

## RESEARCH ARTICLE SUMMARY

## GENOME EDITING

# Targeted nucleotide editing using hybrid prokaryotic and vertebrate adaptive immune systems

Keiji Nishida, Takayuki Arazoe, Nozomu Yachie, Satomi Banno, Mika Kakimoto, Mayura Tabata, Masao Mochizuki, Aya Miyabe, Michihiro Araki, Kiyotaka Y. Hara, Zenpei Shimatani, Akihiko Kondo\*

**INTRODUCTION:** To combat invading pathogens, cells develop an adaptive immune response by changing their own genetic information. In vertebrates, the generation of genetic variation (somatic hypermutation) is an essential process for diversification and affinity maturation of

## ON OUR WEBSITE

Read the full article at <http://dx.doi.org/10.1126/science.aaf8729>

antibodies that function to detect and sequester various foreign biomolecules. The activation-induced cytosine deaminase (AID) carries out hypermutation by modifying deoxycytidine bases in the variable region of the immunoglobulin locus that produces antibody. AID-generated deoxyuridine in DNA is mutagenic as it can be miss-recognized as deoxythymine, resulting in C to T mutations. CRISPR (clustered regularly interspaced short palindromic repeats)/Cas (CRISPR-associated) is a prokaryotic adaptive immune system that records and degrades invasive foreign DNA or RNA. The CRISPR/Cas system cleaves and incorporates foreign DNA/RNA segments into the genomic region called the CRISPR array. The CRISPR array is transcribed to produce crisp-rRNA that serves as guide RNA (gRNA) for recognition of the complementary foreign DNA/RNA in a ribonucleoprotein complex with Cas proteins, which degrade the target. The CRISPR/Cas system has been repurposed as

a powerful genome editing tool, because it can be programmed to cleave specific DNA sequence by providing custom gRNAs.

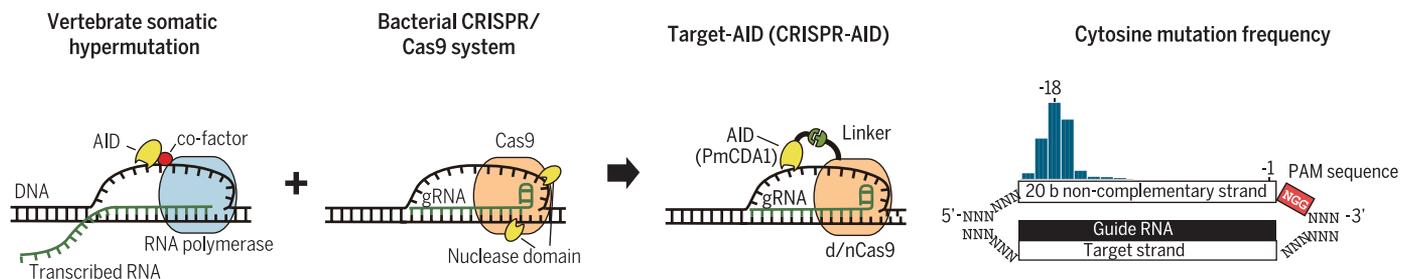
**RATIONALE:** Although the precise mechanism by which AID specifically mutates the immunoglobulin locus remains elusive, targeting of AID activity is facilitated by the formation of a single-stranded DNA region, such as a transcriptional RNA/DNA hybrid (R-loop). The CRISPR/Cas system can be engineered to be nuclease-inactive. The nuclease-inactive form is capable of unfolding the DNA double strand in a protospacer adjacent motif (PAM) sequence-dependent manner so that the gRNA binds to complementary target DNA strand and forms an R-loop. The nuclease-deficient CRISPR/Cas system may serve as a suitable DNA-targeting module for AID to catalyze site-specific mutagenesis.

**RESULTS:** To determine whether AID activity can be specifically targeted by the CRISPR/Cas system, we combined dCas9 (a nuclease-deficient mutant of Cas9) from *Streptococcus pyogenes* and an AID ortholog, PmCDA1 from sea lamprey, to form a synthetic complex (Target-AID) by either engineering a fusion between the two proteins or attaching a SH3 (Src 3 homology) domain to the C terminus of dCas9 and a SHL (SH3 interaction ligand) to the C terminus of PmCDA1. Both of

these complexes performed highly efficient site-directed mutagenesis. The mutational spectrum was analyzed in yeast and demonstrated that point mutations were dominantly induced at cytosines within the range of three to five bases surrounding the -18 position upstream of the PAM sequence on the noncomplementary strand to gRNA. The toxicity associated with the nuclease-based CRISPR/Cas9 system was greatly reduced in the Target-AID complexes. Combination of PmCDA1 with the nickase Cas9 (D10A) mutant, which retains cleavage activity for noncomplementary single-stranded DNA, was more efficient in yeast but also induced deletions as well as point mutations in mammalian cells. Addition of the uracil DNA glycosylase inhibitor protein, which blocks the initial step of the uracil base excision repair pathway, suppressed collateral deletions and further improved targeting efficiency. Potential off-target effects were assessed by whole-genome sequencing of yeast as well as deep sequencing of mammalian cells for regions that contain mismatched target sequences. These results showed that off-target effects were comparable to those of conventional CRISPR/Cas systems, with a reduced risk of indel formation.

