

# Optimized high-fidelity 3DPCR to assess potential mitochondrial targeting by activation-induced cytidine deaminase

Haiyan Wu<sup>1,2</sup>, Kaili Zhang<sup>1,3</sup>, Yue Chen<sup>3</sup>, Jinfeng Li<sup>2</sup>, Matthew P. Strout<sup>4,\*</sup> and Xiwen Gu<sup>1</sup> 

1 Key Laboratory of Shaanxi Province for Craniofacial Precision Medicine Research, Research Center of Stomatology, Xi'an Jiaotong University College of Stomatology, Xi'an, China

2 Department of Oral and Maxillofacial Surgery, Xi'an Jiaotong University College of Stomatology, Xi'an, China

3 Department of Periodontology and Oral Medicine, Xi'an Jiaotong University College of Stomatology, Xi'an, China

4 Section of Hematology, Yale University School of Medicine, New Haven, CT, USA

## Keywords

3DPCR; activation-induced cytidine deaminase; mitochondria; mtDNA

## Correspondence

J. Li, Department of Oral and Maxillofacial Surgery, Xi'an Jiaotong University College of Stomatology, Xi'an, Shaanxi, China  
E-mail: jinfeng@mail.xjtu.edu.cn

M. P. Strout, Section of Hematology, Yale University School of Medicine, New Haven, CT, USA

E-mail: matthew.p.strout@gmail.com

X. Gu, Key Laboratory of Shaanxi Province for Craniofacial Precision Medicine Research, Research Center of Stomatology, Xi'an Jiaotong University College of Stomatology, Xi'an, Shaanxi, China  
E-mail: xiwen.gu@xjtu.edu.cn

## \*Present address

Alexion Pharmaceuticals, New Haven, CT, USA

(Received 17 December 2019, revised 10 May 2020, accepted 3 June 2020)

doi:10.1002/2211-5463.12927

Activation-induced cytidine deaminase (AID) initiates somatic hypermutation and class switch recombination of immunoglobulin genes in B cells, whereas off-targeted AID activity contributes to oncogenic mutations and chromosomal translocations associated with B cell malignancies. Paradoxically, only a minority of AID is allowed to access the nuclear genome, but the majority of AID is retained in the cytoplasm. It is unknown whether cytoplasmic AID can access and target the mitochondrial genome [mitochondrial DNA (mtDNA)]. To address this issue, we developed high-fidelity differential DNA denaturation PCR, which allowed the enrichment of genuine mtDNA mutations and therefore the identification of endogenous mtDNA mutation signatures *in vitro*. With this approach, we showed that AID targeting to mtDNA is a rare event in AID-expressing lymphoma lines. Further biochemical and microscopic analysis revealed that a fraction of cytosol AID is associated with the outer membrane of mitochondria but unable to access the mitochondrial matrix. Together, our data suggested that the mitochondrial genome is protected from AID-mediated mutagenesis by physical segregation of AID from accessing mtDNA within the mitochondrial matrix.

Activation-induced cytidine deaminase (AID) is a single-strand DNA deaminase with versatile and crucial functions. In B lymphocytes, AID initiates somatic hypermutation and class switch recombination of immunoglobulin (Ig) genes via deamination of cytosine (C) into uracil (U). The resulting U:G (guanine)

mismatches are subsequently processed by downstream base excision and DNA mismatch repair pathways to yield DNA mutations and strand lesions [1,2]. Promiscuous off-targeting AID activity has been shown to cause oncogenic point mutations, DNA double-strand breaks and chromosome translocations associated with

## Abbreviations

3DPCR, differential DNA denaturation PCR; AID, activation-induced cytidine deaminase; Ig, immunoglobulin; mtDNA, mitochondrial DNA.

B cell malignancies [3–6]. In addition, inflammatory response can trigger ectopic expression of AID in non-B cells, and AID-associated mutagenesis has been implicated in the pathogenesis of pancreatic, gastric, colon and skin cancers [7–10]. Furthermore, several studies have found that AID can function as a critical epigenetic regulator in a great variety of cell types, including B cells, hematopoietic stem cells and embryonic stem cells [11–15].

Restraining AID access to the nucleus has emerged as one major mechanism to curb deleterious AID activity. For instance, nuclear AID has been shown to be susceptible to proteasomes reactivator REG-gamma-mediated degradation or is subjected to CRM1-dependent active nuclear export [16,17]. As a result, AID targeting of nuclear DNA is transient, and the majority of AID is localized within the cytoplasm [18–20]. It remains unknown whether cytoplasmic AID plays a role other than sequestering AID from the nuclear genome. In this regard, mitochondria are cytoplasmic organelles central for oxidative phosphorylation, fatty acid oxidation and apoptosis signaling. Despite major mitochondrial proteins being encoded by nuclear DNA, mitochondria contain their own mitochondrial genome [mitochondrial DNA (mtDNA)] within the mitochondrial matrix. The circular mtDNA in human consists of 16 569 bp and encodes 13 key components of the mitochondrial respiratory chain complex. It is unknown whether mtDNA is targeted by or protected from AID-mediated mutagenesis.

Here, we developed high-fidelity differential DNA denaturation PCR (3DPCR) to investigate mtDNA mutations in AID-expressing lymphoma cell lines and revealed little evidence for AID targeting of mtDNA. Further biochemical analysis revealed that a fraction of AID is associated with the outer membrane of mitochondria but is protected from accessing the mitochondrial matrix and mtDNA within it, suggesting that mtDNA is protected from AID-mediated mutagenesis by physical segregation.

## Materials and methods

### Cell culture

Human Burkitt lymphoma cell line Ramos [21] and Ramos-A23 (a Ramos-derived line with constitutive high expression of AID driven by overexpression construct) were cultured in RPMI-1640 (BE04-558F; Lonza, Allendale, NJ, USA) supplemented with 10% FBS (HyClone, Cramlington, UK), 2 mM L-glutamine and 1% penicillin/streptomycin (10378016; Thermo, Waltham, MA, USA). Human chronic B cell leukemia line Mec1 was cultured in Iscove's

modified Dulbecco's medium (12-722F; Lonza) supplemented with 10% FBS, 2 mM L-glutamine and 1% penicillin/streptomycin. Chicken DT40 and DT40AID<sup>-/-</sup> cells [22] were cultured in RPMI-1640 supplemented with 10% FBS, 1% chicken serum (C5405; Sigma, St. Louis, MO, USA), 2 mM L-glutamine, 1% penicillin/streptomycin and 0.1 mM  $\beta$ -mercaptoethanol. 293T cells were cultured in Dulbecco's modified Eagle's medium (12-604F; Lonza) supplemented with 10% FBS, 2 mM L-glutamine and 1% penicillin/streptomycin. All cultures were maintained at 37 °C in a humidified tissue culture incubator with 5% CO<sub>2</sub>.

