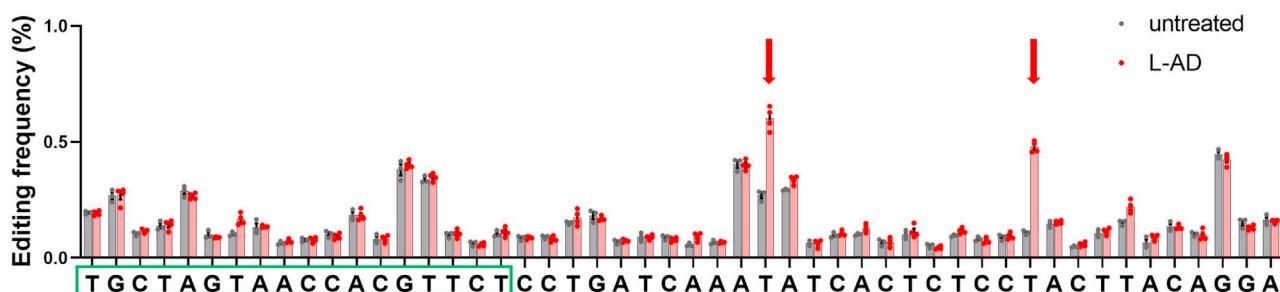
QUANTIFICATION AND STATISTICAL ANALYSIS

No statistical method was used to predetermine the sample sizes. All data are presented as mean and standard error of the mean (SEM).

Supplemental figures

A

Mitochondria: ND4



Mitochondria: ND4

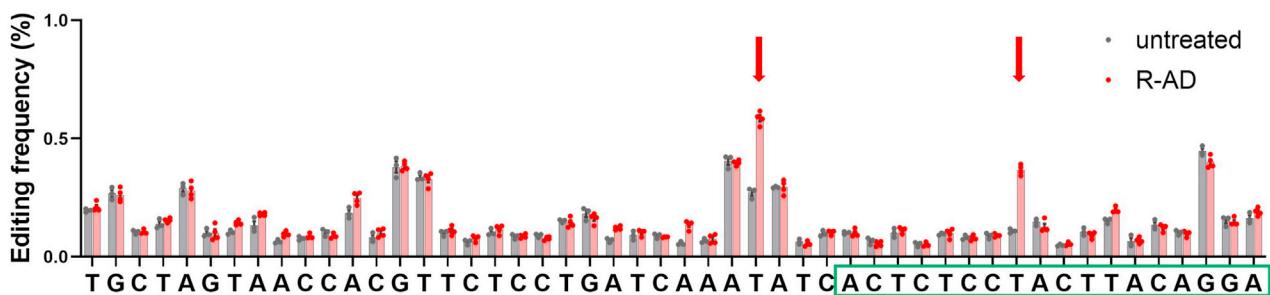
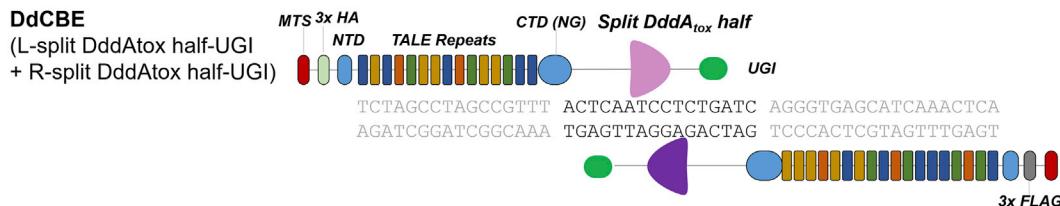


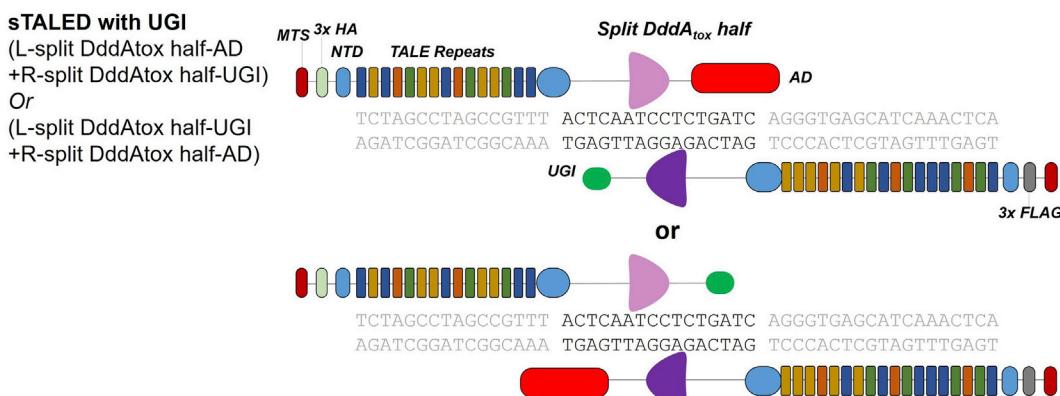
Figure S1. TALE-fused adenine deaminases inducing A-to-G editing at the mitochondrial ND4 site, related to Figure 1

Percentages of total sequencing reads with base conversions induced by TALE-linked adenine deaminases at the ND4 site. TALE-binding sequences are outlined with green boxes. Arrows indicate the positions of A-to-G conversions. Data are shown as means with standard error of the mean (SEM) from n = 4 (L-AD or R-AD) and n = 3 (untreated) biologically independent samples.