**CONCLUSION:** By expanding the genome editing potential of the CRISPR/Cas9 system by deaminase-mediated hypermutation, Target-AID demonstrated a very narrow range of targeted nucleotide substitution without the use of template DNA. Nickase Cas9 and uracil DNA glycosylase inhibitor protein can be used to boost the targeting efficiency. The reduced cytotoxicity will be beneficial for use in cells that are sensitive to artificial nucleases. Use of other types of nucleotide-modifying enzymes and/or other CRISPR-related systems with different PAM requirements will expand our genome-editing repertoire and capacity. ■

The list of author affiliations is available in the full article online.  
\*Corresponding author. Email: [akondo@kobe-u.ac.jp](mailto:akondo@kobe-u.ac.jp)  
Cite this article as K. Nishida *et al.*, *Science* 353, aaf8729 (2016). DOI: 10.1126/science.aaf8729



**A crippled CRISPR/Cas targets AID.** In vertebrate adaptive immunity, cytosine deaminase (AID or PmCDA1) induces somatic hypermutation at single-stranded DNA regions formed during transcription. The bacterial CRISPR/Cas9 immunity system recognizes and cleaves invasive DNA in a gRNA-dependent manner. AID and nuclease-deficient CRISPR/Cas9 are engineered to form a hybrid complex (Target-AID) that performs programmable cytosine mutations in a range of a few bases surrounding the -18 position upstream of PAM sequence of the noncomplementary DNA strand.

## RESEARCH ARTICLE

## GENOME EDITING

# Targeted nucleotide editing using hybrid prokaryotic and vertebrate adaptive immune systems

Keiji Nishida,<sup>1</sup> Takayuki Arazoe,<sup>1</sup> Nozomu Yachie,<sup>2,3,4</sup> Satomi Banno,<sup>1</sup> Mika Kakimoto,<sup>1</sup> Mayura Tabata,<sup>1</sup> Masao Mochizuki,<sup>1</sup> Aya Miyabe,<sup>1</sup> Michihiro Araki,<sup>1</sup> Kiyotaka Y. Hara,<sup>5</sup> Zenpei Shimatani,<sup>1</sup> Akihiko Kondo<sup>1,6\*</sup>

The generation of genetic variation (somatic hypermutation) is an essential process for the adaptive immune system in vertebrates. We demonstrate the targeted single-nucleotide substitution of DNA using hybrid vertebrate and bacterial immune systems components. Nuclease-deficient type II CRISPR/Cas9 (clustered regularly interspaced short palindromic repeats/CRISPR-associated) and the activation-induced cytidine deaminase (AID) ortholog PmCDA1 were engineered to form a synthetic complex (Target-AID) that performs highly efficient target-specific mutagenesis. Specific point mutation was induced primarily at cytidines within the target range of five bases. The toxicity associated with the nuclease-based CRISPR/Cas9 system was greatly reduced. Although combination of nickase Cas9(D10A) and the deaminase was highly effective in yeasts, it also induced insertion and deletion (indel) in mammalian cells. Use of uracil DNA glycosylase inhibitor suppressed the indel formation and improved the efficiency.

The activation-induced cytidine deaminase (AID) is responsible for targeted hypermutation (1) by modifying the deoxycytidine of the variable region of the immunoglobulin locus (2, 3). Although the precise mechanism by which AID specifically targets the immunoglobulin locus remains elusive, targeting of AID is facilitated by the formation of a single-stranded DNA region, such as a transcriptional RNA/DNA hybrid (R-loop) (4, 5). AID generates deoxyuridine in DNA, which is further processed by uracil DNA glycosylase for class-switch recombination (6) as well as repair (7). CRISPR (clustered regularly interspaced short palindromic repeats)/Cas (CRISPR-associated) is a prokaryotic adaptive immune system (8). The CRISPR/Cas system records and cleaves invasive foreign DNA. A ribonucleoprotein complex of Cas nuclease and guide-RNA (gRNA) binds to complementary target DNA sequences, forms an R-loop, and then cleaves the DNA. The CRISPR/Cas system has been developed as a ge-

nome editing tool (9, 10). Here, by using hybrid vertebrate and bacterial immune systems, we provide genome engineering tools that enable targeted nucleotide substitution without integration or the use of template DNA.

## Cytidine deamination is targeted by CRISPR/Cas system

To determine whether the single-stranded DNA region formed by the CRISPR/Cas system also serves as a preferable substrate for AID, we created a synthetic complex (Target-AID) of dCas9 (a nuclease-deficient mutant of Cas9) (11) from *Streptococcus pyogenes* and an AID ortholog, PmCDA1 from sea lamprey by either a fusion protein or attaching a SH3 (Src 3 homology) domain (12) to the C terminus of dCas9 and a SHL (SH3 interaction ligand) to the C terminus of PmCDA1 (Fig. 1A). The protein complex was expressed in *Saccharomyces cerevisiae*, which lacks an endogenous AID-related system. Chimeric gRNAs (13) were designed (13) to target a genomic negative selectable marker gene, CAN1, whose loss of function results in resistance to the drug canavanine. The rate of occurrence of the canavanine-resistant mutant was counted as an on-target mutation. To assess the potential genome-wide, nonspecific mutator phenotype (off-target mutation), another negatively selectable marker, LYPI, was monitored by resistance to S-aminoethyl-L-cysteine. Although the expression of PmCDA1 increased the background mutation frequency, the specific mutation frequency of CAN1 increased by more than 1000-fold over PmCDA1 alone when dCas9-SH3 and SHL-PmCDA1 were expressed simulta-

neously (Fig. 1B, 4). Specific mutation was observed even without synthetic attachment of dCas9 and PmCDA1 (Fig. 1B, 3), suggesting that the single-stranded DNA formed by dCas9-gRNA is a preferred substrate for PmCDA1. The fusion protein of dCas9 and PmCDA1 was also effective, especially when a large peptide linker (100 amino acids) was used (Fig. 1B, 6). Most mutations were found at g782, which is located at the -16 position upstream of the protospacer adjacent motif (PAM) sequence (11), except for PmCDA1 or dCas9 alone (Fig. 1C and table S1), which indicates that the steric effect of the protein complex affects the efficiency but not the location of the mutation. This also suggests that a limited region of DNA is made accessible for PmCDA1 by dCas9-gRNA. Most mutations were found as point mutations at a G and C, which suggests that deamination occurred at the noncomplementary strand of the C (table S1). The fusion protein with the extended linker version (Fig. 1B, 6) was used in the following studies because of its simplicity.

