

# SCIENTIFIC REPORTS



OPEN

## Transient AID expression for *in situ* mutagenesis with improved cellular fitness

Talal Salem Al-Qaisi<sup>1,2,3</sup>, Yu-Cheng Su<sup>4</sup> & Steve R. Roffler<sup>2</sup>

Activation induced cytidine deaminase (AID) in germinal center B cells introduces somatic DNA mutations in transcribed immunoglobulin genes to increase antibody diversity. Ectopic expression of AID coupled with selection has been successfully employed to develop proteins with desirable properties. However, this process is laborious and time consuming because many rounds of selection are typically required to isolate the target proteins. AID expression can also adversely affect cell viability due to off target mutagenesis. Here we compared stable and transient expression of AID mutants with different catalytic activities to determine conditions for maximum accumulation of mutations with minimal toxicity. We find that transient (3–5 days) expression of an AID upmutant in the presence of selection pressure could induce a high rate of mutagenesis in reporter genes without affecting cells growth and expansion. Our findings may help improve protein evolution by ectopic expression of AID and other enzymes that can induce DNA mutations.

Protein engineering is a powerful technique to improve protein activity, stability and other properties for industrial, diagnostic and therapeutic applications<sup>1</sup>. Protein engineering by rational design is very effective, but requires knowledge of the structure and function of the protein of interest<sup>1,2</sup>. By contrast, directed protein evolution performed by alternating rounds of random mutagenesis and selection can be employed in many cases where rational design is not feasible.

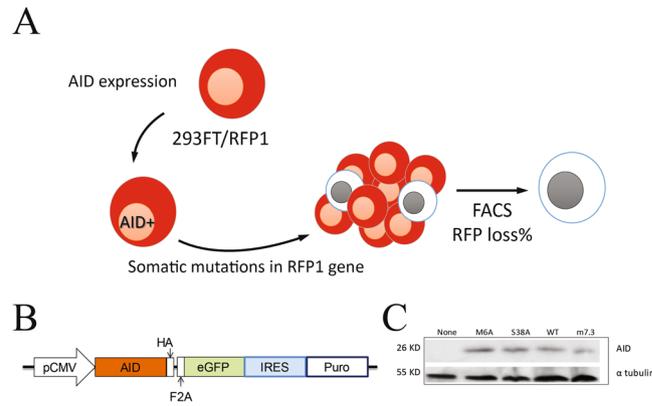
An interesting approach for autonomous mutagenesis in mammalian cells has been inspired by the germinal center reaction in which B cells overexpress the enzyme activation induced cytidine deaminase (AID), which is essential for diversification of antibody genes *in vivo*<sup>3,4</sup>. AID mutagenesis has been applied successfully for evolution of antibody and non-antibody proteins in both B and non-B cells<sup>4–10</sup>. AID mutagenesis coupled with high throughput screening in mammalian cells appears to be an ideal system for protein evolution because it avoids repetitive transfection and recloning and provides selection for proteins with better expression, activity and stability<sup>9</sup>. However, AID mutagenesis for protein evolution suffers from two major drawbacks; low mutation rate and genotoxicity manifested by widespread genomic instability and reduced fitness when harmful mutations accumulate in highly transcribed vital genes<sup>11–15</sup>. Wang *et al.* investigated the possibility of improving AID mutagenesis by increasing the specific activity of AID. Indeed, AID upmutants with increased activity were isolated, but they also led to increased genomic instability<sup>16</sup> and impaired cells expansion<sup>14</sup>, which limits their use for protein evolution applications.

To better understand how to effectively use AID upmutants for mutagenesis in mammalian cells, we investigated *in situ* mutagenesis of a red fluorescence reporter protein by AID mutants with different enzymatic activities. We found that an AID upmutant (m7.3) generates a spike of mutagenesis shortly after expression (less than 10 days). Longer expression did not produce additional mutations in the reporter gene but reduced cell fitness and expansion. Thus, our results suggest that ideally, AID upmutants should be expressed transiently to maximize mutagenesis in target transgenes and minimize off-target toxicity.

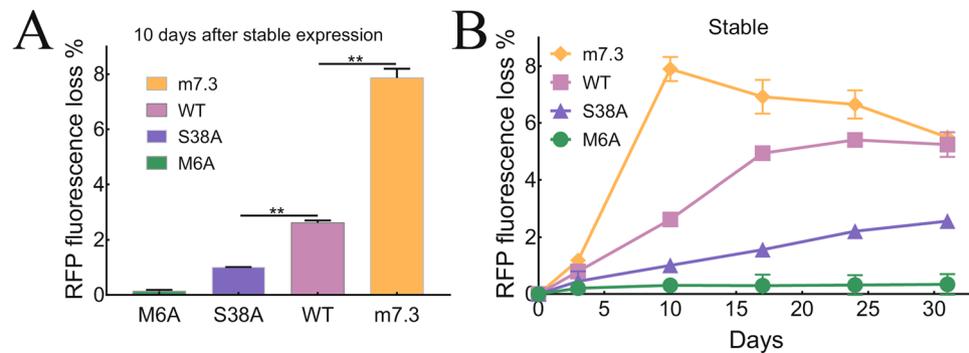
### Results

**Somatic hypermutation reporter system.** Because AID can mutate highly transcribed genes<sup>17–20</sup>, fluorescence reporter proteins are often used to monitor mutation rates<sup>4,21</sup>. We therefore generated a 293FT cell line that stably expresses monomeric RFP (RFP1) (Fig. 1A). Expression of AID in these cells can induce mutations

<sup>1</sup>Taiwan International Graduate Program in Molecular Medicine, National Yang-Ming University and Academia Sinica, Taipei, Taiwan. <sup>2</sup>Institute of Biomedical Sciences, Academia Sinica, Taipei, Taiwan. <sup>3</sup>Institute of Biochemistry and Molecular Biology, National Yang-Ming University, Taipei, Taiwan. <sup>4</sup>Department of Biological Science and Technology, National Chiao Tung University, Hsin-Chu, Taiwan. Correspondence and requests for materials should be addressed to S.R.R. (email: sroff@ibms.sinica.edu.tw)



**Figure 1.** Constructs and screening system. (A) 293FT/RFP1 cells stably express RFP1 fluorescent protein. RFP1 fluorescence loss after AID expression is used to assess AID mutagenic activity. (B) Schematic representation of the AID expression vector. A CMV promoter is followed by a human AID or AID mutant gene which is linked to an HA tag at the C-terminus followed by furin/2A peptide (F2A) bicistronic expression linker and an eGFP reporter gene. An internal ribosome entry site (IRES) is used for bicistronic expression of a puromycin resistance gene. (C) Cell lysates from 293FT/RFP1 cells expressing AID mutants were used to perform immunoblot analysis with antibodies binding to the HA tag on AID or tubulin as a cell loading control.



**Figure 2.** Stable AID expression induces a spike of RFP fluorescence loss. 293FT/RFP1 cells were stably transduced with AID mutants and then  $10^5$  cells were analyzed for somatic hypermutation by RFP fluorescence loss using a flow cytometer (A) 10 days after stable expression or (B) periodically over one month. The results show the mean values of three replicates  $\pm$  S.D. Significant differences between RFP loss after expression of S38A or m7.3 mutants compared to AID-WT are indicated; \*\* $P \leq 0.01$ .