A



B



C

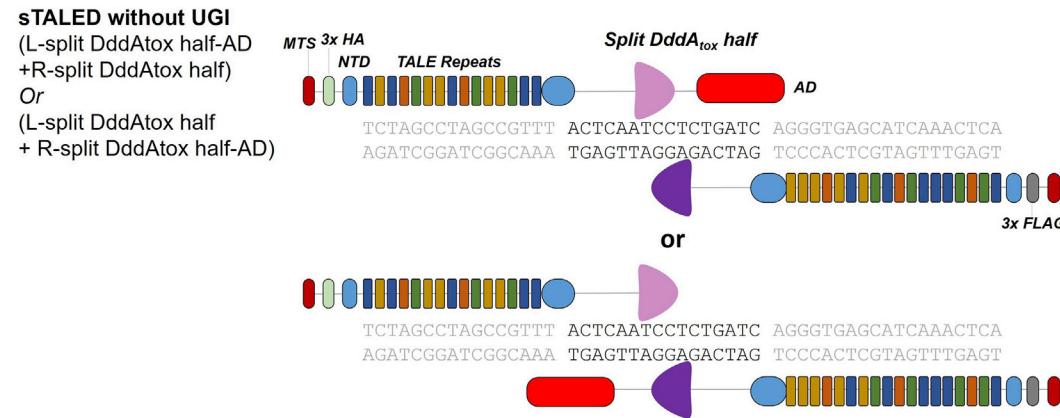


Figure S2. Architecture of DdCBE and sTALED with or without UGI, related to Figure 2

(A) Architecture of DdCBE. The split DddAtox halves and UGI are attached to the C terminus of the TALE array.

(B) Architecture of sTALED with UGI. Adenosine deaminase is linked to the C-terminal end of one half of DddAtox, and UGI is linked to the C-terminal end of the other half.

(C) Architecture of sTALED without UGI. Adenosine deaminase is linked to the C-terminal end of one half of DddAtox, and no UGI is linked to the other half.