## DNA nicking promotes deamination-mediated mutagenesis

Because nucleotide deamination also occurs spontaneously, cells are capable of correcting most deaminated bases, and only a small portion of deamination normally results in mutagenesis (7). To exploit the full potential of Target-AID, we sought ways to bypass the nucleotide repair pathway without compromising overall genome integrity. Because the repair pathway requires base excision and template-dependent polymerization of nucleotides, we hypothesized that the introduction of a single-strand break (nick) in the DNA close to the deaminated site may perturb the repair system and result in a greater rate of mutagenesis. The Cas9 protein contains two nuclease domains, and each cleaves the opposite strand of DNA independently (11). Substituting aspartic acid with alanine in position 10 and histidine with alanine at position 840 prevents cleavage of the noncomplementary and complementary strands of DNA, respectively, to produce the Cas9 nickases nCas9(D10A) and nCas9(H840A) (11). Both nickases alone demonstrated a modest increase in the specific mutation rate, whereas the fusion protein of nCas9(D10A) and PmCDA1 consistently demonstrated an almost saturating mutation frequency (Fig. 2A), indicating that introducing nick on the opposite strand of deamination site facilitates mutagenesis, either by promoting the repair of the nicked, nonedited strand using the edited strand as a template or by providing better accessibility for PmCDA1. The mutational spectrum of nCas9(D10A)-PmCDA1 was similar to that of dCas9-PmCDA1 and distinct from those of Cas9 nickases or the complete nuclease alone (fig. S1). nCas9(D10A) has been shown to induce C-T substitution on the noncomplementary strand of target sites in cancer cell lines that overexpress endogenous cytidine deaminases, suggesting that nCas9-induced base substitution is mediated by the deaminase activity (14). This is consistent with our observation that the single-stranded

<sup>1</sup>Graduate School of Science, Technology and Innovation, Kobe University, 1-1 Rokkodai-cho, Nada-ku, Kobe, Hyogo 657-8501, Japan. <sup>2</sup>Synthetic Biology Division, Research Center for Advanced Science and Technology, the University of Tokyo, 4-6-1 Komaba, Meguro-ku, Tokyo 153-8904, Japan.

<sup>3</sup>Institute for Advanced Bioscience, Keio University, Tsuruoka, Yamagata, Japan. <sup>4</sup>Precursory Research for Embryonic Science and Technology (PRESTO), Japan Science and Technology Agency (JST), Tokyo, Japan.

<sup>5</sup>Department of Environmental Sciences, Graduate School of Nutritional and Environmental Sciences, University of Shizuoka, 52-1 Yada, Suruga-ku, Shizuoka, 422-8526, Japan.

<sup>6</sup>Department of Chemical Science and Engineering, Graduate School of Engineering, Kobe University, 1-1 Rokkodai-cho, Nada-ku, Kobe, Hyogo 657-8501, Japan.

\*Corresponding author. Email: akondo@kobe-u.ac.jp

DNA formation by dCas9-gRNA is sufficient to recruit PmCDA1 (Fig. 1B, 3) and that the mutation is induced at C on the noncomplementary strand, which is facilitated by nickase activity of nCas9(D10A). The combination of nCas9(H840A) and PmCDA1 reduced mutation frequency, suggesting that nicking on the same

strand of the deamination site may prevent the mutagenic effect, presumably by facilitating the excision of deaminated base. It also appeared that Target-AID, using nCas9(D10A) and PmCDA1, was much less toxic than the full CRISPR/Cas9 nuclease, which causes growth defects or cell death in yeast (Fig. 2B). The low efficiency of full

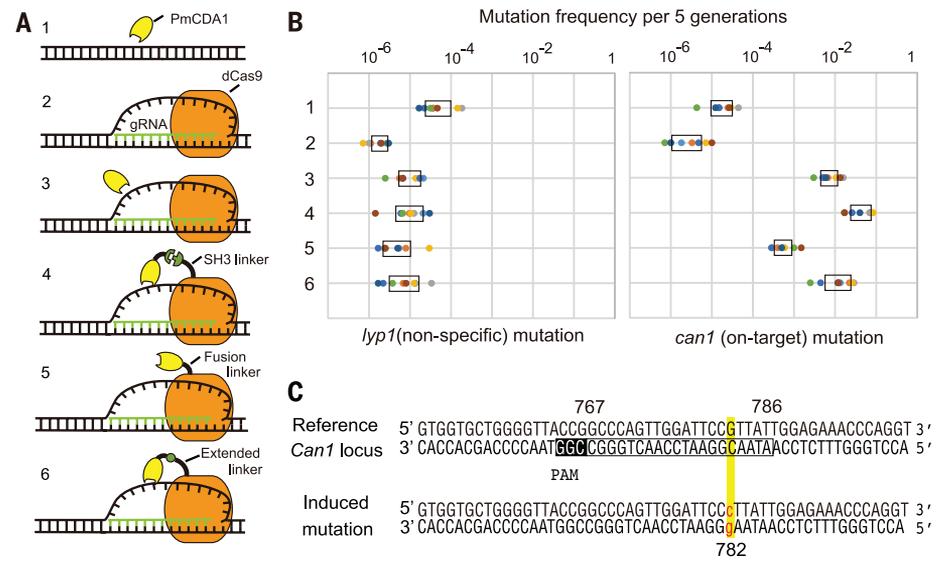
Cas9 nuclease in yeast may be attributed to its toxicity. This facilitates emergence of “escaper cells” that have lost the system (15, 16).