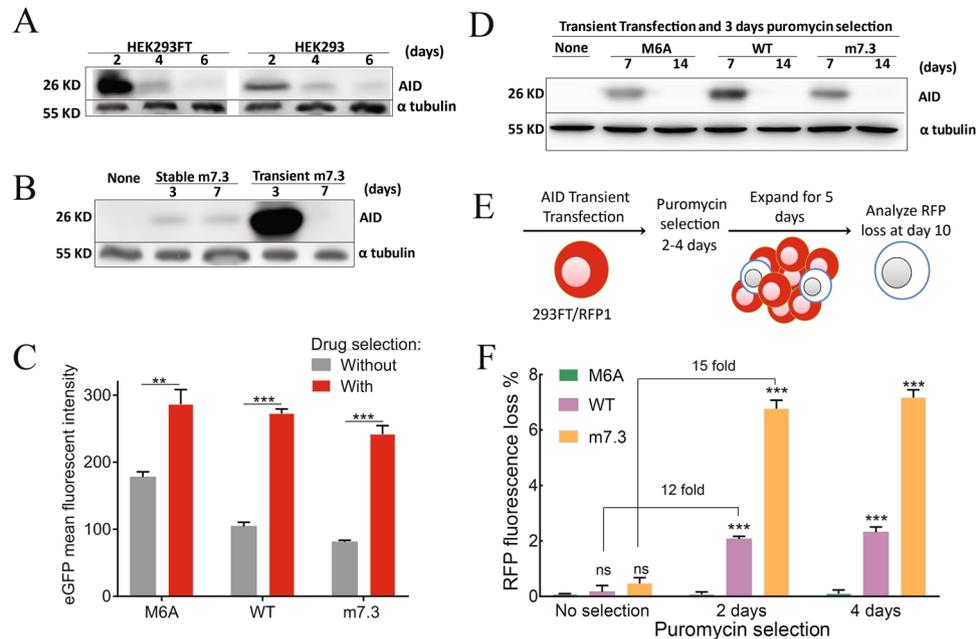
in RFP gene, some of which lead to loss of RFP protein fluorescence. Analyzing the percentage of cells that lose fluorescence provides an estimate of the relative mutation rate induced by different AID mutants.

**Mutagenesis of RFP1 transgene by stably expressed AID.** To investigate mutagenesis by AID, we examined either wild type AID (AID-WT) or three AID mutants; AID m7.3 (m7.3) which displays high catalytic activity<sup>22</sup>, AID S38A (S38A) which displays about 30% of AID-WT activity<sup>23</sup>, as well as M6A which lacks AID activity<sup>24</sup>. AID mutants were cloned into a lentivirus expression plasmid (Fig. 1B) and recombinant lentiviral particles were used to stably infect 293FT/RFP1 cells. Expression of AID in 293FT/RFP1 cells was confirmed 3 days later by immunoblot analysis which detects the HA tag present on the recombinant AID proteins (Fig. 1C).

As expected, the high activity AID m7.3 mutant induced significantly more RFP negative cells on day 10 as compared to AID-WT, consistent with introduction of more mutations in the reporter gene (Fig. 2A). The low activity S38A mutant induced fewer RFP negative cells while the inactive M6A mutant produced almost no loss of RFP fluorescence (Fig. 2A), indicating that the percentage of RFP fluorescence loss can be used as a readout of relative mutation rates.

To study accumulation of mutations over time, we measured the percentage of RFP negative cells for one month (Fig. 2B). The m7.3 AID upmutant achieved maximum RFP fluorescence loss on day 10 but the frequency of RFP loss progressively decreased after this time, consistent with slower growth of cells that accumulated excessive harmful mutations. By contrast, cells that stably expressed AID-WT lost RFP fluorescence at a slower rate, which reached a plateau 24 days after infection. As expected, S38A cells accumulated RFP loss at a slower rate than AID-WT and maximum accumulation of RFP loss was not reached in 32 days.

**Transient expression of AID mutants.** Since the maximum mutation frequency was achieved within 10 days after stable expression of m7.3, we asked whether transient transfection with this AID upmutant is sufficient



**Figure 3.** Optimization of transient transfection for efficient somatic hypermutation. **(A)** Comparison between expression of AID m7.3 in HEK293FT and HEK293 cells over time. Cells transiently transfected with m7.3 plasmid were collected every 2 days for 6 days. Cell lysates were used to perform immunoblot analysis with antibodies against the HA tag on AID or tubulin as a cell loading control. **(B)** Comparison between AID expression level after stable and transient transfection. Cells were transfected stably or transiently with m7.3 AID. Cell lysates prepared after 3 and 7 days were used to perform immunoblot analysis with antibodies binding to the HA tag on AID or tubulin as a cell loading control. **(C)** Mean fluorescence intensity of eGFP in  $10^5$  cells measured at day 7 including 3 days with or without puromycin selection,  $n = 3$ ; Bars, S.D. **(D)** Cells transiently transfected with AID plasmids and selected with puromycin ( $2 \mu\text{g}/\text{mL}$ ) for 3 days. Cell lysates prepared at 7 and 14 days after transient transfection were used to perform immunoblot analysis with antibodies binding to the HA tag on AID or tubulin as a cell loading control. **(E)** Reporter cells (293FT/RFP1) were transfected with different AID mutants and selected for 0, 2 or 4 days with  $2 \mu\text{g}/\text{mL}$  puromycin. RFP loss was measured 10 days after transient transfection. **(F)** RFP loss with or without puromycin selection. Results show mean frequencies of RFP loss measured in  $10^5$  cells at each time ( $n = 3$ ). Significant differences between mean values are indicated;  $**P \leq 0.01$ ,  $***P \leq 0.001$ , ns, not significant.

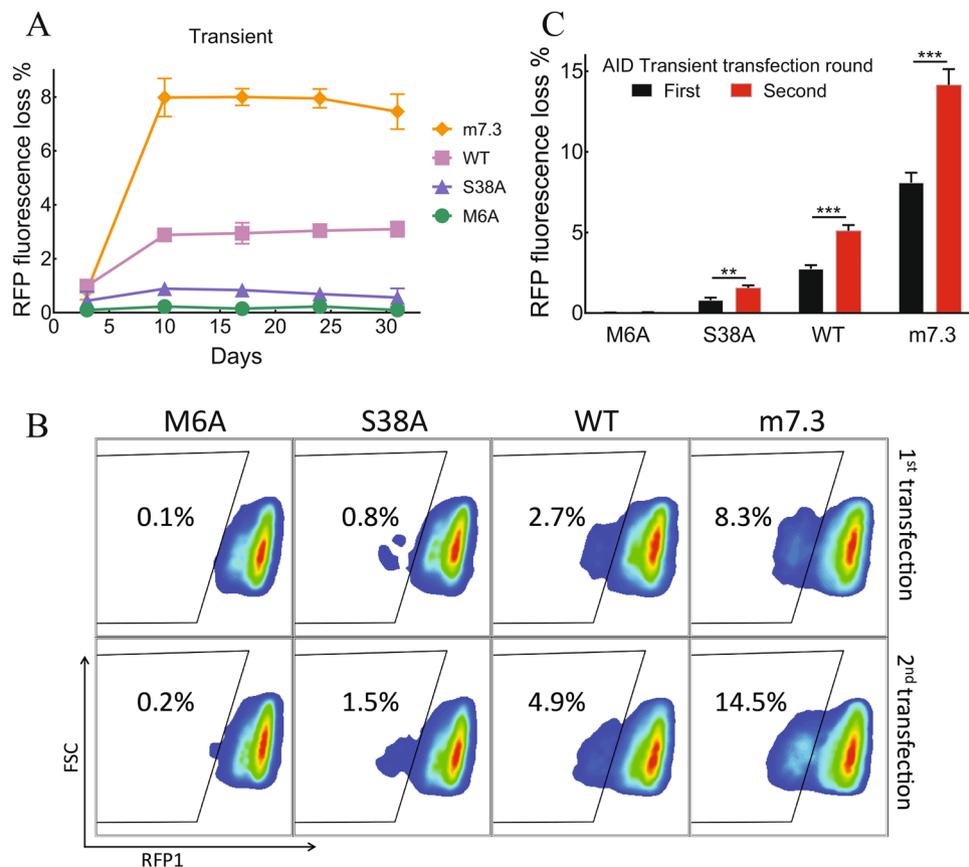
for inducing efficient somatic hypermutation. We first transiently expressed m7.3 and monitored its expression over time by Western blot analysis in HEK293FT and HEK293 cells (Fig. 3A). Although expression is very efficient after 2 days, it decreased dramatically over time probably due to plasmid dilution during cell replication. HEK293FT cells shows clear enhancement of expression after 2 days which may be attributed to replication of the plasmid (pAS3w.Ppuro), which contains the SV40 origin of replication (SV40 ori), in HEK293FT cells which express the SV40 large T antigen<sup>25</sup>.