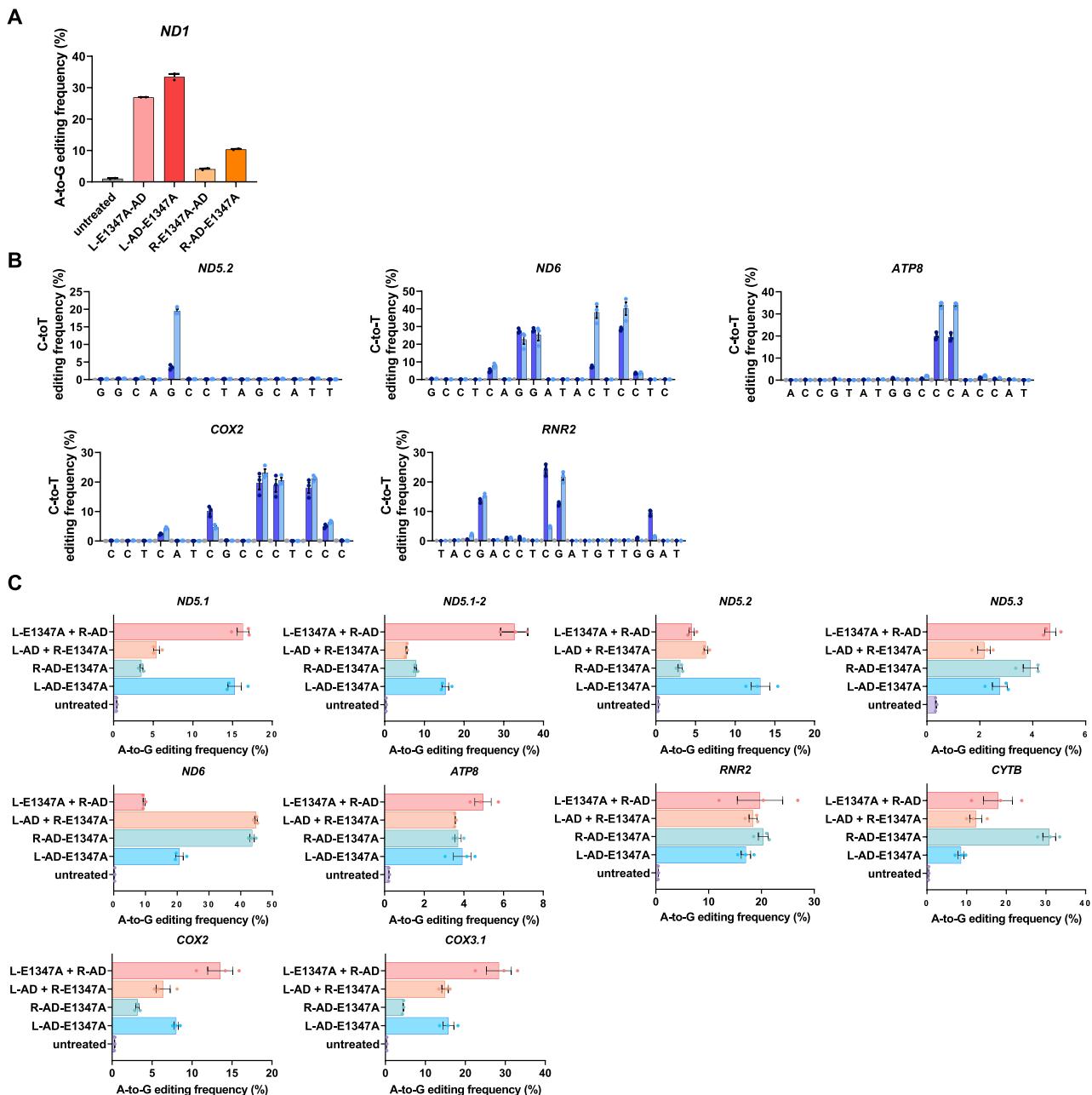


Figure S3. DdCBE, mTALED, and dTALED editing in various mtDNA genes, related to Figure 4

- (A) Comparison of editing frequencies induced at the ND1 site by two different mTALED architectures. Data are shown as means with standard error of the mean (SEM) from n = 2 biologically independent samples.
- (B) Editing frequency of DdCBE at each position in the spacer at 5 sites. Data are shown as means with standard error of the mean (SEM) from n = 3 biologically independent samples.
- (C) Editing frequency of mTALED and dTALED within the spacer at 10 sites. Data are shown as means with standard error of the mean (SEM) from n = 3 biologically independent samples.