**Mutation spectrum induced by Target-AID**

To detect the presence of associated mutations, deep sequencing of the target-flanking region

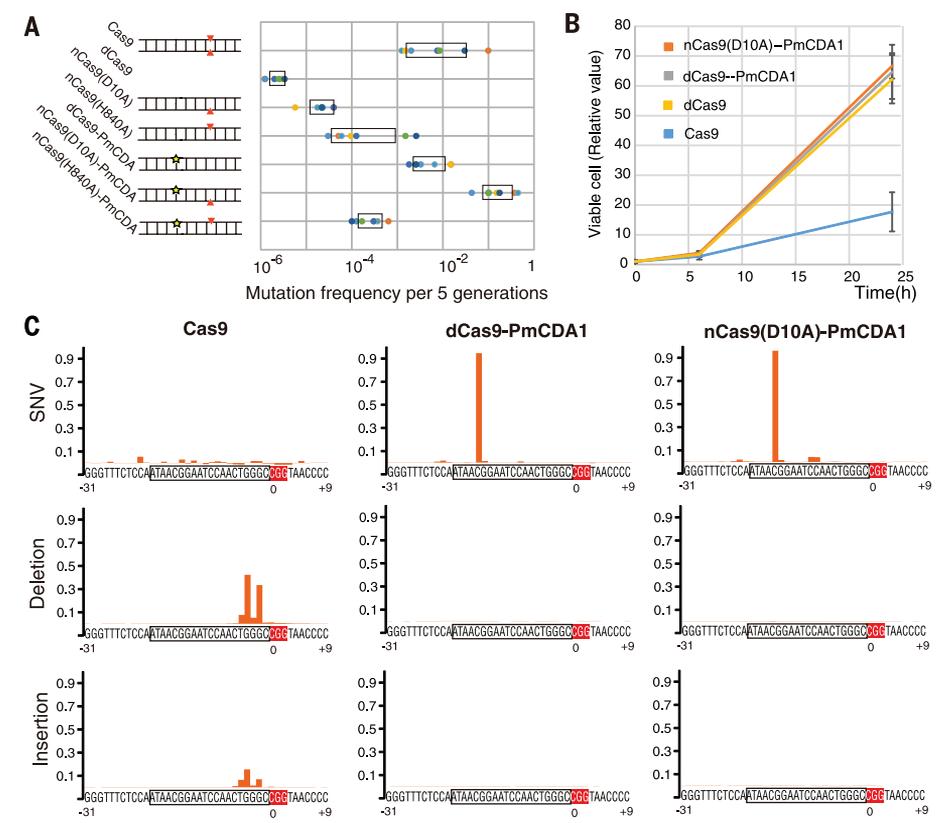
**Fig. 1. Complex formation of dCas9-gRNA and PmCDA1 induces targeted mutation.**

(A) Schematics of the complex forms of dCas9, gRNA, the target DNA, and PmCDA1 examined in this study. (1) PmCDA1 alone. (2) dCas9-gRNA alone. (3) PmCDA1 and dCas9-gRNA without direct attachment. (4) PmCDA1-SHL attached to dCas9-SH3-gRNA via SH3 and SHL interaction (SH3 linker). (5) PmCDA1 fused to dCas9 with short linker (17 amino acids; NLS plus 2x glycine-serine peptide linker). (6) PmCDA1 fused to dCas9 with long linker (100 amino acids). (B) Yeast cells expressing each protein complex numbered as in (A), with gRNA targeting a single site (786 to 767) in CAN1, were analyzed by a colony formation assay. S-aminoethyl-L-cysteine and canavanine select for the cells with the *lyp1* mutation (nonspecific mutation) and the *can1* mutation (on-target mutation), respectively. Mutation frequency per five generations is plotted. Eight biological replicates are shown in different colors. Box bars indicate 95% confidence interval for a sample mean by *t* test statistical analysis. (C) Partial nucleotide sequence of CAN1 with the target site (box) and the PAM sequences (inverted) are indicated with open reading frame nucleotide number. Representative sequence of the canavanine-resistant mutants induced by Target-AID is aligned. Mutation site and mutated nucleotide are highlighted by yellow and red, respectively. A mutation list can be found in table S1.