### Retroviral transfection and GFPstop reporter assay

293T cells were cotransfected with the pKAT2 retroviral packaging plasmid and a viral vector encoding an inactive GFP (MMLV-GFPstop-IRES-Puro) [23] using Lipofectamine 3000 reagent (Invitrogen, Waltham, MA, USA) following the manufacturer's protocol. Two days posttransfection, viral supernatant was harvested and added to Ramos, Ramos-A23 and Mec1 cells, followed by centrifugation at room temperature (420 g for 45 min). Forty-eight hours postinfection, retrovirus-transduced cells were selected for 3 weeks with optimal concentration of puromycin (0.25  $\mu\text{g}\cdot\text{mL}^{-1}$  for Ramos and Ramos-A23 cells and 1  $\mu\text{g}\cdot\text{mL}^{-1}$  for Mec1 cells). To detect GFP reversion mutants, we resuspended  $1 \times 10^6$  transduced or control cells in PBS and evaluated for GFP expression with a FACSCalibur flow cytometer using CELLQUEST software (BD, Franklin Lakes, NJ, USA).

### Immunofluorescence analysis

Where indicated, cells were stained with 200 nM MitoTracker Red (Invitrogen) in RPMI-1640 for 15 min at 37 °C. Cells were then fixed and permeabilized at -20 °C with prechilled 100% methanol for 10 min. After three washes with PBS, nonspecific antibody binding was blocked with 5% goat serum in PBS. The cells were then incubated for 1 h with mouse anti-AID IgG1 (1 : 100; 392500; Invitrogen) primary antibody. After three washes with PBS, the cells were incubated with Alexa Fluor 488-conjugated goat anti-(mouse IgG1) (1 : 200; Invitrogen) secondary antibodies for 1 h. Nuclei were counterstained with DAPI-containing mounting medium (Vector Shield). Immunofluorescence images were captured using a Nikon eclipse-TI confocal microscopy (Perkin Elmer, Waltham, MA, USA).

### Mitochondria treatment and western blot analysis

Mitochondria were purified from  $2 \times 10^7$  cultured cells using Qproteome mitochondria isolation kit (Qiagen,

Hilden, Germany) following the manufacturer's protocol, which allows simultaneous isolation of cytosol and mitochondrial proteins. For proteinase digestion, isolated mitochondria were incubated with 5–50  $\mu\text{g}\cdot\text{mL}^{-1}$  trypsin on ice for 30 min. Samples were separated by SDS/PAGE, transferred to poly(vinylidene difluoride) membranes and subsequently probed with mouse anti-AID IgG1 (1 : 1000; 392500; Invitrogen), rabbit anti-VDAC (Voltage-dependent anion channel) Ig (1 : 1000; 4866; Cell Signaling, Danvers, MA, USA), rabbit anti- $\beta$ -Tubulin Ig (1 : 2000; sc-9104; Santa Cruz) and rabbit anti-Hsp60 Ig (1 : 1000; 4870; Cell Signaling) primary antibodies, followed by corresponding goat anti-(rabbit Ig) (1 : 5000; SC-2004; SCBT, Santa Cruz, CA, USA) and rabbit anti-(mouse Ig) (1 : 10 000; AB\_2340074; Jackson Immuno, West Grove, PA, USA) secondary antibodies coupled to horseradish peroxidase according to the manufacturer's protocol.

### 3DPCR error rate

An 897-bp Cytb product was amplified from Ramos cells, cloned into pMD19-T vector (pMD19-Cytb) and used as a template for evaluating 3DPCR error spectrum. For Taq-based two-step 3DPCR, the first-round standard PCR was amplified using 1 U Taq polymerase (Takara, Dalian, China), 1 ng pMD19-Cytb template and M13 universal primers (M13F, 5'-TGTAACACGACGGCCAGT-3'; M13R, 5'-CAGGAAACAGCTATGACC-3'), according to the following program: 95 °C for 5 min, 30 cycles at 95 °C for 30 s, 53 °C for 30 s, 72 °C for 60 s and a final extension step of 72 °C for 10 min. The second-round 3DPCR was amplified using 1 U Taq, 1/50 of the first-round product and Cytb primers (CytbF, 5'-AACCATCGTTGTATTTCAAC-3'; CytbR, 5'-TAGTTTGTAGGGACGGATC-3'), according to the following program: 95 °C for 5 min, 35 cycles at 85–90 °C for 30 s, 53 °C for 30 s, 72 °C for 60 s and a final extension step of 72 °C for 10 min. Taq-based single 3DPCR was performed following the same second-round 3DPCR program except that 1 ng pMD19-Cytb DNA was used directly as a template. Similarly, Phusion-based single 3DPCR was performed using 1 ng pMD19-Cytb DNA and Cytb primers, but with 1 U high-fidelity Phusion polymerase (NEB), according to the following program: 95 °C for 5 min, 35 cycles at 85–90 °C for 30 s, 58 °C for 30 s, 72 °C for 60 s and a final extension step of 72 °C for 10 min. For 3DPCR, only the products amplified with the least denaturation temperature were subjected to cloning and sequencing.

### Amplification, cloning and analysis of mtDNA mutations

Subclones of Ramos and Ramos-A23 were generated by serial dilution in 96-well plates and further expanded for at least 2 weeks before sequencing. Genomic DNA and

mtDNA were prepared by DNeasy blood and tissue kit (Qiagen) from cultured cells and isolated mitochondria, respectively. For analysis of mutations, all of the amplifications were performed with Phusion polymerase (NEB). Phusion-based 3DPCR amplification of Cytb has been described, except using mtDNA as a template. The PCR products of interest were purified with QIAquick Gel Extraction Kit (Qiagen), A-tailed with Taq and cloned into TOPO (Thermo) or pMD19-T vector (Takara) for bidirectional sequencing. The sequences were checked for quality and nucleotide changes using SEQUENCHER software (Gene Codes, Ann Arbor, MI, USA). Primers used for analyzing mtDNA in human lymphoma cells were: D-loop (D-loopF, 5'-GAAAACAAAATACTCAAATGGGCC-3'; D-loopR, 5'-GCTGTGCAGACATTCAATTGTTATT-3'), 12s rRNA (12sF, 5'-CCCATCTACCCAGCACACA-3'; 12sR, 5'-CTGGTTCGTCCTCAAGTGCACACT-3'), 16s rRNA (16sF, 5'-GTAAACCTAGCCCCAAACC-3'; 16sR, 5'-GTTGGGTCTGTCTCCGAGGT-3'); primers for analyzing IgH variable region in Ramos (VH4\_Ldr, 5'-ATGAAACACCTGTGGTTCTT-3'; JH\_Cons, 5'-CTTACCTGAGGAGACGGTGACC-3') [24]; and primers for analyzing Cytb in chicken DT40 (chkCytbF, 5'-ATGGCACCCAACATTTCGAAA-3'; chkCytbR, 5'-CGATTGTGGGGAAGAGGATA-3').