We then compared expression level of m7.3 after transient and stable expression (Fig. 3B). Stable expression using lentiviral transduction produced continuous expression of AID over time, while transient transfection generated a large amount of the protein for a short time.

We then asked if we can benefit from the high efficiency of drug selection to extend AID expression by simply using short term drug selection to select cells that retain the plasmid longer. Indeed, we found that three days of drug selection was sufficient to enhance expression of the eGFP reporter by about three fold, which is expressed bicistronically with AID (Fig. 3C). Extension of AID expression was further confirmed by Western blot analysis of M6A, WT and m7.3 (Fig. 3D). AID protein was apparent at day 7 when drug selection is used, but decreased to undetectable levels by day 14 (Fig. 3D), indicating that expression is still transient, and stable transfectants were not selected by short-term exposure to puromycin.

Finally, we tested whether transient transfection with short term selection has an impact on inducing mutations in the reporter 293FT/RFP1 cells (Fig. 3E). When puromycin was not used, the mutation frequency was extremely low (Fig. 3F). By contrast, addition of puromycin for 2 or 4 days significantly enriched RFP fluorescence loss by 12 and 15 fold for AID-WT and m7.3, respectively (Fig. 3F). Enhancement of RFP loss by puromycin treatment could be due to preferential selection of cells that express AID for longer times.

**Time course and multiple AID transient transfections.** Loss of RFP fluorescence was monitored for 4 weeks after transient transfection of AID mutants and selection in puromycin for 3 days. In all cases, maximum mutagenesis was achieved within 10 days after transient transfection, reaching around 3% and 8% for AID-WT and m7.3 respectively (Fig. 4A). The percentage of RFP negative cells did not change over a month follow up, consistent with the transient expression of AID and similar growth rates of the RFP positive and RFP negative



**Figure 4.** Robust RFP1 loss by transient expression of AID. **(A)** 293FT/RFP1 cells transiently transfected with plasmids encoding AID mutants, selected with puromycin for 3 days and analyzed for somatic hypermutation by RFP fluorescence loss. Results show mean values of three experiments  $\pm$  SD. **(B)** Representative flow cytometry results and **(C)** analysis of accumulative RFP loss after one and two rounds of transient AID expression. Results show the mean values of three replicates  $\pm$  SD. Significant differences between RFP loss after the first and the second AID transfections are indicated; \*\* $P \leq 0.01$ , \*\*\* $P \leq 0.001$ .

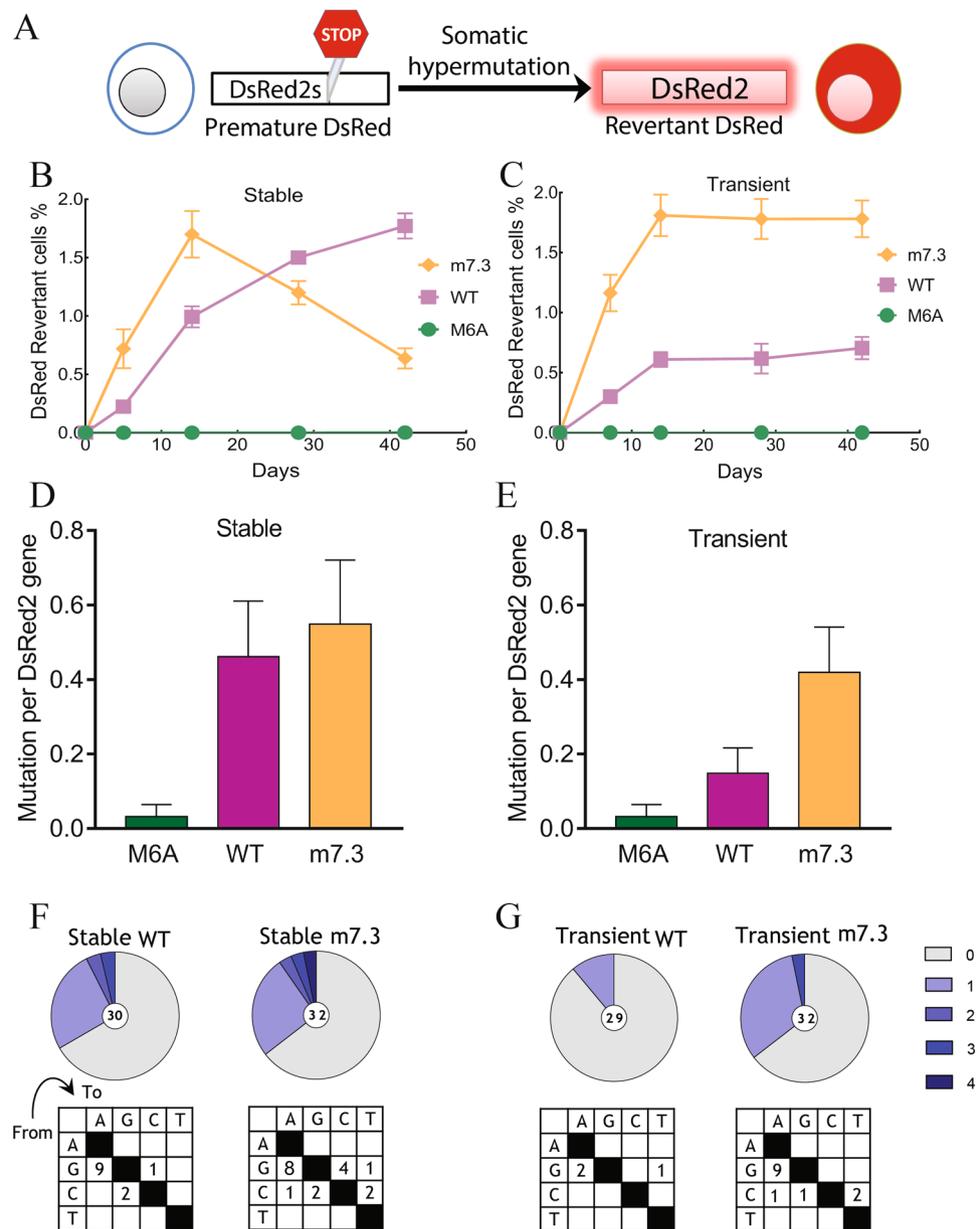
cells. Limiting the time cells are exposed to AID may be important to balance between diversity generation and accumulation of deleterious mutations<sup>26,27</sup>.

We further examined if a second round of transient transfection can continue to produce mutations in the reporter gene, because repeated rounds of mutagenesis and selection are typically required for successful protein evolution. Indeed, the frequency of RFP negative cells significantly increased after a second round of transient transfection of AID (Fig. 4B,C), reaching around 13% of total cells with m7.3 AID.

**Transient AID expression in a DsRed reversion assay.** To further test the usefulness of transient AID expression, we investigated reversion of a single fatal base pair in the red fluorescent protein (DsRed2) gene. We introduced a pre-mature stop codon within an AID hotspot motif, which renders the DsRed2 gene non-functional (Fig. 5A). Emergence and retention of DsRed2 fluorescence was studied over time after mutagenesis by AID-WT and m7.3 by either stable or transient transfection protocols. Stable expression of m7.3 (Fig. 5B) induced rapid emergence of red fluorescence, which reached a maximum frequency of 1.7% after two weeks and then gradually decreased over time. Stable expression of AID-WT required about 42 days to accumulate a similar number of DsRed2 positive cells. Transient transfection of m7.3 also produced about 1.8% DsRed2 positive cells after two weeks (Fig. 5C), but in contrast to stable expression of m7.3, there was no obvious outgrowth of DsRed2 negative cells at later times, indicating less damage to cells.