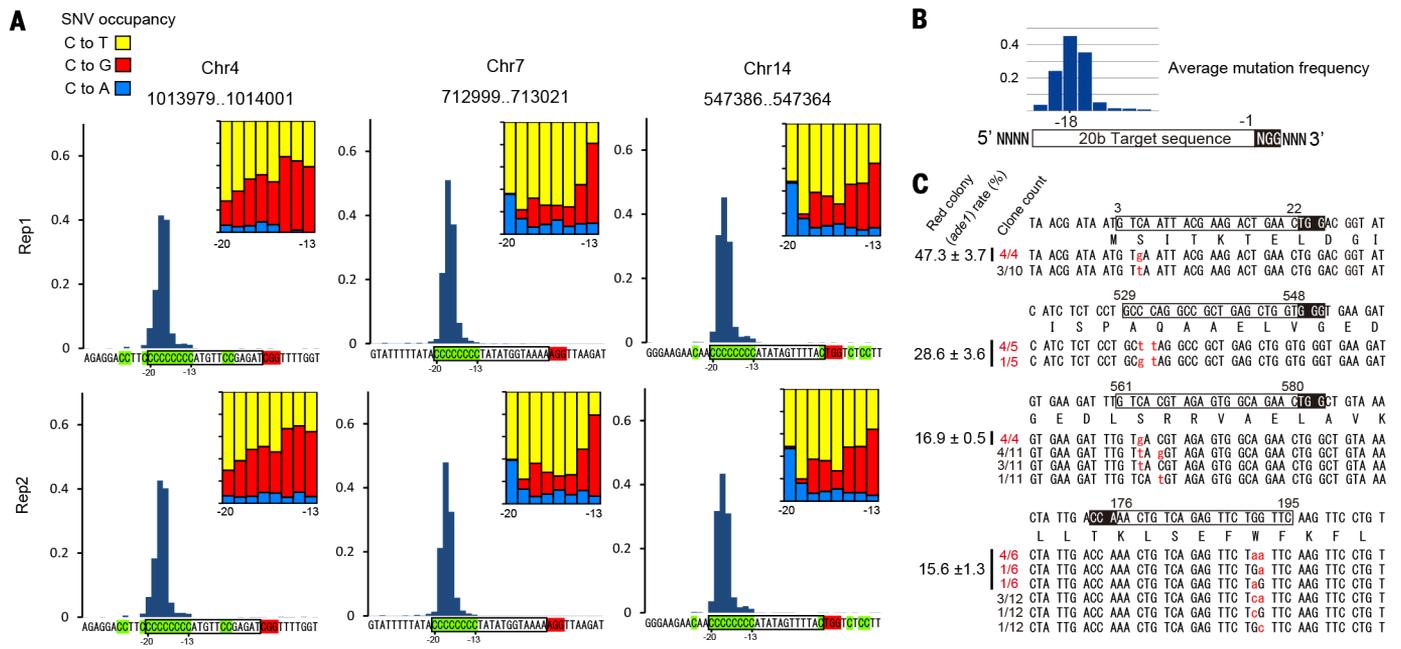


**Fig. 2. Effect of nickase activity on Target-AID.**

(A) Forms of fusion protein of Cas9 nuclease mutants [Cas9, dCas9, nCas9(D10A), or nCas9(H840A)] and PmCDA1 are indicated on the left. The schematic illustration in the middle depicts the process of DNA modification by the fusion proteins. Ladder indicates target DNA. Red arrowhead indicates DNA-strand break introduction. Yellow star indicates cytosine deamination introduction. Yeast cells expressing each protein, with gRNA targeting a single site (791 to 772) in CAN1, were analyzed for canavanine resistance, and the mutation frequencies are plotted. Seven biological replicates are shown in different colors. Box bars indicate the 95% confidence interval for a sample mean by *t* test statistical analysis. (B) Yeast cells expressing Cas9, dCas9, dCas9-PmCDA1 or nCas9(D10A)-PmCDA1, and gRNA targeting CAN1 were measured for cell viability. Viable cells were counted and plotted based on colony formation at each time point. Relative values (1 at 0 hour) of the mean of the biological triplicate are plotted. Error bars indicate SD. (C) Mutation spectra of canavanine-resistant cells produced by Cas9, dCas9-PmCDA1, or nCas9-PmCDA1, analyzed by deep sequencing at the targeted genomic site (32681 to 32700 at chromosome V or 786 to 767 in Can1). Frequencies of SNVs and deletions or insertions at each nucleotide position were plotted from the -31 to the 9-bp position relative to target PAM sequence. The nucleotide sequence is shown with the target sequence (box) and the PAM sequence (inverted red).



Downloaded from https://www.science.org on January 25, 2024



**Fig. 3. Mutable position and target design for nCas9(D10A)-PmCDA1.**

(A) Mutation spectra of the targeted poly-C-containing sequences. Three intergenic target sequences that contain poly-C at 5' half were edited by nCas9(D10A)-PmCDA1 and analyzed by deep sequencing. Frequencies of SNV at each position from -31 to 9-bp position relative to target PAM sequence were plotted. The upper right insets indicate the ratio of nucleotide changes at positions from -20 to -13 as C to T (yellow), C to G (red), and C to A (blue). Two biological replicates (Rep1 and Rep2) are shown. (B) Mutational position and frequency by Target-AID. Blue bars indicate average mutation frequency of A at each base position in the 20-base target sequence. (C) Four

target sites in *Ade1* gene were selected for editing by Target-AID based on the criteria that the cytosine mutations in the -19 to -16 position may introduce a stop codon, resulting in a red colony phenotype. Red colony frequency (mean  $\pm$  SD of biological triplicate) for each target site is indicated on the left. Sequences of identified mutations are aligned with the number of clones over the number of total sequenced clones for each colony color (red or none). Reference wild-type sequences with translated amino acid sequences are shown with the target site (box) and the PAM sequence (inverted). The bottom target sequence is complementary to the sequence shown.

was performed for canavanine-selected cells mutated at the *Can1* locus by Cas9, dCas9-PmCDA1, or nCas9-PmCDA1 (Fig. 2C). Cas9 induced deletions and insertions at the -1 to -5 positions. dCas9-PmCDA and nCas9-PmCDA induced a point mutation predominantly at the -16 position with a frequency of 94.7 and 96.1%, respectively. The second most frequent single-nucleotide variants (SNVs) were at the -22 position for dCas9-PmCDA and the -10 position for nCas9-PmCDA, with a frequency of 1.7 and 4.3%, respectively. Outside of the -22 to -10 position, no SNVs over 1% were detected for dCas9-PmCDA1, and one SNV at the +24 position for nCas9-PmCDA1 was found (1.2%). Insertions and deletions (indels) were rarely observed for both dCas9-PmCDA and nCas9-PmCDA with the highest frequency of 0.18 and 0.24%, respectively, at the -16 position.

The high efficiency of Target-AID using nCas9(D10A) and PmCDA1 allowed us to obtain mutants without canavanine-selection and to perform nonbiased sequencing analyses, which shows that the mutations were focused at cytosine bases located 15 to 19 bases upstream of the PAM sequence on the noncomplementary strand to gRNA, while the length of the gRNA target sequence (i.e., the 5' end position of gRNA) did not affect mutation location (fig. S2). C to G and C to T mutations of the noncomplementary strand were observed with nearly equal frequency

(9 of 18 and 8 of 18, respectively), whereas C to A was rare (1 of 18).