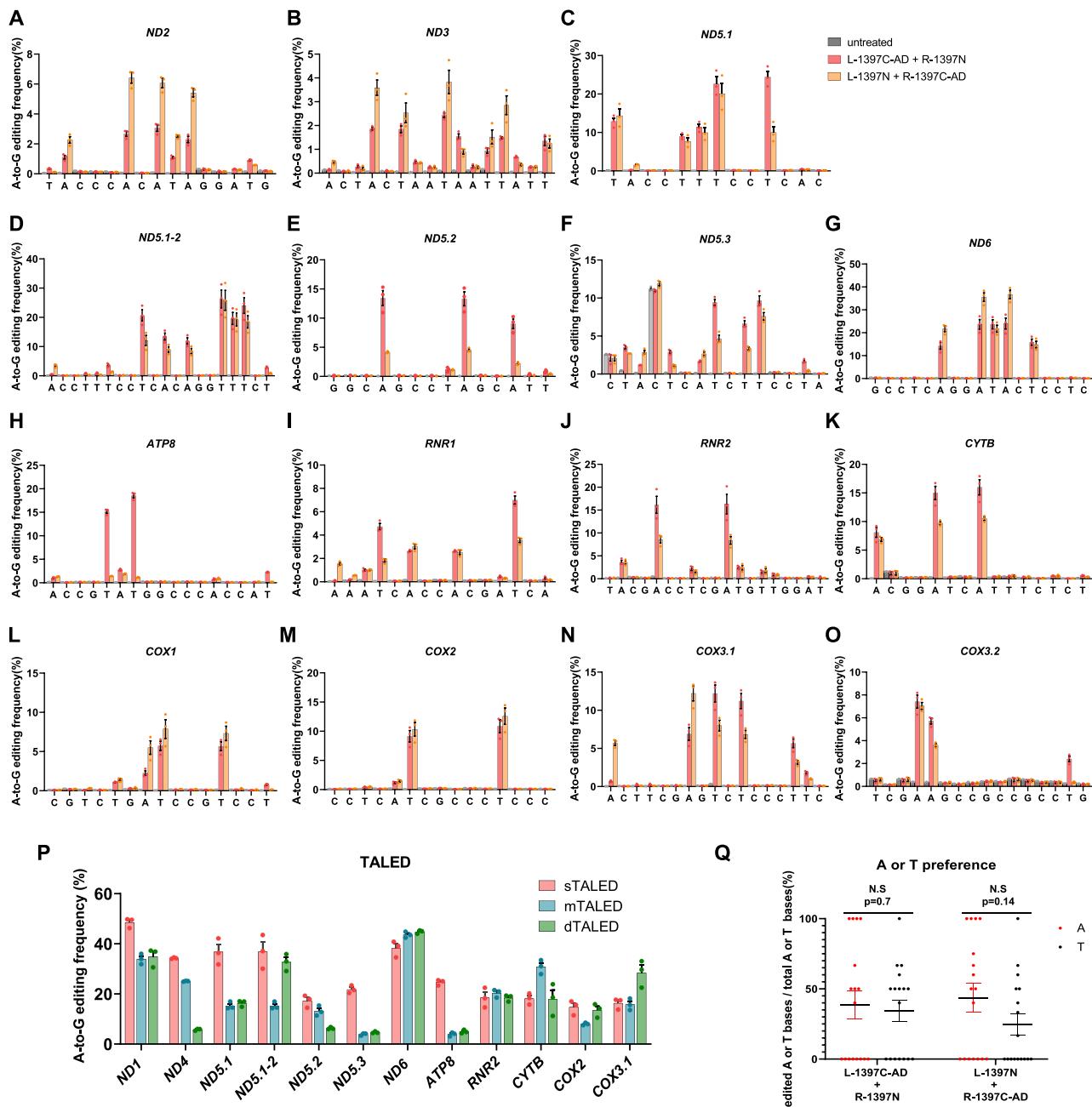


Figure S4. sTALED editing in various mtDNA genes, related to Figure 4D

(A–O) Editing frequency of sTALED at each position in the spacer at 15 sites in mtDNA. Data are shown as means with standard error of the mean (SEM) from $n = 3$ biologically independent samples.

(P) Editing frequency of sTALED, mTALED, and dTALED, all with the same TALE array, at various sites. Data are shown as means with standard error of the mean (SEM) from $n = 3$ biologically independent samples.

(Q) Percentages of A or T bases in the editing window that were respectively converted to G or A, with editing frequencies of at least 5%. Editing frequencies were obtained from a total of 34 sTALED pairs at 17 sites.

A ND4

TGCTAGTAACCA CGTTT CCTGATCAAATTC ACTCTCC TACTTACAGGA
ACGATCATGGTGCAAGA GGACTAGTTTAG TGAGGAGATGAA GTGCCT

	sub (%)		sub (%)
ND4-untreated1	0.3	ND4-mTAELD1	0.3
ND4-untreated2	0.3	ND4-mTAELD2	12.1
ND4-untreated3	0.3	ND4-mTAELD3	1.0
ND4-untreated4	0.3	ND4-mTAELD4	11.8
ND4-untreated5	0.3	ND4-mTAELD5	5.7
ND4-untreated6	0.2	ND4-mTAELD6	0.3
ND4-untreated7	0.2	ND4-mTAELD7	20.0
ND4-sTALED1	0.3	ND4-mTAELD8	0.3
ND4-sTALED2	10.3	ND4-mTAELD9	69.1
ND4-sTALED3	66.5	ND4-mTAELD10	0.2
ND4-sTALED4	0.3	ND4-mTAELD11	76.2
ND4-sTALED5	1.5	ND4-mTAELD12	0.7
ND4-sTALED6	0.3	ND4-mTAELD13	5.5
ND4-sTALED7	0.3	ND4-mTAELD14	14.1
ND4-sTALED8	0.6	ND4-mTAELD15	27.2
ND4-sTALED9	0.3	ND4-mTAELD16	67.1
ND4-sTALED10	2.9	ND4-mTAELD17	2.0
		ND4-mTAELD18	8.1
		ND4-mTAELD19	22.0
		ND4-mTAELD20	30.1

C ND1

TCTAGCCTAGCGTTT ACTCAATCCTCTGATC AGGGTGAGCATCAAACCTCA
AGATCGGATCGGCAA TGAGTTAGGAGACTAG TCCCACCTGTAGTTGAGT

	sub (%)		sub (%)
ND1-untreated 1	0.3	ND1-mTALED 1	0.3
ND1-untreated 2	0.3	ND1-mTALED 2	0.2
ND1-untreated 3	0.2	ND1-mTALED 3	0.4
ND1-untreated 4	0.3	ND1-mTALED 4	0.3
ND1-untreated 5	0.2	ND1-mTALED 5	0.2
ND1-untreated 6	0.2	ND1-mTALED 6	0.2
ND1-untreated 7	0.2	ND1-mTALED 7	0.3
ND1-sTALED 1	0.2	ND1-mTALED 8	46.3
ND1-sTALED 2	0.3		
ND1-sTALED 3	0.3		
ND1-sTALED 4	0.3		
ND1-sTALED 5	0.3		
ND1-sTALED 6	0.3		
ND1-sTALED 7	0.3		
ND1-sTALED 8	0.3		
ND1-sTALED 9	1.1		
ND1-sTALED 10	83.6		
ND1-sTALED 11	0.3		
ND1-sTALED 12	38.9		
ND1-sTALED 13	20.2		

B Human MT-ND4 Ref. TCTGTGCTAGTAACCACGTTCTCTGATCAAATATCACTCTCTACTTACAGGA (H-strand)