### Data analysis and statistics

Analysis of mutation frequency and hotspot enrichment was performed as previously described [3,25]. Statistical comparison of mutation frequencies was performed using Fisher's exact test (one-tailed), and the significance of AID hotspot and CpG motif enrichment was assessed using the binomial test. Welch's two-sample *t*-test was used to compare independent means. All statistical analyses were performed using R software. *P* values <0.05 were considered significant.

## Results

### AID mutational activities in mec1, Ramos and Ramos-A23 cells

To determine whether AID can target mtDNA, we used multiple B cell lines with constitutive AID expression, including human chronic lymphocytic leukemia line Mec1, human Burkitt lymphoma line Ramos and Ramos-A23, a Ramos derivative that expresses high levels of AID [21,26]. We also used chicken lymphoma line DT40 and its derivative carrying targeted deletion of AID (DT40 AID<sup>-/-</sup>) [22]. Immunofluorescence analysis confirmed the absence of AID in DT40 AID<sup>-/-</sup> cells, moderate expression of AID in Ramos cells and abundant expression of AID in DT40, Mec1 and Ramos-A23 cells (Fig. 1A). To confirm that AID can target nuclear DNA in these cells, we used a GFPstop

reporter construct, which encodes an inactive GFP that can regain fluorescence upon AID-mediated reversion of a premature stop codon, TAG [23]. We delivered the GFPstop reporter into Mec1, Ramos and Ramos-A23 cells via retroviral infection, and measured the occurrence of GFP<sup>+</sup> cells by flow cytometry (Fig. 1B,C). Although the protein level of AID is high, Mec1 cells exhibited the lowest AID mutational activity, which is consistent with the fact that the Mec1 line is derived from leukemic cells carrying unmutated Ig, possibly because of the high-fidelity processing of AID-induced C-to-U mutations in this line [26]. In comparison, Ramos and Ramos-A23 cells exhibited much higher AID mutational activity than Mec1 cells, with the highest mutational activity observed in Ramos-A23 cells, which correlates well with AID dosage and also is consistent with previous reports that Ramos cells constitutively mutate their Ig [21].

**Few mtDNA mutations in Mec1, Ramos and Ramos-A23 cells**

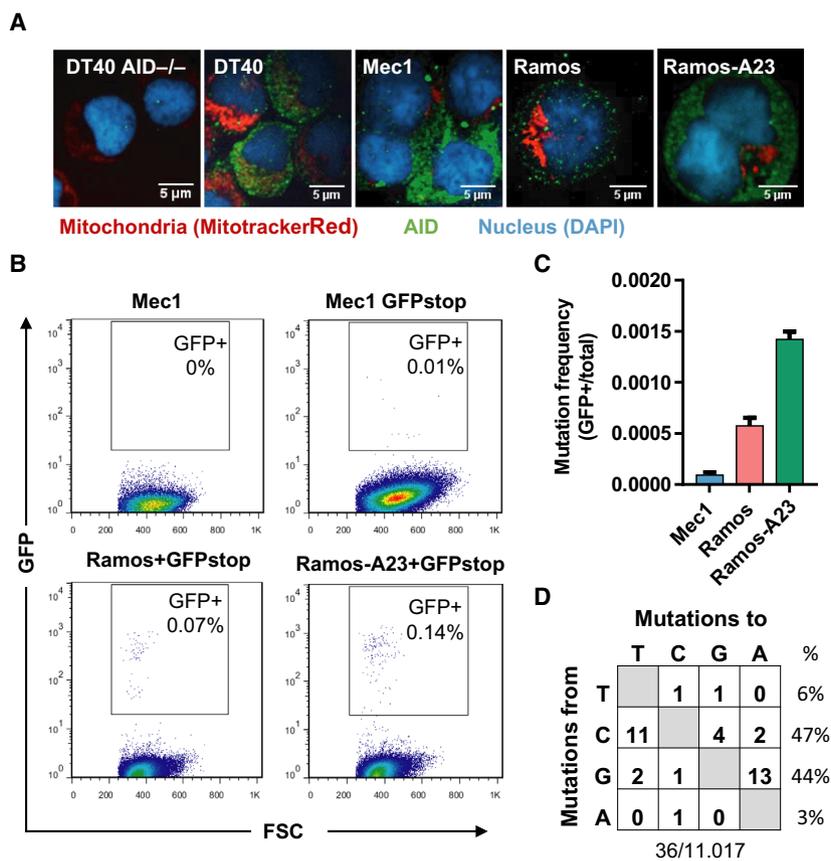
We next screened different mtDNA regions (D-loop, 12s rRNA, 16s rRNA and Cytb gene) for mutation using mtDNA from purified mitochondria of Mec1, Ramos

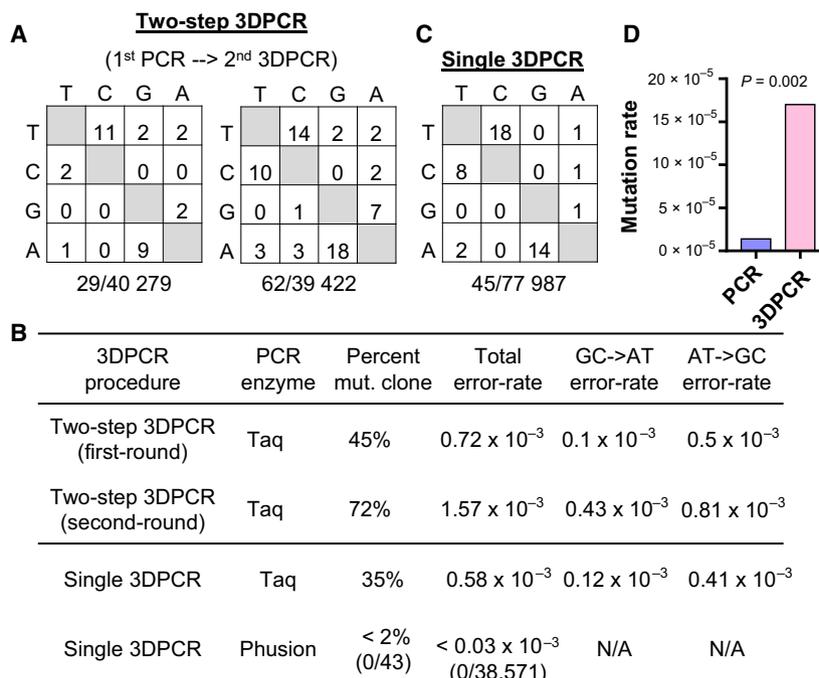
and Ramos-A23 cells. For each region, an ~1-kb segment was amplified, cloned and sequenced for point mutations, with around 40 kb nucleotides of sequence information collected for each gene. We identified dozens of preexisting single-nucleotide polymorphisms (14 in Ramos and Ramos-A23 and 21 in Mec1 cells; Table S1) that are clonal and differed from the mitochondrial reference sequence (rCRS) [27]. However, we identified only six nonclonal point mutations (C16223G and C33T in Ramos, C1352T and C15298T in Ramos-A23, and C144T and A16524G in Mec1) in a total of 473 504 sequenced nucleotides ( $1.27 \times 10^{-5}$  mutations/bp). In contrast, sequencing of the 0.5-kb IgH variable fragment, the physiological target of AID in Ramos cells, revealed 36 point mutations in 11 017 sequenced nucleotides ( $3.3 \times 10^{-3}$  mutations/bp; Fig. 1D). This surprisingly low level of mtDNA mutation in cell lines with different AID expression and mutational activity argued against a significant role for AID targeting of mtDNA.