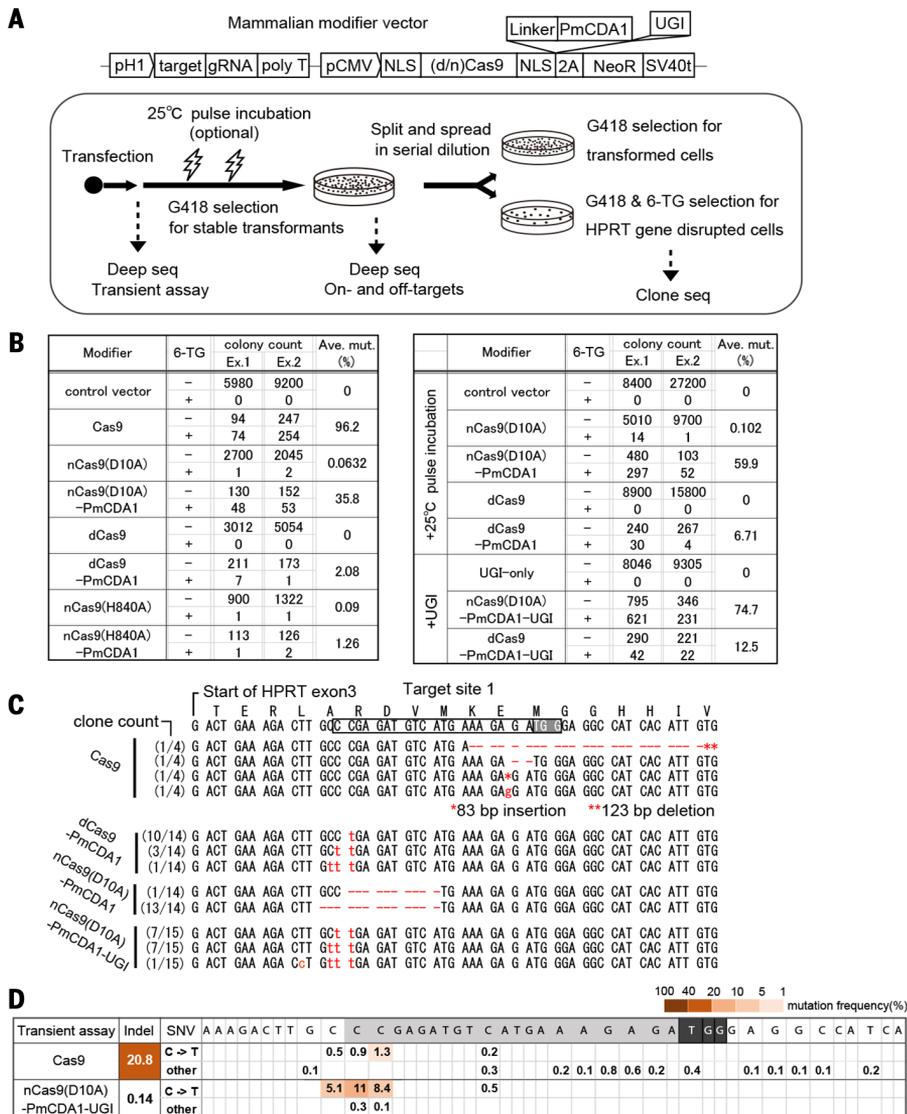
To obtain a comprehensive mutational spectrum for nCas9-PmCDA1, three intergenic target sites that were C-rich were selected (Fig. 3A). At the three targets, a peak of SNV frequency was consistently observed at the -18 position (41 to 51%), with lower frequencies at the -17 and -19 positions (18 to 40%). The SNV frequencies decreased at the -16 position (4.5 to 6.5%) and -20 position (1.3 to 6.3%) and further decreased in the surrounding area. Frequencies of more than 1% were rarely observed outside the -13 to -21 positions. This suggests that the mutational efficiency is highly dependent on the relative position within the target sequence. It also implies that processive deamination of poly-C by PmCDA1 is limited, if it occurs at all. On the other hand, changed nucleotides showed positional bias within the poly-C region (Fig. 3A, insets). C to T mutations were more common at the 5' side of the poly-C region, whereas the C to G mutation ratio increased on the 3' side. In *Saccharomyces cerevisiae*, the translesion DNA polymerase eta prefers to insert C opposite an abasic lesion, resulting in a G insertion at an abasic site (17), whereas a T insertion at an abasic site is common in most other organisms (18). The 3' side of the poly-C region is more likely to be subjected to error-prone repair by the translesion polymer-

ase eta, whereas mutations on the 5' side are left as the deaminated base (i.e., uracil), which is then recognized as T in DNA synthesis.

To assess the general effectiveness of Target-AID, four sites were systematically selected from the *Ade1* gene [918 base pairs (bp)] to introduce a stop codon, following the criteria that C to G or C to T mutations should be introduced at -16 to -19 positions upstream of the PAM sequence (Fig. 3, B and C). The four targets resulted in an approximately 47, 29, 17, and 16% rate of *ade1* gene disruption (red colony phenotype), respectively (Fig. 3C). Sequencing analysis of both red and white colonies showed that the genetic mutation frequency was even greater because not all mutations resulted in gene disruption. All sequenced mutations were localized at the -16 to -18 position (Fig. 3C). There seems to be no strong context preference as reported previously for PmCDA1 (19).