Furthermore, the reporter DsRed2 reporter system was validated by sequence analysis, which revealed that mutation load of the unsorted cells after 14 days of stable or transient expression of AID mutants correlates with the percentage of revertant cells in the DsRed revertant system (Fig. 5D,E). Mutations in all groups were predominantly transition mutations and mainly G to A mutations (Fig. 5F,G).

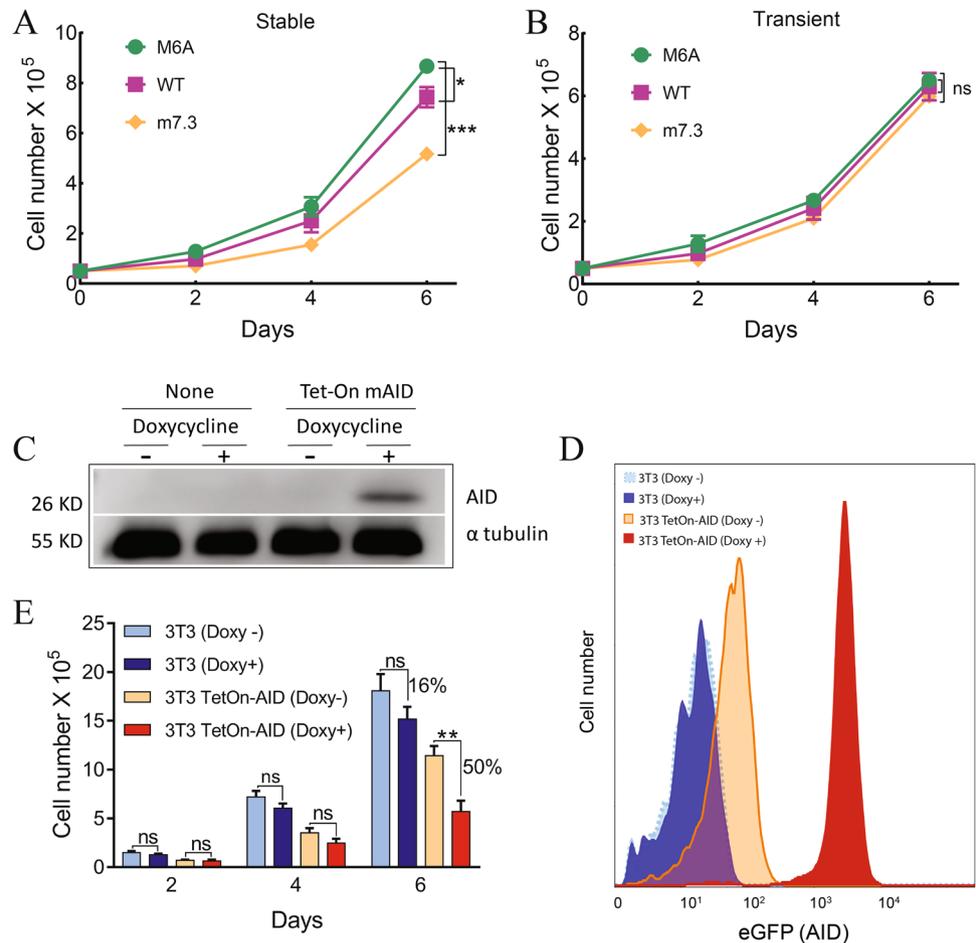
**Toxicity of Stable AID expression.** Both transient and stable expression of AID can effectively introduce mutations into reporter genes, but stable expression of AID appeared to adversely affect cell viability, in agreement with studies demonstrating toxicity induced by uncontrolled expression of AID<sup>11–15</sup>. To confirm the adverse effect of AID expression on cell expansion, we stably or transiently expressed M6A, AID-WT or m7.3 in 293FT cells, seeded similar number of cells and counted viable cells every 2 days for 6 days. Indeed, we found that stable



**Figure 5.** Robust DsRed reversion by transient AID expression. (A) 293FT/DsRed2s reporter cells stably express DsRed with a premature stop codon within an AID hotspot motif. AID expression can remove the premature stop codon and restore DsRed fluorescence. (B) 293FT/DsRed2s cells stably transduced with AID mutants ( $10^6$  cells) were analyzed for gain of red fluorescence over time.  $n = 3$ ; Bars, SD. (C) 293FT/DsRed2s cells transiently transfected with AID mutants and selected with puromycin for 3 days ( $10^6$  cells) were analyzed for gain of red fluorescence over time.  $n = 3$ ; Bars, SD. (D) Frequency of mutations in the DsRed2 gene obtained by PCR amplification from unsorted cells 14 days after expression of AID mutants stably or (E) transiently with puromycin selection. Bars, SEM. (F) Pie charts illustrating the number of mutations in PCR-amplified sequences from unsorted cells 14 days after expression of AID mutants stably or (G) transiently. The number in the middle of the pie chart indicates the total number of sequences analyzed. Below the pie charts, mutation matrix shows nucleotide substitution pattern identified from sequencing results.

(Fig. 6A) but not transient (Fig. 6B) expression of both AID-WT and m7.3 significantly decreased cell growth as compared to the control group expressing the inactive M6A mutant.

Toxicity of AID was further confirmed in 3T3/TetOn-AID, a mouse cell line stably expressing a tetracycline-inducible AID expression cassette<sup>10</sup>. Controllable induction of AID expression was confirmed by Western blot analysis (Fig. 6C), and eGFP reporter (Fig. 6D). Continuous induction of AID expression in these cells for 6 days shows 50% growth inhibition compared to cells cultured in doxycycline-free medium. Growth retardation can be partially attributed to doxycycline itself, but this is a minor effect because parental 3T3 cells are only inhibited by 16% after exposure to the same dose of doxycycline for the same time period.



**Figure 6.** Transient expression of AID reduces its adverse effect on cell growth. Growth curve of HEK293FT cells (A) virally transduced with AID mutants or (B) transiently transfected with AID expressing plasmids. Results show mean values  $\pm$  SD ( $n = 3$ ). (C) Western blots show AID expression in cell extracts from 3T3 and 3T3 TetOn-AID cells after 4 days of culturing the cells with or without 0.5  $\mu\text{g}/\text{mL}$  doxycycline; tubulin was used as loading control. (D) AID induction as confirmed by bicistronic GFP expression in  $10^5$  cells of each 3T3 and 3T3 TetOn-AID cells after 4 days of culturing the cells with or without 0.5  $\mu\text{g}/\text{mL}$  doxycycline. (E) Growth curve of 3T3 and 3T3 TetOn-AID cells with or without 0.5  $\mu\text{g}/\text{mL}$  for 6 days. Results show mean values  $\pm$  SD ( $n = 3$ ). Significant differences between growth curves are indicated; \* $p \leq 0.05$ ; \*\* $p \leq 0.01$ ; \*\*\* $p \leq 0.001$ . ns, not significant.

## Discussion

We report a simple and reproducible protocol to generate somatic hypermutations in transgenes with reduced toxicity using the m7.3 hyperactive AID mutant. The protocol relies on transient bicistronic expression of m7.3 and a puromycin resistance gene combined with short-term puromycin selection. The devised protocol was based on the finding that the mutation rate declines quickly with concurrent decline of cell proliferation rate after stable expression of m7.3. We predicted that limiting the expression time of m7.3 could reduce its adverse effects. Indeed, transient expression of m7.3 for 3–5 days under puromycin selection was an efficient to introduce mutations in a reporter gene with almost no adverse effects on cell fitness. We confirmed that mutagenesis can be controlled by puromycin selection, mutagenesis is very efficient when AID is introduced under selection and removal of puromycin greatly reduced AID level and mutation rate. Taken together, our study reports a simple way to diversify highly transcribed genes with reduced toxicity and without the need for library construction, protein engineering or recombinant antibody production.