**Development of high-fidelity 3DPCR to detect rare mtDNA mutations**

The multicopy nature of the mitochondrial genome imposes a significant challenge for the detection and

**Fig. 1.** AID expression and mutational activity in Mec1, Ramos and Ramos-A23 cells. (A) Confocal microscopy images of DT40 AID<sup>-/-</sup>, DT40, mec1, Ramos and Ramos-A23 cells stained with anti-AID antibody (green) and MitoTracker Red (red). Nucleus was counterstained with DAPI (blue). Scale bars: 5 μm. (B) Representative flow cytometry analysis of GFP reversion mutations. Mec1, Ramos and Ramos-A23 cells were transduced with a GFPstop reporter construct and analyzed for GFP expression 21 days after transduction and puromycin selection. FSC, forward scatter. (C) AID mutational activity in Mec1, Ramos and Ramos-A23 lines measured by reversion of GFPstop reporter. For each sample, at least 100 000 cells were analyzed, and mutation frequency was calculated by dividing the GFP<sup>+</sup> count by the total number of counts. Data presented are mean and standard deviation of three biological replicates. (D) Mutation spectrum detected in the IgH variable region of Ramos cells. The total numbers of mutations and sequenced nucleotides are shown at the bottom.





**Fig. 2.** Development of high-fidelity 3DPCR for detection of mtDNA mutations. (A) Error spectrum of Taq-based two-step 3DPCR using a plasmid template harboring an 897-bp *Cytb* fragment. Errors generated during the first-round standard PCR (left panel) and second-round 3DPCR (right panel) were compared. In each panel, nucleotides in the left column are mutated to the nucleotides on the top row. The total numbers of mutations and sequenced nucleotides are shown below. (B) Summary of error rates for Taq-based two-step 3DPCR, Taq-based single 3DPCR and high-fidelity Phusion-based single 3DPCR. The level of errors was comparable between single 3DPCR and standard PCR, with fidelity of polymerase being the major determinant. (C) Error spectrum of Taq-based single 3DPCR using the same plasmid template as above. (D) Enrichment of *Cytb* mutations by Phusion-based single 3DPCR. *Cytb* gene was amplified from mtDNA of Ramos-A23 cells by either standard PCR or single 3DPCR, both with Phusion polymerase, and sequenced for mutations. A total of 144 kb sequencing data was obtained. Significance was assessed by Fisher's exact test. The graph summarizes the data from three biological replicates.

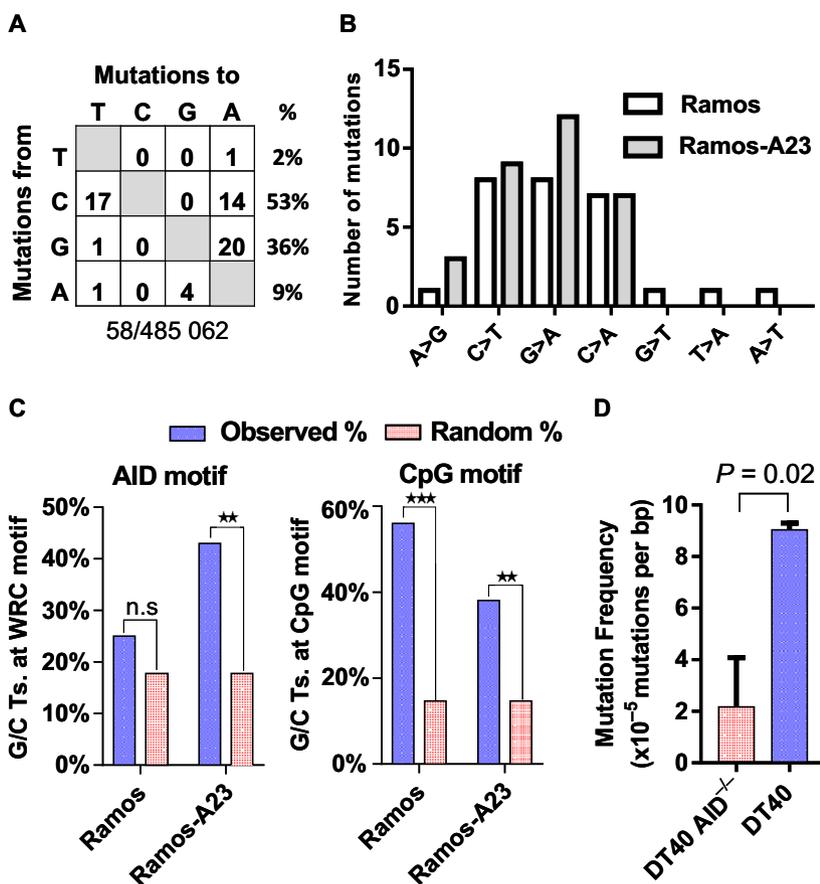
analysis of rare mtDNA mutations. Although 3DPCR can enrich mutations via biased amplification, Suspene *et al.* [28,29] reported the cautious existence of high 3DPCR replication errors (4–20 errors  $\text{kb}^{-1}$ ). To enable the detection of rare mtDNA mutations, we first dissected the origin of 3DPCR errors in our context with a cloned 897-bp fragment of mitochondrial *Cytb* gene using Suspene's procedure, which consisted of first-round standard PCR, followed by second-round 3DPCR (two-step 3DPCR) (Table S2). Sequencing of first-round standard PCR revealed a total error rate of  $0.72 \text{ kb}^{-1}$  with ~45% of sequences (21 of 47) harboring 1–3 errors featuring A>G/T>C predominance, which is identical to the reported error rate and spectrum of Taq polymerase (Fig. 2A,C) [30]. The second-round 3DPCR used the first-round product as templates, which led to a doubling of error rate ( $1.57 \text{ kb}^{-1}$ ) with 72% of sequences (33 of 46) harboring 1–4 errors (chi-square test,  $P = 0.01$ ; Fig. 2A,C). Notably, although the signature of A>G/T>C predominance was still evident, the G>A/C>T errors showed 4-fold enrichment