### Multiplex and biallelic editing by Target-AID

Multiplex editing was performed for the *ADE1* and *CAN1* genes by simultaneously expressing *ADE1* and *CAN1* targeting gRNAs. The independent mutation frequencies for *ADE1* and *CAN1* were ~54 and 51%, respectively, and double-phenotypic mutants occurred ~31% of the time (fig. S3A). Sequencing analysis revealed that all



**Fig. 4. Target-AID in mammalian cells.** (A) Schematics of the mammalian experiment. The mammalian modifier vector expressing gRNA and the modifier protein is depicted on the top. Linker plus PmCDA1 is inserted before 2A peptide. UGI is also inserted between PmCDA1 and 2A peptide. Experimental flow is depicted at the bottom. CHO cells were transfected using each modifier vector and harvested after 3 days for transient assay or transferred to G418-containing media for selection of the transformed cells. Pulse incubations at 25°C are optionally introduced. After G418 selection, cells were analyzed by deep sequencing. For the HPRT gene target, cells were further tested for 6-TG resistance. The resistant cells were counted and then analyzed by Sanger sequencing. (B) 6-TG-resistant (HPRT null) frequencies induced by each modifier plasmid are shown. Cells were spread onto media with (+) or without (-) 6-TG in serial dilution, and resistant colonies were counted at optimal colony density for each sample. Experiment was repeated (Ex.1 and Ex.2) and combined to give average mutation frequency [Ave. mut (%)]. (C) Mutation alignment of the targeted HPRT gene locus. About 0.4-kb fragment around the target site of 6-TG resistant cells were PCR-amplified and subcloned for sequencing. Frequencies of the same mutant sequence were indicated as clone count over total reads of the same vector construct. Mutations are shown in red. Large deletion or insertion is indicated by an asterisk. (D) Mutation spectra obtained by transient transfection. Cells were transfected with a vector expressing Cas9 or nCas9(D10A)-PmCDA1-UGI and gRNA targeting HPRT. Whole-cell culture was subjected to deep sequencing without any selection or enrichment. Total indel frequency (Indel) within the indicated region (33 nucleotides surrounding the target site) and SNV frequencies above 0.1% at each nucleotide position are shown and highlighted as indicated at the top right. Target sequence is shadowed and PAM sequence is inverted.

clones (10 of 10) contained mutations at both sites, whereas phenotypic alterations were linked to specific types of mutations (fig. S3B).

Biallelic editing was performed by using yeast diploid strain YPH501 (fig. S4). Occurrence of biallelic disruption of *Adel* and *Can1* were ana-

lyzed by their recessive phenotype of red colony and canavanine resistance, respectively. Biallelic mutants were obtained more frequently than expected if independent allelic mutagenesis was assumed (fig. S4A), whereas sequencing analysis verified heteroallelic mutant from phenotypically wild-type cells (fig. S4B). These imply that microhomology-mediated repair or gene conversion is induced during deamination-mediated mutagenesis, which results in a higher rate of biallelic mutagenesis (20).

### gRNA mismatch tolerance of Target-AID

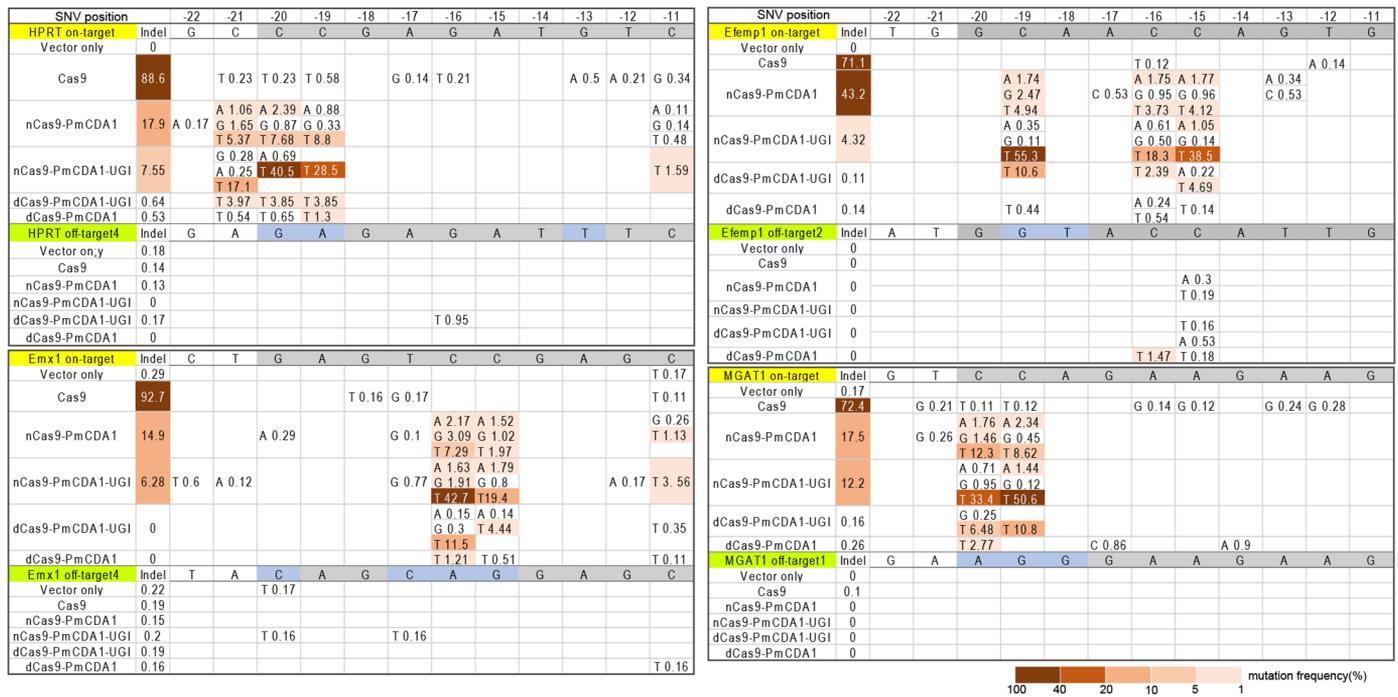
Off-targeting effect was assessed by introducing single mismatches at each base position of the gRNA sequence (fig. S5). Mismatches proximal to PAM sequence (3, 6, or 9 bases away) were not tolerated, whereas mismatches at 12 or 15 base position showed a reduced but detectable mutation rate (27- or 35-fold reduction, respectively). Mismatch at the 18 base position, where deamination is expected to occur on the opposite strand, was not tolerated. Mismatches at more than 20 base positions showed no additional effect. These results are consistent with Cas9 specificity and thus will be improved by Cas9 protein engineering (21).