Protein evolution requires repetitive cycles of introducing mutations and selection<sup>28</sup>. The efficiency of mutagenesis is a key factor for fast and efficient protein evolution<sup>29,30</sup>. *In vitro* mutagenesis by error prone PCR or DNA shuffling generates high mutation rates, but requires repeated cloning and transfection. *In vivo* mutagenesis methods require minimal intervention because selection and mutagenesis are coupled in the same system. Most *in vivo* mutagenesis methods suffer from low mutation rates and increased off-target mutagenesis<sup>31</sup>. Therefore, there is a need for an *in vivo* mutagenesis systems which can diversify a chosen gene of interest with a high frequency.

Natural somatic hypermutation is an efficient system for antibody diversification. *In vivo*, about one point mutation is generated in an immunoglobulin gene per cell division<sup>32</sup>. This requires the enzyme activation-induced cytidine deaminase (AID) which deaminates deoxycytidine residues to uracil in actively transcribed genes<sup>33–36</sup>.

Uracil is mutagenic when paired with guanosine in DNA, since dU mimics dT during replication, and the U:G mismatch triggers error-prone DNA repair in B cells. Thus, AID introduces somatic hypermutation (SHM) by converting dC to dU<sup>37–40</sup>. The power of SHM for protein evolution was first shown in a cell line that constitutively expresses AID (Ramos cells); Cumbers *et al.* were able to evolve endogenous IgM from unknown specificity to selectively bind streptavidin with nanomolar affinity after 19 rounds of FACS sorting<sup>41</sup>. Since that time, several studies reported using hypermutating cell lines like Ramos and DT40 for protein evolution<sup>4,42,43</sup>. Interestingly, ectopic expression of AID in non-B cells also turns on somatic hypermutation and diversifies highly transcribed transgenes<sup>18</sup>. Successful antibody evolution was reported after ectopic AID expression in human lung carcinoma<sup>44</sup>, HEK293<sup>10,45–47</sup> and CHO<sup>48</sup> cells.

A major drawback of inducing mutations by endogenous AID is that mutations are still far below that observed *in vivo*<sup>20,49,50</sup>. Ectopic overexpression of AID can partially solve this problem, but it increases off-target mutagenesis of highly expressed genes and transgenes such as AID itself<sup>20,51</sup>. Hypermutation rate *in vivo* has been estimated to be approximately  $10^{-3}$  mutations per base pair per generation; this frequency seems to be optimal as a lower frequency might generate poor diversity, whereas a higher mutation rate might cause increased cellular toxicity<sup>52–57</sup>.

To improve the efficiency of AID mutagenesis, Wang *et al.* generated human AID upmutants with improved activity, but unfortunately the increase in catalytic activity was accompanied with increased adverse effects on DNA instability and cells fitness<sup>22</sup>. However, the increased genomic instability associated with the AID upmutants was assayed by stable expression of AID upmutants in B cells<sup>22</sup>. By contrast, AID is transiently expressed for short periods in B cells during antibody affinity maturation<sup>26,58</sup>. Therefore, we decided to test transient expression of AID to induce mutations in transgenes with reduced cellular toxicity.

Transient transfection of AID for 3–5 days under puromycin selection was ideal to generate SHM with reduced toxicity. By contrast, transient transfection without any selection pressure produced extremely low mutation frequency. Since the puromycin resistance gene is a highly transcribed transgene, it has a chance to be targeted by AID. We used the shortest possible time for drug selection to avoid killing the cells that accumulate more mutations in the puromycin-resistance gene, which are probably the cells with better SHM activity. Puromycin was useful for pulse selection because it can kill the cells that did not express the gene in as short as 1 or 2 days using low drug concentration<sup>59–61</sup>.

An interesting application of the findings in this study is to mimic the natural cancer progression and metastasis mechanisms. There is a mounting evidence that mutagenesis by AID and other closely related deaminases play important roles in cancer evolution and metastasis<sup>62–64</sup>. It was hypothesized that transient but not constitutive mutagenesis facilitates cancer progression<sup>12,65</sup>. Indeed our data support this hypothesis because it seems that continuous mutagenesis might have eventual adverse effects on cancer fitness due to continuous accumulation of mutations, which might adversely affect cancer growth. This concept might be useful for researchers interested in studying cancer pathogenesis. For example, a cytidine deaminase can be expressed for a short time and cells with interesting phenotype can be selected. This could be a useful approach to allow researchers to isolate customized cancer clones based on the selection criteria they want.

In this study, we provide a simple solution for fast, controllable, low toxicity mutagenesis of highly transcribed genes by using transient hyperactive AID expression. AID gene transfection generates a library of mutants for the gene of interest inside the cells without the need for using error-prone PCR or DNA shuffling. Proteins with improved properties can be isolated, propagated and further mutagenized by additional rounds of mutagenesis by AID upmutant. Besides selection for improved protein function, mutants can also be selected for other desired properties such as solubility, expression, and post-translational modification.

## Material and Methods

**Cell lines and culture conditions.** Ramos human Burkitt's lymphoma B cells (CRL-1596, American Type Culture Collection, Manassas, VA, USA) were grown in RPMI 1640 medium supplemented with 2 mM L-glutamine and 10% heat-inactivated fetal bovine serum (HyClone, Logan, UT, USA). BALB/3T3 fibroblasts and HEK293 cells were obtained from ATCC (Manassas, VA, USA). HEK293FT human embryonic kidney cells (Thermo Fisher Scientific, San Jose, CA, USA). BALB/3T3 fibroblasts, HEK293 and HEK293FT were maintained in Dulbecco's modified Eagle's medium (Sigma, Saint Louis, MO, USA) supplemented with 10% fetal calf serum, 100 µg/mL streptomycin and 100 U/ml penicillin grown in a humidified incubator at 37 °C in 5% CO<sub>2</sub> and 95% air. 3T3/TetOn-AID cells<sup>10</sup> were maintained in Dulbecco's modified Eagle's medium supplemented with 10% dialyzed fetal calf serum, 100 µg/mL streptomycin and 100 U/ml penicillin grown in a humidified incubator at 37 °C in 5% CO<sub>2</sub> and 95% air.

**Generation of Reporter cells.** 293FT/mRFP1 and 293FT/DsRed2s reporter cells were generated to assess mutation rates. The mRFP1 gene was amplified from mRFP1 pRSET B plasmid which was a kind gift from Dr. R. Y. Tsien (University of California, San Diego, USA). mRFP1 was amplified by PCR with a 5' primer 5'-ctagcggcgcgcccatgtgagcaagggcgagg-3' that encodes an *AscI* restriction site and a 3' primer 5'-gagaggtaacttagcgcgggtggagtgccggc-3' that encodes a *HpaI* site. mRFP was ligated into the backbone of pLKO\_AS3w-Hygro (National RNAi Core Facility, Institute of Molecular Biology/Genomic Research Center, Academia Sinica, Taiwan) to generate pLKO\_AS3w-RFP1 lentiviral vector for mammalian cell expression.

The DsRed2s gene was amplified by PCR from the pLHCX-DsRed2s.hyg plasmid<sup>10</sup> using the forward primer 5'-ggaggtggcgcgcccatgtgagcaagggcgagg-3' and reverse primer 5'ggtggagttaaaccggtcaactgtatctggccactg-3' and then inserted into *AscI* and *PmeI* sites downstream of the CMV promoter in pLKO\_AS3w-Hygro to generate pLKO\_AS3w-DsRed2s lentiviral vector.