(0.1 versus  $0.43 \text{ kb}^{-1}$ ) during the second-round 3DPCR, which is consistent with the continued accumulation of Taq polymerase errors and biased amplification of G>A/C>T errors generated during the first-round PCR (Fig. 2A,C). We thus postulate that omitting first-round PCR and using high-fidelity polymerase may suppress 3DPCR errors. Indeed, Taq amplification of the cloned *Cytb* segment with single 3DPCR revealed an error rate and spectrum that was similar to standard PCR (Fig. 2B), whereas single 3DPCR with Phusion polymerase, which has more than 50-fold higher fidelity than Taq, revealed zero errors in 38 571 sequenced nucleotides (Fig. 2C). We next tested amplification of *Cytb* with Phusion-based single 3DPCR using mtDNA from Ramos-A23 cells as a template. Further sequencing revealed 13 point mutations in 74 599 sequenced nucleotides, corresponding to a 12-fold enrichment compared with the Phusion-based standard PCR (Fig. 2D), suggesting that a Phusion-based single 3DPCR approach can efficiently enrich mutations in mtDNA.

### Few AID targeting of mtDNA in AID-high cells

We thus compared Cytb mutations in Ramos (AID-low) and Ramos-A23 (AID-high) cells with a Phusion-based single 3DPCR approach. Extensive sequencing of Ramos (228 819 sequenced nucleotides) and Ramos-A23 cells (256 243 sequenced nucleotides) revealed a total of 58 point mutations (27 in Ramos and 31 in Ramos-A23) (Fig. 3A and Table S3). Spectrum analysis revealed an abundance of G/C targeting mutations (G>A/C>T transition and C>A/G>T transversion) and paucity of A>G/T>C transitions, which distinguished them from the observed 3DPCR errors (Fig. 2A) and the reported Phusion

amplification errors [30]. Furthermore, after correcting for base composition (guanine is 2.8-fold less common than C in the Cytb fragment), the G>A transitions exhibited more than 3-fold enrichment over C>T mutations ( $P < 0.001$ , binomial test). A similar trend of strand bias was evident for C>A/G>T transversions (14 C>A, 1 G>T). This mutational strand bias is also absent in 3DPCR errors but consistent with the heavy-strand biased mutation signature in the mitochondrial genome [31,32], confirming that these mutations are not PCR artifact but genuine mtDNA mutations.

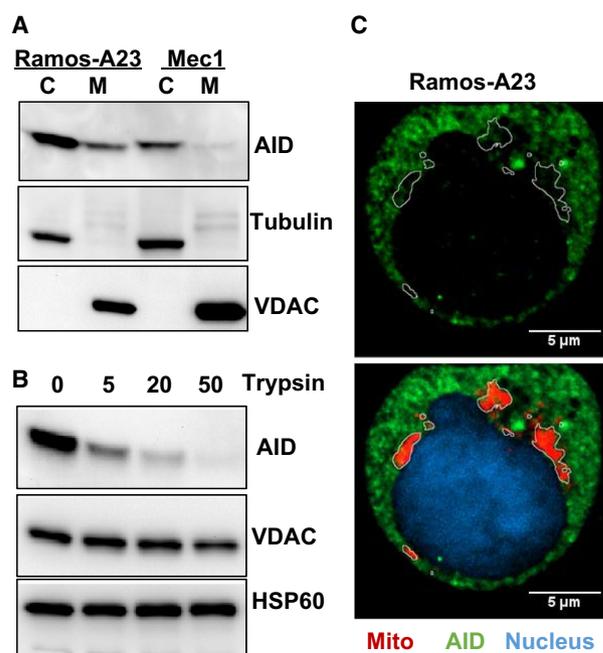


**Fig. 3.** Weak targeting of mtDNA by AID. (A) Combined spectrum of Cytb mutations in Ramos and Ramos-A23 cells recovered by Phusion-based single 3DPCR. The total numbers of mutations and sequenced nucleotides are shown below. Percentage of mutations occurring at a specific base is calculated in the last column. (B) Comparable pattern of Cytb mutations recovered by Phusion-based single 3DPCR in Ramos and Ramos-A23 cells. The mutations are focused on G/C nucleotides, mostly C>T/G>A transitions and C>A/G>T transversions. (C) Distribution of Cytb mutations by AID hotspot motif (WRC/GYW) (left panel) and CpG dinucleotide motif (right panel). The mutation context of G/C transitions in Ramos ( $n = 3$ ) and Ramos-A23 cells ( $n = 3$ ) was compared. Random % indicates the theoretical percentage of G/C transitions occurring at the AID motif or CpG motif upon random targeting, which is estimated based on the total number of G/C and the number of G/C within the AID or CpG motif. Significance was assessed by binomial test.  $**P < 0.01$ ,  $***P < 0.001$ ; n.s., not significant. (D) Analysis of Cytb mutations recovered by Phusion-based single 3DPCR using mtDNA from AID-deficient or -proficient chicken DT40 lymphoma cells. The graph depicts the mean and standard deviation of mutation frequency from three biological replicates. Significance was calculated by Welch's two-sample  $t$ -test.

Although Ramos-A23 cells exhibited nearly 3-fold higher AID mutational activity than Ramos cells (Fig. 1C), analysis of 3DPCR-recovered Cytb mutations revealed comparable frequency ( $11.9 \pm 3.5 \times 10^{-5}/\text{bp}$  and  $13.4 \pm 3.4 \times 10^{-5}/\text{bp}$ , respectively) and similar spectrums in both Ramos and Ramos-A23 cells (Fig. 3B). AID has been well recognized for preferentially targeting C at WRC/GYW hotspot motifs (where W = A/T, R = A/G and Y = C/T) [33]. A closer analysis of mutation context revealed that only 25% of G/C transitions (4 of 16) in Ramos coincided with AID's hotspot, which is expected by random targeting (binomial test,  $P = 0.5$ ). However, 43% of G/C transitions (9 of 21) in AID-high Ramos-A23 cells was within the AID hotspot, which corresponded to a 2.4-fold enrichment, supporting AID targeting of mtDNA in AID-high Ramos-A23 cells (binomial test,  $P = 0.007$ ) (Fig. 3C). We also compared 3DPCR recovered Cytb mutations in AID-high chicken lymphoma line DT40 (Fig. 1A) [22]. Sequencing of DT40 cells that were proficient or deficient of AID (DT40 versus DT40 AID<sup>-/-</sup>) revealed a 4.2-fold higher mutation rate in AID-proficient cells (Fig. 3D), further supporting AID targeting of mtDNA. However, the comparable mtDNA mutation rate between AID-low Ramos and AID-high Ramos-A23 cells suggested that AID targeting of mtDNA is a rare event.