Sample	Read AA change (%)
ND4-S3 : 66.5 %	Allele-1 TCTGTGCTAGTAACCACGTTCTCTGATCAAATATCACTCTCTACTTACAGGA S V L V T T F S W S N I T L L L T G - 33.0
	Allele-2 TCTGTGCTAGTAACCACGTTCTCTGATCgAAATATCACTCTCTACTTACAGGA S V L V T T F S W S N I T L L L T G - 28.2
	Allele-3 TCTGTGCTAGTAACCACGTTCTCTGATCgAAATATCACTCTCTACTTACAGGA S V L V T T F S W S N I T L L L T G - 9.9
	Allele-4 TCTGTGCTAGTAACCACGTTCTCTGATCgAAATATCACTCTCTACTTACAGGA S V L V T T F S W S N I T L L L T G - 6.1
	Allele-5 TCTGTGCTAGTAACCACGTTCTCTGATCgAAATATCACTCTCTACTTACAGGA S V L V T T F S W S N I T L L L T G - 2.5
	Allele-6 TCTGTGCTAGTAACCACGTTCTCTGATCgAAATATCACTCTCTACTTACAGGA S V L V T T F S W S N I T L L L T G - 1.9
ND4-Fm9 : 69.1 %	Allele-1 TCTGTGCTAGTAACCACGTTCTCTGATCAAATATCACTCTCTACTTACAGGA S V L V T T F S W S N I T L L L T G - 29.2
	Allele-2 TCTGTGCTAGTAACCACGTTCTCTGATCgAAATATCACTCTCTACTTACAGGA S V L V T T F S W S N I T L L L T G - 18.4
	Allele-3 TCTGTGCTAGTAACCACGTTCTCTGATCgAAATATCACTCTCTACTTACAGGA S V L V T T F S W S N I T L L L T G - 5.7
	Allele-4 TCTGTGCTAGTAACCACGTTCTCTGATCgAAATATCACTCTCTACTTACAGGA S V L V T T F S W S N I T L L L T G - 3.5
	Allele-5 TCTGTGCTAGTAACCACGTTCTCTGATCgAAATATCACTCTCTACTTACAGGA S V L V T A F S W S N I T L L L T G T385A 3.1
	Allele-6 TCTGTGCTAGTAACCACGTTCTCTGATCgAAATATCACTCTCTACTTACAGGA S V L V T T F S W S N V T L L L T G - 2.7
ND4-Fm9 : 76.2 %	Allele-1 TCTGTGCTAGTAACCACGTTCTCTGATCAAATATCACTCTCTACTTACAGGA S V L V T T F S W S N I T L L L T G - 21.8
	Allele-2 TCTGTGCTAGTAACCACGTTCTCTGATCgAAATATCACTCTCTACTTACAGGA S V L V T T F S W S N I T L L L T G - 18.7
	Allele-3 TCTGTGCTAGTAACCACGTTCTCTGATCgAAATATCACTCTCTACTTACAGGA S V L V T T F S W S N I T L L L T G - 4.3
	Allele-4 TCTGTGCTAGTAACCACGTTCTCTGATCgAAATATCACTCTCTACTTACAGGA S V L V T T F S W S N I T L L L T G - 4.2
	Allele-5 TCTGTGCTAGTAACCACGTTCTCTGATCgAAATATCACTCTCTACTTACAGGA S V L V T T F S W S N I T L L L T G T385A 3.6
	Allele-6 TCTGTGCTAGTAACCACGTTCTCTGATCgAAATATCACTCTCTACTTACAGGA S V L V T V F S W S N I T L L L T G T385A 3.4
ND4-Fm9 : 67.1 %	Allele-1 TCTGTGCTAGTAACCACGTTCTCTGATCAAATATCACTCTCTACTTACAGGA S V L V T T F S W S N I T L L L T G - 30.9
	Allele-2 TCTGTGCTAGTAACCACGTTCTCTGATCgAAATATCACTCTCTACTTACAGGA S V L V T T F S W S N I T L L L T G - 19.6
	Allele-3 TCTGTGCTAGTAACCACGTTCTCTGATCgAAATATCACTCTCTACTTACAGGA S V L V T T F S W S N I T L L L T G T385A 4
	Allele-4 TCTGTGCTAGTAACCACGTTCTCTGATCgAAATATCACTCTCTACTTACAGGA S V L V T T F S W S N I T L L L T G - 4
	Allele-5 TCTGTGCTAGTAACCACGTTCTCTGATCgAAATATCACTCTCTACTTACAGGA S V L V T T F S W S N I T L L L T G - 3.9
	Allele-6 TCTGTGCTAGTAACCACGTTCTCTGATCgAAATATCACTCTCTACTTACAGGA S V L V T T F S W S N I T L L L T G - 2.8

D Human MT-ND1 Ref. TCTAGCCTAGCGTTTACTCAITCTCTGATCAGGGTGAGCATCAAACCTCA (H-strand)