Lentiviral plasmids encoding mRFP1 and DsRed2s were co-transfected with pCMVΔR8.91 packaging plasmid and pMD.G VSV-G envelope plasmid in HEK293FT cells to produce lentiviral particles<sup>10</sup>. Two days later, the

culture medium was filtered, mixed with 8 µg/ml polybrene, and incubated with HEK293FT cells at a multiplicity of infection (MOI) of 0.1 to ensure single copy of reporter gene per cell. Stable cell lines were selected with 200 µg/ml hygromycin for two weeks.

**AID gene amplification.** Ramos cells ( $1 \times 10^6$ ) were lysed with 1 ml Super RNApure reagent and RNA was extracted according to the manufacturer's instructions (Genesis Biotech Inc., Taipei, Taiwan). Five micrograms of RNA were reverse transcribed using the Superscript III first-strand synthesis system (Thermo Fisher Scientific). Phusion high fidelity DNA polymerase was used to amplify the AID gene using primers 5'-ctagtgcctagcgccaccatggacagcctcttgatgaac-3' and 5'-cgtgagttaacacctccaagtccaaa-gtacgaa-atgc-3'.

**AID mutant construction.** PCR products were cloned by Zero Blunt PCR cloning kit (Thermo Fisher Scientific) and sequenced using standard M13 forward and reverse primers. Multiple site mutagenesis was performed as described<sup>66</sup> in the wild-type human AID gene to generate the AID m7.3 mutant (K10E, T82I, and E156G) with increased catalytic activity<sup>22</sup>. Site directed mutagenesis using QuikChange™ Site-Directed Mutagenesis Kit (Stratagene, La Jolla, CA, USA) was used to generate the AID S38A mutant with compromised activity<sup>67</sup> and AID-M6A with total loss of activity<sup>24</sup>. Wild type AID and mutants AID genes were cloned into a lentiviral vector pAS3w.Ppuro<sup>10</sup> by the *HpaI* and *NheI* restriction sites (New England BioLabs, Ipswich, MA, USA). An HA tag epitope (YPYDVPDYA) was fused to the C-terminus of AID and AID mutants to allow detection by immunoblot analysis.

**Stable expression of AID.** Lentiviral plasmids encoding AID mutants were used to generate recombinant lentiviral particles. Virus packaging was done by co-transfection of 7.5 µg of AID, m7.3, S38A, or M6A with pCMVΔR8.91 (6.75 µg) and pMD.G (0.75 µg) using TransIT-LT1 transfection reagent (Mirus Bio, Madison, WI, USA) (45 µL) in HEK293FT cells grown in a 10 cm culture dish (90% confluency). After 48 h, lentiviral particles were suspended in culture medium containing 8 µg/mL polybrene and filtered through a 0.45 µm filter. 293FT/mRFP1 or 293FT/DsRed2s cells were seeded in 6-well plates ( $1 \times 10^5$  cells/well) one day before viral infection. Lentivirus containing medium was added to the plate and then centrifuged for 1.5 h (500 g, 32 °C). The cells were selected in puromycin (2 µg/mL) for 4 days to generate stable cell lines.

**Transient expression of AID.** Transient transfection was performed using TransIT-LT1 Transfection Reagent (Mirus Bio) according to the manufacturer's instructions. Briefly, 293FT, 293FT/RFP1 or 293FT/DsRed2s cells were seeded in 6-well plates in DMEM containing 10% FBS without antibiotics one day before transfection, such that they were 70–90% confluent at the time of transfection. On the day of transfection, 2.5 µg of AID lentiviral vectors and 7.5 µL of TransIT-LT1 were combined in 500 µL of Opti-MEM Reduced Serum Medium (Thermo Fisher Scientific). The mixture was incubated for an additional 25 min at room temperature and then added to each well, and the cells were incubated at 37 °C in a CO<sub>2</sub> incubator for 24 hours. Cells were then refreshed with DMEM containing 10% FBS for one day, and then selected with 2 µg/mL puromycin for 2 or 4 days.

**Immunoblotting.** To measure AID protein levels,  $5 \times 10^6$  cells were lysed in 0.5 mL RIPA buffer (1% NP-40, 150 mM NaCl, 0.5% sodium deoxycholate, 0.1% SDS, 50 mM Tris, pH 8.0) for 1 h at 4 °C. Fifty µg of total protein from the clarified lysates were separated on a 12.5% reducing SDS-PAGE, transferred to nitrocellulose paper and sequentially stained with biotinylated goat anti-HA antibody (Vector Laboratories, CA, USA) or rabbit anti-tubulin α antibody (NeoMarkers, CA, USA) followed by streptavidin-HRP and goat anti-rabbit Ig-HRP, respectively (Jackson ImmunoResearch Laboratories, PA, USA). Bands were visualized by ECL detection (Thermo Fisher Scientific) and analyzed with a LAS-3000 Mini Fujifilm imaging system (FujiFilm, Tokyo, Japan).

**Flow cytometry.** The percentage of RFP1 negative or DsRed2 positive cells was determined by flow cytometry. Cells were washed with phosphate buffered saline (PBS) and then incubated with 1 µM Sytox Blue viability dye (Thermo Fisher Scientific) in PBS containing 0.05% bovine serum albumin (BSA) at room temperature for 30 minutes before measurement on a SRII flow cytometry (Becton Dickinson, Mountain View, CA, USA), with analysis by FlowJo software (Tree Star, San Carlos, CA, USA).

**Cell proliferation assay.** Two days after transient or stable transfection of HEK293FT cells with AID mutants, cells were selected in medium containing 2 µg/mL puromycin for 4 days, and refreshed with puromycin-free media for 1 day. Similar numbers of cells were seeded in 6 well plates and viable cells were counted every 2 days by trypan blue dye exclusion using a Countess Automated Cell Counter (Thermo Fisher Scientific).

Stably infected 3T3/TetOn-AID was reported previously. Cells were seeded in 6 cm plates with or without 0.5 µg/mL doxycycline, and viable cells are counted every 2 days for 6 days.

**Sequencing.** Mutations in the DsRed2 gene were characterized by sequencing of genomic DNA that was PCR-amplified from  $10^6$  cells 2 weeks after transfection. The PCR product was ligated to zero-blut vector (Thermo Fisher Scientific) and sequenced to confirm the mutations.

## References

1. Tobin, P. H., Richards, D. H., Callender, R. A. & Wilson, C. J. Protein engineering: a new frontier for biological therapeutics. *Curr Drug Metab* **15**, 743–756 (2014).
2. Chen, Z. & Zeng, A. P. Protein engineering approaches to chemical biotechnology. *Curr Opin Biotechnol* **42**, 198–205, <https://doi.org/10.1016/j.copbio.2016.07.007> (2016).
3. Petersen-Mahrt, S. K., Harris, R. S. & Neuberger, M. S. AID mutates E. coli suggesting a DNA deamination mechanism for antibody diversification. *Nature* **418**, 99–103, <https://doi.org/10.1038/nature00862> (2002).