### Physical separation of AID from mtDNA

To understand the mechanism underlying rare AID targeting of mtDNA, we tested whether AID can access mtDNA in the mitochondrial matrix. Western blot analysis of cytosol and mitochondrial fractions from both Ramos-A23 and Mec1 revealed that AID is not exclusively confined within the cytosol proteins and a subset of AID is observed in the mitochondrial fraction (Fig. 4A). To further analyze AID's submitochondrial localization, we treated isolated mitochondria with trypsin, which can remove the exposed proteins on the mitochondrial outer membrane without compromising outer membrane integrity [34]. As shown in Fig. 4B, trypsin digestion resulted in dose-dependent loss of mitochondria-associated AID, whereas integrated outer membrane protein VDAC and mitochondrial matrix protein HSP60 were not affected, suggesting that mitochondria-associated AID is only externally attached to the mitochondrial outer membrane, but not located within the mitochondrial inner membrane space or matrix. In support of this, western blot analysis did not identify truncated AID variants, which would be expected for proteins with dual mitochondrial and cytosol localization, because



**Fig. 4.** Physical separation of AID from mtDNA. (A) Western blot analysis of cytosol and mitochondrial AID proteins from Ramos-A23 and Mec1 cells, showing that a fraction of AID is associated with mitochondria. The purity of proteins was assessed by cytosol-specific Tubulin and mitochondrial outer membrane protein VDAC. (B) Mitochondria-associated AID is attached to the mitochondrial outer membrane. Purified Ramos-A23 mitochondria were treated with different concentrations of trypsin (0–50  $\mu\text{g}\cdot\text{mL}^{-1}$ ) for 30 min on ice before immunoblot analysis. The integrated mitochondrial outer membrane protein VDAC and mitochondrial matrix protein HSP60 were used as control. All immunoblots were representative of three independent experiments. (C) Confocal microscopy analysis of AID–mitochondria colocalization. Ramos-A23 cells were stained with AID antibody (green) and mitochondrial dye MitoTracker Red (red). Nucleus was counterstained with DAPI (blue). Mitochondrial contours were outlined at the bottom panel and overlaid onto the top panel. The marked mitochondrial areas were essentially free of AID staining. Scale bars: 5  $\mu\text{m}$ . C, cytosol protein; M, mitochondrial protein.

the N-terminal mitochondrial targeting peptide will be cleaved upon mitochondrial entrance. Consistently, immunofluorescence analysis revealed that the mitochondrial contour can clearly label areas that are essentially free of AID staining (Fig. 4C), supporting that AID cannot get access to mtDNA in the mitochondrial matrix.

### Recapitulation of endogenous mtDNA mutation signature by 3DPCR

Further analysis of Cytb mutations revealed that 56% of G/C transitions (9 of 16) in Ramos occurred within

the CpG dinucleotide motif (3.8-fold enrichment,  $P = 0.0001$ ), whereas 38% of G/C transitions (8 of 21) in Ramos-A23 were within the CpG motif (2.6-fold enrichment,  $P = 0.007$ ) (Fig. 3C), supporting the CpG motif as a mutational hotspot in mtDNA. This CpG mutation enrichment has also been observed in somatic mtDNA mutations across multiple human cancers [31,35]. In addition, C>A/G>T transversions identified in both Ramos and Ramos-A23 cells are hallmark patterns of oxidative DNA damage, consistent with mitochondria being the primary source of cellular reactive oxygen species [36]. Thus, Phusion-based single 3DPCR also enabled us to capture at least three endogenous mtDNA mutation signatures: the mutational strand bias, the CpG mutational hotspot and the reactive oxygen species-associated mutations.

## Discussion

The AID/APOBEC3 family proteins are single-stranded DNA deaminases with key immunological functions. AID initiates somatic hypermutation and class switch recombination of Ig in the adaptive arm of immunity, whereas APOBEC3 members restrict virus infections in the innate arm of immunity [37]. In addition, off-targeting of both AID and APOBEC3 members has been implicated as an important cause of genomic instability and cancer [4,5,38]. Because mitochondria play multiple critical roles and have their own mtDNA, a question of interest is whether mtDNA is targeted by these dangerous DNA mutators. Several APOBEC3 members, especially A3A and A3G, have been shown to be able to introduce a large number of point mutations into cytoplasmic mtDNA [39]. Our study here showed that AID can introduce very low levels of mtDNA mutations in AID-high Ramos-A23 and DT40 cells but cannot enter the mitochondrial matrix. As extensive literature has suggested the presence of mtDNA in the cytoplasm of cells [40–42], it is likely that cytoplasmic mtDNA is similarly targeted by AID, but at a much lower level. Thus, the mitochondrial genome is in general protected from AID/APOBEC3-mediated mutagenesis by the physical segregation of AID/APOBEC3 mutators from accessing the mitochondrial matrix.

The 3DPCR technique was initially developed to detect viral genomes that are hypermutated by APOBEC3 proteins, because APOBEC3-edited DNA is richer in A+T and thus can be amplified at relatively lower denaturation temperature [28]. 3DPCR has also been extended to detect lightly mutated C>T/G>A

transitions in cancer genomes [38,43]. Recently, Suspene *et al.* [29] cast strong doubt on the application of 3DPCR beyond viral DNA editing by showing that 3DPCR can generate high levels of C>T/G>A errors at the rate of 4–20 kb<sup>-1</sup>. This high error rate was mainly estimated using low-fidelity Taq polymerase and a two-step 3DPCR procedure (first-round standard PCR followed by second-round 3DPCR). We here revisited the origin of errors in Taq-based two-step 3DPCR using a cloned 897-bp Cytb fragment as a template. Our analysis showed that first-round standard PCR generated an error rate of 0.72 kb<sup>-1</sup> with mostly A>G/T>C and few C>T/G>A errors, which is identical to the error rate and spectrum of Taq polymerase [30]. Analysis of second-round 3DPCR confirmed a higher error rate (1.57 kb<sup>-1</sup>) with 52% A>G/T>C and 27% G>A/C>T errors, which is consistent with the continued accumulation of Taq errors together with the biased amplification of G>A/C>T errors generated during the first-round PCR. Thus, as shown here, Phusion-based single 3DPCR amplification of Cytb is essentially error free and can successfully recover genuine rare mutations in mtDNA. The enrichment of A>G/T>C errors in our 3DPCR sequencing is not observed in Suspene's analysis but may be explained by the difference in amplicon size. Compared with the 897-bp Cytb template here, Suspene *et al.* [29] used much shorter templates (<300 bp), which may significantly magnify the amplification bias toward templates with G>A/C>T errors. Taken together, we propose that the level of 3DPCR errors is context dependent, with polymerase fidelity, amplification cycle and template size being three major determining factors. A limitation of Phusion-based single 3DPCR is that Phusion cannot polymerize U-containing template, but this can be overcome by a U-compatible high-fidelity polymerase, such as Phusion U or PfuTurbo Cx. In addition, Phusion can efficiently detect AID-generated mutations in genomic DNA because AID-generated U is rapidly processed by downstream low-fidelity DNA repair pathways [44].