Sample	AA change	Read (%)
ND1-S10 : 83.6 %	Allele-1 TCTAGCCTAGCGTTTACTCAITCTCTGATCAGGGTGAGCATCAAACCTCA S S L A V Y S I L W S G W A S N S - 15.4	
	Allele-2 TCTAGCCTAGCGTTTACTCAACCTCTGATCAGGGTGAGCATCAAACCTCA S S L A V Y S T L W S G W A S N S I116T 15.2	
	Allele-3 TCTAGCCTAGCGTTTACTCAACCTCCGATCAGGGTGAGCATCAAACCTCA S S L A V Y S I P W S G W A S N S L117P 12.9	
	Allele-4 TCTAGCCTAGCGTTTACTCAACCTCCGATCAGGGTGAGCATCAAACCTCA S S L A V Y S I P W S G W A S N S I116T,L117P 12.0	
	Allele-5 TCTAGCCTAGCGTTTACTCAACCTCCGATCAGGGTGAGCATCAAACCTCA S S L A V Y S T P R S G W A S N S I116T,L117P,W118R 10.3	
	Allele-6 TCTAGCCTAGCGTTTACTCAACCTCCGATCAGGGTGAGCATCAAACCTCA S S L A V Y S I P R S G W A S N S I116T,W118R 9.3	
	Allele-7 TCTAGCCTAGCGTTTACTCAACCTCCGATCAGGGTGAGCATCAAACCTCA S S L A V Y S I L R S G W A S N S W118R 9.1	
	Allele-8 TCTAGCCTAGCGTTTACTCAACCTCCGATCAGGGTGAGCATCAAACCTCA S S L A V Y S I P R S G W A S N S L117P,W118R 6.8	

Figure S5. Single-cell-derived clones containing mtDNA edits, related to Figure 4

- (A) The editing frequency in clones obtained through single cell expansion after treatment with ND4-targeted sTALED or mTALED. The TALE-binding site is indicated in green, and the target adenines previously identified with targeted deep sequencing are indicated in red.
- (B) Allele analysis in single-cell-derived clones that exhibited high frequencies of base editing in ND4. The H strand and the encoded amino acid (AA) sequence are shown. The TALE-binding site is indicated in green and base changes are shown in red. Only the top 6 sequence reads among hundreds of edited ones are shown.
- (C) The editing frequency in clones obtained through single cell expansion after treatment with ND1-targeted sTALED or mTALED. The TALE-binding site is indicated in green, and the target adenines previously identified with targeted deep sequencing are indicated in red.
- (D) Allele analysis in single cell-derived clones that exhibited high frequencies of base editing in ND1. The H strand and the encoded AA sequence are shown. The TALE-binding site is indicated in green and base changes are shown in red. Only the top 8 sequence reads among hundreds of edited ones are shown.

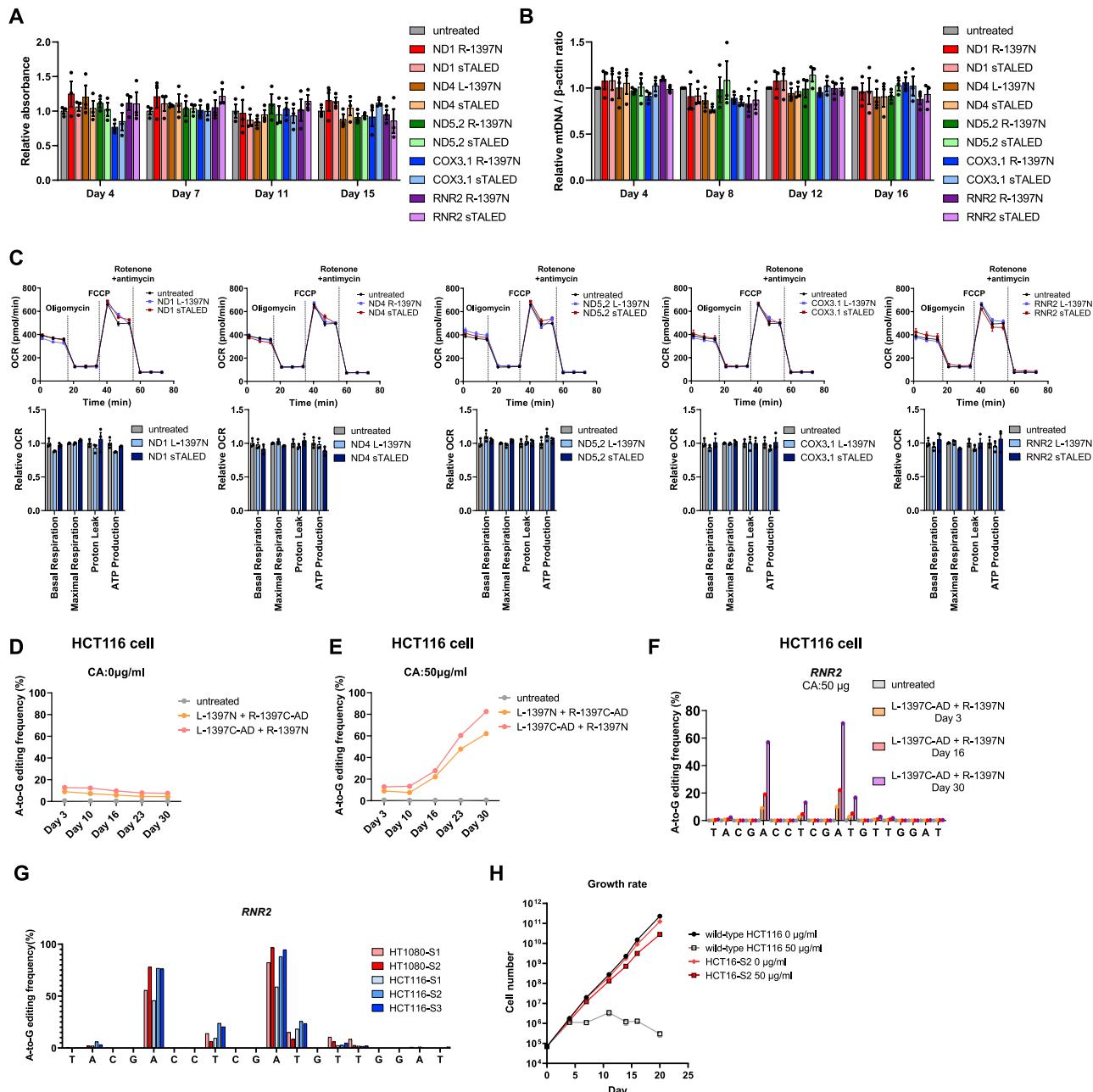
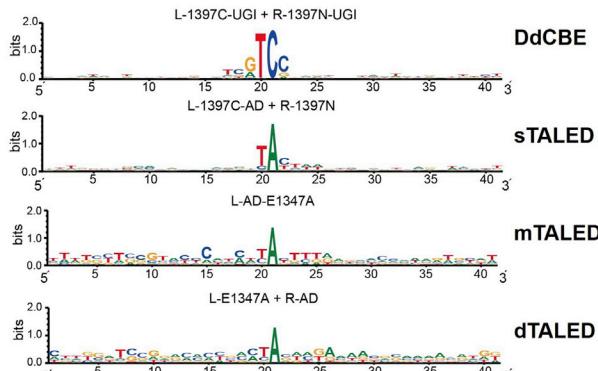