4. Wang, L., Jackson, W. C., Steinbach, P. A. & Tsien, R. Y. Evolution of new nonantibody proteins via iterative somatic hypermutation. *Proc Natl Acad Sci USA* **101**, 16745–16749, <https://doi.org/10.1073/pnas.0407752101> (2004).
5. Ho, M., Nagata, S. & Pastan, I. Isolation of anti-CD22 Fv with high affinity by Fv display on human cells. *Proc Natl Acad Sci USA* **103**, 9637–9642, <https://doi.org/10.1073/pnas.0603653103> (2006).
6. Akamatsu, Y., Pakabunto, K., Xu, Z., Zhang, Y. & Tsurushita, N. Whole IgG surface display on mammalian cells: Application to isolation of neutralizing chicken monoclonal anti-IL-12 antibodies. *J Immunol Methods* **327**, 40–52, <https://doi.org/10.1016/j.jim.2007.07.007> (2007).
7. Ho, M. & Pastan, I. Display and selection of scFv antibodies on HEK-293T cells. *Methods Mol Biol* **562**, 99–113, [https://doi.org/10.1007/978-1-60327-302-2\\_8](https://doi.org/10.1007/978-1-60327-302-2_8) (2009).
8. McConnell, A. D. *et al.* High affinity humanized antibodies without making hybridomas; immunization paired with mammalian cell display and *in vitro* somatic hypermutation. *PLoS One* **7**, e49458 (2012).
9. Bowers, P. M. *et al.* Coupling mammalian cell surface display with somatic hypermutation for the discovery and maturation of human antibodies. *Proceedings of the National Academy of Sciences* **108**, 20455–20460 (2011).
10. Su, Y.-C. *et al.* Mimicking the germinal center reaction in hybridoma cells to isolate temperature-selective anti-PEG antibodies. *MAbs* **6**, 39–38 (2014).
11. Lada, A. G. *et al.* Genome-wide mutation avalanches induced in diploid yeast cells by a base analog or an APOBEC deaminase. *PLoS genetics* **9**, e1003736, <https://doi.org/10.1371/journal.pgen.1003736> (2013).
12. Roberts, S. A. & Gordenin, D. A. Clustered and genome-wide transient mutagenesis in human cancers: Hypermutation without permanent mutators or loss of fitness. *BioEssays: news and reviews in molecular, cellular and developmental biology*, <https://doi.org/10.1002/bies.201300140> (2014).
13. Klemm, L. *et al.* The B cell mutator AID promotes B lymphoid blast crisis and drug resistance in chronic myeloid leukemia. *Cancer cell* **16**, 232–245, <https://doi.org/10.1016/j.ccr.2009.07.030> (2009).
14. Zahn, A. *et al.* Activation induced deaminase C-terminal domain links DNA breaks to end protection and repair during class switch recombination. *Proc Natl Acad Sci USA* **111**, E988–997, <https://doi.org/10.1073/pnas.1320486111> (2014).
15. Wang, C. L., Harper, R. A. & Wabl, M. Genome-wide somatic hypermutation. *Proc Natl Acad Sci USA* **101**, 7352–7356, <https://doi.org/10.1073/pnas.0402009101> (2004).
16. Wang, M., Yang, Z., Rada, C. & Neuberger, M. S. AID upmutants isolated using a high-throughput screen highlight the immunity/cancer balance limiting DNA deaminase activity. *Nat. Struct. Mol. Biol.* **16**, 769–776 (2009).
17. Fukita, Y., Jacobs, H. & Rajewsky, K. Somatic hypermutation in the heavy chain locus correlates with transcription. *Immunity* **9**, 105–114 (1998).
18. Yoshikawa, K. *et al.* AID enzyme-induced hypermutation in an actively transcribed gene in fibroblasts. *Science* **296**, 2033–2036, <https://doi.org/10.1126/science.1071556296/5575/2033> [pii] (2002).
19. Bachl, J., Carlson, C., Gray-Schopfer, V., Dessing, M. & Olsson, C. Increased transcription levels induce higher mutation rates in a hypermutating cell line. *J Immunol* **166**, 5051–5057 (2001).
20. Martin, A. & Scharff, M. D. Somatic hypermutation of the AID transgene in B and non-B cells. *Proc Natl Acad Sci USA* **99**, 12304–12308, <https://doi.org/10.1073/pnas.192442899> (2002).
21. Wang, X., Fan, M., Kalis, S., Wei, L. & Scharff, M. D. A source of the single-stranded DNA substrate for activation-induced deaminase during somatic hypermutation. *Nat Commun* **5**, 4137, <https://doi.org/10.1038/ncomms5137> (2014).
22. Wang, M., Yang, Z., Rada, C. & Neuberger, M. S. AID upmutants isolated using a high-throughput screen highlight the immunity/cancer balance limiting DNA deaminase activity. *Nature structural & molecular biology* **16**, 769–776, <https://doi.org/10.1038/nsmb.1623> (2009).
23. Cheng, H. L. *et al.* Integrity of the AID serine-38 phosphorylation site is critical for class switch recombination and somatic hypermutation in mice. *Proc Natl Acad Sci USA* **106**, 2717–2722, <https://doi.org/10.1073/pnas.0812304106> (2009).
24. Okazaki, I. M. *et al.* Histone chaperone Spt6 is required for class switch recombination but not somatic hypermutation. *Proc Natl Acad Sci USA* **108**, 7920–7925, <https://doi.org/10.1073/pnas.1104423108> (2011).
25. Mahon, M. J. Vectors bicistronically linking a gene of interest to the SV40 large T antigen in combination with the SV40 origin of replication enhance transient protein expression and luciferase reporter activity. *Biotechniques* **51**, 119–128, <https://doi.org/10.2144/000113720> (2011).
26. Greeve, J. *et al.* Expression of activation-induced cytidine deaminase in human B-cell non-Hodgkin lymphomas. *Blood* **101**, 3574–3580, <https://doi.org/10.1182/blood-2002-08-2424> (2003).
27. Cattoretti, G. *et al.* Stages of germinal center transit are defined by B cell transcription factor coexpression and relative abundance. *J Immunol* **177**, 6930–6939 (2006).
28. Mossner, E. & Pluckthun, A. Directed evolution with fast and efficient selection technologies. *Chimia* **55**, 325–329 (2001).
29. Heydenreich, F. M. *et al.* High-throughput mutagenesis using a two-fragment PCR approach. *Sci Rep-Uk* **7**, <https://doi.org/10.1038/s41598-017-07010-4> (2017).
30. Drummond, D. A., Iverson, B. L., Georgiou, G. & Arnold, F. H. Why high-error-rate random mutagenesis libraries are enriched in functional and improved proteins. *Journal of molecular biology* **350**, 806–816, <https://doi.org/10.1016/j.jmb.2005.05.023> (2005).
31. Bratulic, S. & Badran, A. H. Modern methods for laboratory diversification of biomolecules. *Curr Opin Chem Biol* **41**, 50–60, <https://doi.org/10.1016/j.cbpa.2017.10.010> (2017).
32. Rajewsky, K., Forster, I. & Cumano, A. Evolutionary and Somatic Selection of the Antibody Repertoire in the Mouse. *Science* **238**, 1088–1094, <https://doi.org/10.1126/science.3317826> (1987).
33. Muramatsu, M. *et al.* Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. *Cell* **102**, 553–563 (2000).
34. Bransteitter, R., Pham, P., Scharff, M. D. & Goodman, M. F. Activation-induced cytidine deaminase deaminates deoxycytidine on single-stranded DNA but requires the action of RNase. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 4102–4107, <https://doi.org/10.1073/pnas.0730835100> (2003).
35. Pham, P., Bransteitter, R., Petruska, J. & Goodman, M. F. Processive AID-catalysed cytosine deamination on single-stranded DNA simulates somatic hypermutation. *Nature* **424**, 103–107, <https://doi.org/10.1038/nature01760> (2003).
36. Peters, A. & Storb, U. Somatic hypermutation of immunoglobulin genes is linked to transcription initiation. *Immunity* **4**, 57–65, [https://doi.org/10.1016/S1074-7613\(00\)80298-8](https://doi.org/10.1016/S1074-7613(00)80298-8) (1996).
37. Maul, R. W. & Gearhart, P. J. AID and somatic hypermutation. *Advances in immunology* **105**, 159–191, [https://doi.org/10.1016/S0065-2776\(10\)05006-6](https://doi.org/10.1016/S0065-2776(10)05006-6) (2010).
38. Zanotti, K. J. & Gearhart, P. J. Antibody diversification caused by disrupted mismatch repair and promiscuous DNA polymerases. *DNA Repair* **38**, 110–116, <https://doi.org/10.1016/j.dnarep.2015.11.011> (2016).
39. Chahwan, R., Edelmann, W., Scharff, M. D. & Roa, S. AIDing antibody diversity by error-prone mismatch repair. *Semin Immunol* **24**, 293–300, <https://doi.org/10.1016/j.simm.2012.05.005> (2012).
40. Supek, F. & Lehner, B. Clustered Mutation Signatures Reveal that Error-Prone DNA Repair Targets Mutations to Active Genes. *Cell* **170**, 534–+, <https://doi.org/10.1016/j.cell.2017.07.003> (2017).
41. Cumbers, S. J. *et al.* Generation and iterative affinity maturation of antibodies *in vitro* using hypermutating B-cell lines. *Nature biotechnology* **20**, 1129–1134, <https://doi.org/10.1038/nbt752> (2002).