## Acknowledgements

We thank Dr. David G. Schatz (Yale University School of Medicine) for providing GFPstop construct and Dr. Gur Yaari (Yale University School of Medicine) for assistance in data analysis. Part of the work was performed at Yale University School of Medicine. This work was supported by the Fundamental Research Funds for the Central Universities of China (to XG), the Scientific Research Foundation for Returned Scholars of Shaanxi Province (to XG) and the Natural Science

Foundation of Shaanxi Province (Grant 2020JM-412 to XG).

## Conflict of interest

The authors declare no conflict of interest.

## Data accessibility

The data that support the findings of this study are available from the corresponding author (XG) upon reasonable request.

## Author contributions

XG and MPS conceived the idea. XG performed the experiments, analyzed the data and wrote the paper. HW and KZ performed the experiments and analyzed the data. YC and JL cowrote the paper. All authors reviewed the manuscript.

## References

- Muramatsu M, Kinoshita K, Fagarasan S, Yamada S, Shinkai Y and Honjo T (2000) Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. *Cell* **102**, 553–563.
- Maizels N (2005) Immunoglobulin gene diversification. *Annu Rev Genet* **39**, 23–46.
- Liu M, Duke JL, Richter DJ, Vinuesa CG, Goodnow CC, Kleinstein SH and Schatz DG (2008) Two levels of protection for the B cell genome during somatic hypermutation. *Nature* **451**, 841–845.
- Robbiani DF and Nussenzweig MC (2013) Chromosome translocation, B cell lymphoma, and activation-induced cytidine deaminase. *Annu Rev Pathol* **8**, 79–103.
- Casellas R, Basu U, Yewdell WT, Chaudhuri J, Robbiani DF and Di Noia JM (2016) Mutations, kataegis and translocations in B cells: understanding AID promiscuous activity. *Nat Rev Immunol* **16**, 164–176.
- Gu X, Shivarov V and Strout MP (2012) The role of activation-induced cytidine deaminase in lymphomagenesis. *Curr Opin Hematol* **19**, 292–298.
- Sawai Y, Kodama Y, Shimizu T, Ota Y, Maruno T, Eso Y, Kurita A, Shiokawa M, Tsuji Y, Uza N *et al.* (2015) Activation-induced cytidine deaminase contributes to pancreatic tumorigenesis by inducing tumor-related gene mutations. *Cancer Res* **75**, 3292–3301.
- Matsumoto Y, Marusawa H, Kinoshita K, Niwa Y, Sakai Y and Chiba T (2010) Up-regulation of activation-induced cytidine deaminase causes genetic aberrations at the CDKN2b-CDKN2a in gastric cancer. *Gastroenterology* **139**, 1984–1994.
- Takai A, Marusawa H, Minaki Y, Watanabe T, Nakase H, Kinoshita K, Tsujimoto G and Chiba T (2012) Targeting activation-induced cytidine deaminase prevents colon cancer development despite persistent colonic inflammation. *Oncogene* **31**, 1733–1742.
- Nonaka T, Toda Y, Hiai H, Uemura M, Nakamura M, Yamamoto N, Asato R, Hattori Y, Bessho K, Minato N *et al.* (2016) Involvement of activation-induced cytidine deaminase in skin cancer development. *J Clin Invest* **126**, 1367–1382.
- Bhutani N, Brady JJ, Damian M, Sacco A, Corbel SY and Blau HM (2010) Reprogramming towards pluripotency requires AID-dependent DNA demethylation. *Nature* **463**, 1042–1047.
- Cortellino S, Xu J, Sannai M, Moore R, Caretti E, Cigliano A, Le Coz M, Devarajan K, Wessels A, Soprano D *et al.* (2011) Thymine DNA glycosylase is essential for active DNA demethylation by linked Deamination-base excision repair. *Cell* **146**, 67–79.
- Kunimoto H, McKenney AS, Meydan C, Shank K, Nazir A, Rapaport F, Durham B, Garrett-Bakelman FE, Pronier E, Shih AH *et al.* (2017) AID is a key regulator of myeloid/erythroid differentiation and DNA methylation in hematopoietic stem/progenitor cells. *Blood* **129**, 1779–1790.
- Teater M, Dominguez PM, Redmond D, Chen Z, Ennishi D, Scott DW, Cimmino L, Ghione P, Chaudhuri J, Gascoyne RD *et al.* (2018) AICDA drives epigenetic heterogeneity and accelerates germinal center-derived lymphomagenesis. *Nat Commun* **9**, 222.
- Kumar R, DiMenna L, Schrode N, Liu TC, Franck P, Munoz-Descalzo S, Hadjantonakis A-K, Zarrin AA, Chaudhuri J, Elemento O *et al.* (2013) AID stabilizes stem-cell phenotype by removing epigenetic memory of pluripotency genes. *Nature* **500**, 89–92.
- Ito S, Nagaoka H, Shinkura R, Begum N, Muramatsu M, Nakata M and Honjo T (2004) Activation-induced cytidine deaminase shuttles between nucleus and cytoplasm like apolipoprotein B mRNA editing catalytic polypeptide 1. *Proc Natl Acad Sci USA* **101**, 1975–1980.
- Uchimura Y, Barton LF, Rada C and Neuberger MS (2011) REG- $\gamma$  associates with and modulates the abundance of nuclear activation-induced deaminase. *J Exp Med* **208**, 2385–2391.
- Hasler J, Rada C and Neuberger MS (2011) Cytoplasmic activation-induced cytidine deaminase (AID) exists in stoichiometric complex with translation elongation factor 1 $\alpha$  (eEF1A). *Proc Natl Acad Sci USA* **108**, 18366–18371.
- Orthwein A, Patenaude AM, Affar EB, Lamarre A, Young JC and Di Noia JM (2010) Regulation of