Figure S6. Characterization of cells containing TALED-generated mutant mtDNA, related to Figure 6

- (A) Cell viability was measured by recording the color change caused by the generation of formazan at the indicated time points. Absorbance values were normalized to values from untreated cells. Data are shown as means with standard error of the mean (SEM) from n = 3 biologically independent samples.
- (B) The mtDNA copy number was measured by qPCR at the indicated time points. mtDNA levels in edited cells were normalized to values from untreated cells. Data are shown as means with standard error of the mean (SEM) from n = 3 biologically independent samples.
- (C) The oxygen consumption rate (OCR) of cells treated with sTALEDs targeting ND1, ND4, ND5.2, COX3.1, and RNR2 (above). FCCP, carbonyl cyanide-4-(trifluoromethoxy)phenylhydrazone. Each respiratory parameter was normalized to the value from untreated cells (below). Data are shown as means with standard error of the mean (SEM) from n = 3 biologically independent samples.
- (D) Editing frequencies at day 3, 10, 16, 23, and 30 post-transfection in HCT116 cells cultured in 0 μ g/mL chloramphenicol.
- (E) Editing frequencies at day 3, 10, 16, 23, and 30 post-transfection in HCT116 cells cultured in 50 μ g/mL chloramphenicol.
- (F) Editing frequencies at each position in the spacer at day 3, 16, and 30 post-transfection in HCT116 cells cultured in 50 μ g/mL chloramphenicol.
- (G) Editing frequencies at each position in the spacer in single cell-derived clones with a genetic mutation in the RNR2 site induced by sTALED.
- (H) Growth rate at different chloramphenicol concentrations in single-cell-derived clones with a high mutation frequency in the RNR2 site and wild-type cells. Data are shown as means with standard error of the mean (SEM) from n = 2 biologically independent samples.