42. Seo, H. *et al.* An *ex vivo* method for rapid generation of monoclonal antibodies (ADLib system). *Nat Protoc* **1**, 1502–1506, <https://doi.org/10.1038/nprot.2006.248> (2006).
43. Arakawa, H. *et al.* Protein evolution by hypermutation and selection in the B cell line DT40. *Nucleic Acids Res* **36**, e1, <https://doi.org/10.1093/nar/gkm616> (2008).
44. Chen, S. *et al.* Affinity maturation of anti-TNF-alpha scFv with somatic hypermutation in non-B cells. *Protein & Cell* **3**, 460–469, <https://doi.org/10.1007/s13238-012-2024-7> (2012).
45. McConnell, A. D. *et al.* High affinity humanized antibodies without making hybridomas; immunization paired with mammalian cell display and *in vitro* somatic hypermutation. *PLoS One* **7**, e49458, <https://doi.org/10.1371/journal.pone.0049458> (2012).
46. Horlick, R. A. *et al.* Simultaneous surface display and secretion of proteins from mammalian cells facilitate efficient *in vitro* selection and maturation of antibodies. *Journal of Biological Chemistry* **288**, 19861–19869 (2013).
47. J King, D., M Bowers, P., R Kehry, M. & A Horlick, R. Mammalian cell display and somatic hypermutation *in vitro* for human antibody discovery. *Current drug discovery technologies* **11**, 56–64 (2014).
48. Chen, C., Li, N., Zhao, Y. & Hang, H. Coupling recombinase-mediated cassette exchange with somatic hypermutation for antibody affinity maturation in CHO cells. *Biotechnology and bioengineering* **113**, 39–51, <https://doi.org/10.1002/bit.25541> (2016).
49. Peled, J. U. *et al.* The biochemistry of somatic hypermutation. *Annual review of immunology* **26**, 481–511, <https://doi.org/10.1146/annurev.immunol.26.021607.090236> (2008).
50. Parsa, J. Y. *et al.* AID mutates a non-immunoglobulin transgene independent of chromosomal position. *Molecular immunology* **44**, 567–575, <https://doi.org/10.1016/j.molimm.2006.02.003> (2007).
51. Woo, C. J., Martin, A. & Scharff, M. D. Induction of somatic hypermutation is associated with modifications in immunoglobulin variable region chromatin. *Immunity* **19**, 479–489, [https://doi.org/10.1016/S1074-7613\(03\)00261-9](https://doi.org/10.1016/S1074-7613(03)00261-9) (2003).
52. Clarke, S. H. *et al.* Inter- and intraclonal diversity in the antibody response to influenza hemagglutinin. *The Journal of experimental medicine* **161**, 687–704 (1985).
53. Berek, C. & Milstein, C. Mutation drift and repertoire shift in the maturation of the immune response. *Immunol Rev* **96**, 23–41 (1987).
54. Celada, F. & Seiden, P. E. Affinity maturation and hypermutation in a simulation of the humoral immune response. *European journal of immunology* **26**, 1350–1358, <https://doi.org/10.1002/eji.1830260626> (1996).
55. Kepler, T. B. & Perelson, A. S. Somatic hypermutation in B cells: an optimal control treatment. *J Theor Biol* **164**, 37–64, <https://doi.org/10.1006/jtbi.1993.1139> (1993).
56. Kepler, T. B. & Perelson, A. S. Modeling and optimization of populations subject to time-dependent mutation. *Proc Natl Acad Sci USA* **92**, 8219–8223 (1995).
57. Kleinstein, S. H., Louzoun, Y. & Shlomchik, M. J. Estimating hypermutation rates from clonal tree data. *J Immunol* **171**, 4639–4649 (2003).
58. Allen, C. D., Okada, T. & Cyster, J. G. Germinal-center organization and cellular dynamics. *Immunity* **27**, 190–202, <https://doi.org/10.1016/j.immuni.2007.07.009> (2007).
59. Taniguchi, M. *et al.* Efficient production of Cre-mediated site-directed recombinants through the utilization of the puromycin resistance gene, *pac*: a transient gene-integration marker for ES cells. *Nucleic Acids Res* **26**, 679–680 (1998).
60. Li, M. A. *et al.* Mobilization of giant piggyBac transposons in the mouse genome. *Nucleic Acids Res* **39**, e148, <https://doi.org/10.1093/nar/gkr764> (2011).
61. Haemmerle, G. *et al.* Hormone-sensitive lipase deficiency in mice causes diglyceride accumulation in adipose tissue, muscle, and testis. *J Biol Chem* **277**, 4806–4815 (2002).
62. Caval, V., Suspene, R., Shapira, M., Vartanian, J. P. & Wain-Hobson, S. A prevalent cancer susceptibility APOBEC3A hybrid allele bearing APOBEC3B 3'UTR enhances chromosomal DNA damage. *Nat Commun* **5**, 5129, <https://doi.org/10.1038/ncomms6129> (2014).
63. Henderson, S. & Fenton, T. APOBEC3 genes: retroviral restriction factors to cancer drivers. *Trends Mol Med* **21**, 274–284, <https://doi.org/10.1016/j.molmed.2015.02.007> (2015).
64. Roberts, S. A. & Gordenin, D. A. Hypermutation in human cancer genomes: footprints and mechanisms. *Nature reviews. Cancer* **14**, 786–800, <https://doi.org/10.1038/nrc3816> (2014).
65. Chan, K. & Gordenin, D. A. Clusters of Multiple Mutations: Incidence and Molecular Mechanisms. *Annu Rev Genet* **49**, 243–267, <https://doi.org/10.1146/annurev-genet-112414-054714> (2015).
66. Sawano, A. & Miyawaki, A. Directed evolution of green fluorescent protein by a new versatile PCR strategy for site-directed and semi-random mutagenesis. *Nucleic acids research* **28**, E78 (2000).
67. Basu, U. *et al.* Regulation of activation-induced cytidine deaminase DNA deamination activity in B-cells by Ser38 phosphorylation. *Biochemical Society transactions* **37**, 561–568, <https://doi.org/10.1042/BST0370561> (2009).

## Acknowledgements

This work was supported by Academia Sinica and a research grant from the National Science Council, Taipei, Taiwan (MOST 106-2321-B-001-030). We thank Ms. Chia-Chen Tai and Ms. Tzu-Wen Tai of the Flow Cytometry Core, Scientific Instrument Center at the Institute of Biomedical Sciences, Academia Sinica, Taipei, Taiwan, for assistance with the LSR II flow cytometer. We appreciate the technical help of Ms. Wei-Chiao Huang and Ms. Chia-Yi Kao.

## Author Contributions

T.S.A.Q. and S.R.R. conceived the project, designed the experiments, analyzed data and wrote the manuscript. T.S.A.Q. performed all experiments except for experiments presented in Fig. 6(C–E) which was done by Y.C.S.

## Additional Information

**Competing Interests:** The authors declare no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018