- activation-induced deaminase stability and antibody gene diversification by Hsp90. *J Exp Med* **207**, 2751–2765.
- 20 Wang Q, Kieffer-Kwon KR, Oliveira TY, Mayer CT, Yao K, Pai J, Cao Z, Dose M, Casellas R, Jankovic M *et al.* (2017) The cell cycle restricts activation-induced cytidine deaminase activity to early G1. *J Exp Med* **214**, 49–58.
  - 21 Sale JE and Neuberger MS (1998) TdT-accessible breaks are scattered over the immunoglobulin V domain in a constitutively hypermutating B cell line. *Immunity* **9**, 859–869.
  - 22 Arakawa H, Hauschild J and Buerstedde JM (2002) Requirement of the activation-induced deaminase (AID) gene for immunoglobulin gene conversion. *Science* **295**, 1301–1306.
  - 23 Wang CL, Harper RA and Wabl M (2004) Genome-wide somatic hypermutation. *Proc Natl Acad Sci USA* **101**, 7352–7356.
  - 24 Campbell MJ, Zelenetz AD, Levy S and Levy R (1992) Use of family specific leader region primers for PCR amplification of the human heavy chain variable region gene repertoire. *Mol Immunol* **29**, 193–203.
  - 25 Gu X, Booth CJ, Liu Z and Strout MP (2016) AID-associated DNA repair pathways regulate malignant transformation in a murine model of BCL6-driven diffuse large B-cell lymphoma. *Blood* **127**, 102–112.
  - 26 Stacchini A, Aragno M, Vallario A, Alfàrano A, Circosta P, Gottardi D, Faldella A, Rege-Cambrin G, Thunberg U, Nilsson K *et al.* (1999) MEC1 and MEC2: two new cell lines derived from B-chronic lymphocytic leukaemia in prolymphocytoid transformation. *Leukemia Res* **23**, 127–136.
  - 27 Behar DM, van Oven M, Rosset S, Metspalu M, Loogvali EL, Silva NM, Kivisild T, Torroni A and Villems R (2012) A "Copernican" reassessment of the human mitochondrial DNA tree from its root. *Am J Hum Genet* **90**, 675–684.
  - 28 Suspene R, Henry M, Guillot S, Wain-Hobson S and Vartanian J-P (2005) Recovery of APOBEC3-edited human immunodeficiency virus G hypermutants by differential DNA denaturation PCR. *J Gen Virol* **86**, 125–129.
  - 29 Suspene R, Caval V, Henry M, Bouzidi MS, Wain-Hobson S and Vartanian JP (2014) Erroneous identification of APOBEC3-edited chromosomal DNA in cancer genomics. *Br J Cancer* **110**, 2615–2622.
  - 30 McInerney P, Adams P and Hadi MZ (2014) Error rate comparison during polymerase chain reaction by DNA polymerase. *Mol Biol Int* **2014**, 287430.
  - 31 Ju YS, Alexandrov LB, Gerstung M, Martincorena I, Nik-Zainal S, Ramakrishna M, Davies HR, Papaemmanuil E, Gundem G, Shlien A *et al.* (2014) Origins and functional consequences of somatic mitochondrial DNA mutations in human cancer. *eLife* **3**, e02935.
  - 32 Gu X, Kang X and Liu J (2019) Mutation signatures in germline mitochondrial genome provide insights into human mitochondrial evolution and disease. *Hum Genet* **138**, 613–624.
  - 33 Rogozin IB and Diaz M (2004) Cutting edge: DGYW/WRCH is a better predictor of mutability at G: C bases in Ig hypermutation than the widely accepted RGYW/WRCY motif and probably reflects a two-step activation-induced cytidine deaminase-triggered process. *J Immunol* **172**, 3382–3384.
  - 34 Bellot G, Cartron PF, Er E, Oliver L, Juin P, Armstrong LC, Bornstein P, Mihara K, Manon S and Vallette FM (2006) TOM22, a core component of the mitochondria outer membrane protein translocation pore, is a mitochondrial receptor for the proapoptotic protein Bax. *Cell Death Differ* **14**, 785–794.
  - 35 Stewart JB, Alaei-Mahabadi B, Sabarinathan R, Samuelsson T, Gorodkin J, Gustafsson CM and Larsson E (2015) Simultaneous DNA and RNA mapping of somatic mitochondrial mutations across diverse human cancers. *PLoS Genet* **11**, e1005333.
  - 36 Kauppila JH and Stewart JB (2015) Mitochondrial DNA: radically free of free-radical driven mutations. *Biochim Biophys Acta* **1847**, 1354–1361.
  - 37 Conticello S (2008) The AID/APOBEC family of nucleic acid mutators. *Genome Biol* **9**, 229.
  - 38 Burns MB, Lackey L, Carpenter MA, Rathore A, Land AM, Leonard B, Refsland EW, Kotandeniya D, Tretyakova N, Nikas JB *et al.* (2013) APOBEC3B is an enzymatic source of mutation in breast cancer. *Nature* **494**, 366–370.
  - 39 Suspene R, Aynaud MM, Guetard D, Henry M, Eckhoff G, Marchio A, Pineau P, Dejean A, Vartanian J-P and Wain-Hobson S. (2011) Somatic hypermutation of human mitochondrial and nuclear DNA by APOBEC3 cytidine deaminases, a pathway for DNA catabolism. *Proc Natl Acad Sci USA* **108**, 4858–4863.
  - 40 McArthur K, Whitehead LW, Heddleston JM, Li L, Padman BS, Oorschot V, Geoghegan ND, Chappaz S, Davidson S, San Chin H *et al.* (2018) BAK/BAX macropores facilitate mitochondrial herniation and mtDNA efflux during apoptosis. *Science* **359**, eaao6047.
  - 41 West AP, Houry-Hanold W, Staron M, Tal MC, Pineda CM, Lang SM, Bestwick M, Duguay BA, Raimundo N, MacDuff DA *et al.* (2015) Mitochondrial DNA stress primes the antiviral innate immune response. *Nature* **520**, 553–557.
  - 42 West AP and Shadel GS (2017) Mitochondrial DNA in innate immune responses and inflammatory pathology. *Nat Rev Immunol* **17**, 363–375.
  - 43 Shinohara M, Io K, Shindo K, Matsui M, Sakamoto T, Tada K, Kobayashi M, Kadowaki N and Takaori-Kondo A (2012) APOBEC3B can impair genomic stability by inducing base substitutions in genomic DNA in human cells. *Sci Rep* **2**, 806.

- 44 Rada C, Di Noia JM and Neuberger MS (2004) Mismatch recognition and uracil excision provide complementary paths to both Ig switching and the A/T-focused phase of somatic mutation. *Mol Cell* **16**, 163–171.

### Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**Table S1.** List of clonal single-nucleotide polymorphisms in the sequenced mtDNA regions of Ramos and Mec1 cells.

**Table S2.** List of 3DPCR errors generated by two-step 3DPCR and single 3DPCR using cloned 897-bp Cytb template and Taq polymerase.

**Table S3.** List of 58 point mutations in Ramos and Ramos-A23 subclones.