A



B

mtDNA ND4

TGCTAGTAACCACGTTCT CCT₃GA₅T₆CA₈A₉T₁₀A₁₁A₁₂T₁₃C ACTCTCTACTTACAGGA

ACGATCATTGGTGCAAGA GGA CT A GT T T A T A G TGAGAGGATGAATGTCCT

**Nuclear MTND4
Pseudogene (Chr5)**

TGCTAGTAACCACATTCT CCT₃GA₅T₆CA₈A₉T₁₀A₁₁A₁₂T₁₃C ACTCTCTACTTACAGGA

ACGATCATTGGTGCAAGA GGA CT A GT T T A T A G TGAGAGGATGAATGTCCT

mtDNA ND5.1-2

TAGCATTAGCAGGAAT A₁CCT₄T₅T₆CCT₅CA₁₁CA₁₃GGT₁₆T₁₇T₁₈CT₂₀ ACTCCAAAGACCACATCA

ATCGTAATCGTCCTTA T GGA A A GGA GT GT CCA A A GA TGAGGTTCTGGTGTAGT

**Nuclear MTND5.1-2
Pseudogene (Chr5)**

TAGCATTGGCAGGAAT A₁CCC T₅T₆CCT₅CA₁₁CA₁₃GGT₁₆T₁₇T₁₈CT₂₀ ACTCCAAAGACCACATCA

ATCGTAATCGTCCTTA T GGG A A GGA GT GT CCA A A GA TGAGGTTCTGGTGTAGT

mtDNA ND5.2

TTCAACCTCCCTCACCATT GGCA₄G CCT₈A₉GCA₁₂T₁₃T₁₄ AGCAGGAATACCTTCTCA

AAGTTGGAGGGAGTGGTAA CGGT C GGA T CGT A A TCGTCCTATGGAAAGGAGT

**Nuclear MTND5.2
Pseudogene(Chr5)**

TTCAACCTCCCTCACCATT GGCA₄A₅CCT₈A₉GCA₁₂T₁₃T₁₄ GGCAGGAATACCTTCTCA

AAGTTGGAGGGAGTGGTAA CGGT T GGA T CGT A A CCGTCCTATGGAAAGGAGT

mtDNA ND6

TCAACCCCTGACCCCCAT GCCT₄CA₆GG₉T₁₀A₁₁CT₁₃CCT₁₆C AATAGCCATCGCTGTAGTA

AGTTGGGACTGGGGTA CGGA GT CCT A T GA GGA G TTATCGGTAGCGACATCAT

**Nuclear MTND6
Pseudogene(Chr18)**

TCAACCCCTGACCCCCAT GCCT₄CA₆GG₉T₁₀A₁₁CT₁₃CCT₁₆C AATAGCCATCGCTGTAGTG

AGTTGGGACTGGGGTA CGGA GT CCT A T GA GGA G TTATCGGTAGCGACATCAT

mtDNA ND6

TCAACCCCTGACCCCCAT GCCT₄CA₆GG₉T₁₀A₁₁CT₁₃CCT₁₆C AATAGCCATCGCTGTAGTA

AGTTGGGACTGGGGTA CGGA GT CCT A T GA GGA G TTATCGGTAGCGACATCAT

**Nuclear MTND6
Pseudogene(Chr5)**

TCAACCCCTGACCCCCAT GCCT₄CA₆GG₉T₁₀A₁₁CT₁₃CCT₁₆C AATAGCCACCGCTGTAGTA

AGTTGGGACTGGGGTA CGGA GT CCT A T GA GGA G TTATCGGTAGCGACATCAT

mtDNA CYTB

TCGGGGAGGCCCTATATT A₁CGGA₅T₆CA₈T₉T₁₀T₁₁CT₁₃C T₁₅ ACTCAGAACCTGAAACA

AGCCCGCTCGGATATAA T GCCT A GT A A A GA G A TGAGTCTTGACTTTGT

**Nuclear MTCYTB
Pseudogene(Chr5)**

TCGGGGAGGCCCTATATT A₁CGGC T₆CA₈T₉T₁₀T₁₁CT₁₃T₁₄T₁₅ ACCTAGAACCTGAAACA

AGCCGGCTCGGATATAA T GCCG A GT A A A GA A A TGGATCTTGACTTTGT

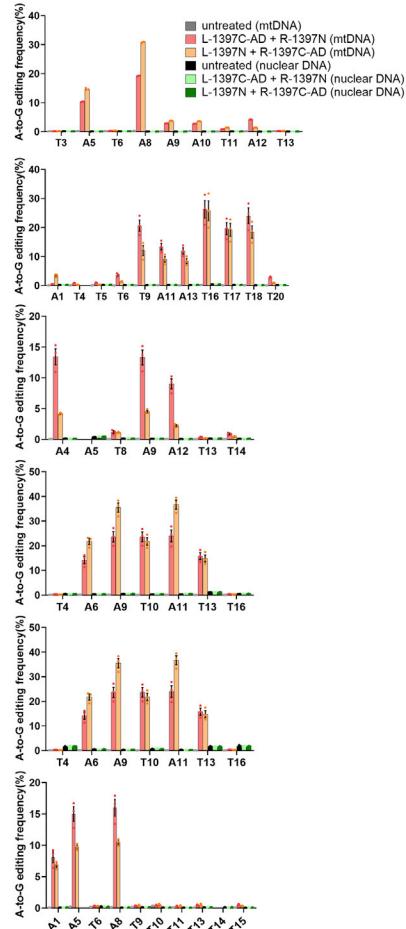


Figure S7. TALED off-target sites in mtDNA and genomic DNA, related to Figure 7

(A) Sequence logos obtained using DNA sequences at C-to-T and A-to-G off-target edits induced with frequencies of $\geq 1\%$ by the ND1-specific DdCBE and TALEDs.

(B) Editing frequencies at potential off-target sites in the nuclear genome. Data are shown as means with standard error of the mean (SEM) from n = 3 biologically independent samples.