### Whole-genome sequencing of Target-AID-treated yeast

To further assess nonspecific mutator effects of the deaminase in this system, whole genome sequencing was performed for each of three clones expressing Cas9, nCas9, Cas9, PmCDA1, dCas9-PmCDA1, and nCas9-PmCDA1 (tables S2 and S3). Deaminase-expressing strains (PmCDA1, dCas9-PmCDA1, and nCas9-PmCDA1) were found to contain seven unique SNVs and two structural variants in a total of nine strains, whereas one SNV and three structural variants were found in a total of nine no-deaminase strains (dCas9, nCas9, and Cas9). A slight increase in SNVs in the deaminase-expressing strains may be attributed to the mutator effect of the deaminase. Overexpression of deaminases are mutagenic, and PmCDA1 expression in a yeast strain lacking uracil DNA glycosylase (*ung1*-) has been shown to induce genome-wide mutation (20). On the other hand, vertebrate cells that possess AID family proteins do not suffer from AID-related mutagenicity at wild-type expression levels. Compared with DNA double-strand breaks, DNA deamination is a more frequent spontaneous DNA lesion and much easier to repair precisely. Only intensive deamination or a compromised repair system may cause mutation. Cas9 expression did not show obvious off-target or nonspecific mutagenic effects in yeast, possibly because random DNA double-strand breaks may lead to cell death in yeast and thus be underrepresented.

### DNA nicking plus deamination induces deletion in mammalian cells

To investigate the feasibility of Target-AID in mammalian cell, Chinese hamster ovary (CHO) cell was transfected using a vector expressing Cas9, nCas9 (D10A), nCas9(D10A)-PmCDA, dCas9-PmCDA,



**Fig. 5. Deep-sequencing analysis of target and off-target mutations induced by Cas9 and a series of Target-AID in CHO cells.** Using the indicated modifier vector (left column) expressing gRNA targeting HPRT1, EMX1, Efempl1, or MGAT1, transfected cell culture was analyzed 7 days after transfection. The on-target site (yellow) and the top six off-target sites determined by CCTop (<http://crispr.cos.uni-heidelberg.de/index.html>) as listed in tables S5 to S8 were analyzed by deep sequencing. Full sets of off-target analysis are shown in figs. S5 to S8, whereas the most highly mutated off-target sites (green) for each target site are shown here. Target sequence and

or nCas9(H840A)-PmCDA, with a gRNA targeting the hypoxanthine-guanine phosphoribosyltransferase (HPRT) gene. As HPRT converts a purine analog 6-thioguanine (6TG) into a toxic derivative, homozygous disruption of the HPRT gene confers 6TG resistance to the cell. After transfection, equal numbers of cells were spread and grown to form colonies in the presence or absence of 6TG to calculate mutation frequency (Fig. 4A). nCas9(D10A)-PmCDA1 induced a higher number of resistant colonies than dCas9-PmCDA1 (Fig. 4B). Sequencing analysis showed base substitutions for clones from dCas9-PmCDA1, as expected (Fig. 4C). However, nCas9(D10A)-PmCDA1 induced short deletions with a central focus on the 5' distal side of the target sequence, where deamination was expected to occur. This is likely due to a combination of nick and deamination that cannot be repaired properly in mammalian cell.

### Suppression of deletion by uracil DNA glycosylase inhibitor

To assess and bypass the repair pathway, we introduced uracil DNA glycosylase inhibitor protein (UGI) (22), which blocks removal of uracil in DNA and the subsequent repair pathway. UGI was expressed as a fusion protein with d/nCas9-PmCDA1 to demonstrate improved mutation frequency (Fig. 4B), as well as a substantial

change in the mutation spectra, showing a majority of C to T substitutions instead of deletions (Fig. 4C, nCas9(D10A)-PmCDA1-UGI). Pulse incubations at 25°C after transfection improved the mutation frequency to 59.9 and 6.71% for nCas9-PmCDA1 and dCas9-PmCDA1, respectively (PmCDA1 is originally from sea lamprey, which is presumably adapted to low temperatures). Transient transfection was tested by harvesting the cell culture 3 days after transfection without any marker selection. Deep sequencing showed that nCas9-PmCDA1-UGI induced 8.4, 11, and 5.1% of C to T base substitutions at -19, -20 and -21 positions, respectively, whereas Cas9 induced 20.8% indels in the same condition (Fig. 4D).

### Off-target assessment in mammalian cells

General effectiveness and potential off-target effect were assessed by using three target loci—MGAT1 (23), Efempl1 (24), and EMX1 (25)—that were previously tested as CRISPR/Cas9 targets in CHO cells or in human cells, in addition to HPRT locus. After G418 selection for transformants, cells were harvested and subjected to deep sequencing. Potential off-target genomic regions were selected using CCTop (<http://crispr.cos.uni-heidelberg.de/index.html>) (26) (tables S6 to S9), and the top six off-target sites were analyzed (Fig. 5 and figs. S6 to S9). Compared with nCas9-

mismatched bases are highlighted in gray and light blue, respectively, with nucleotide position (-22 to -11) relative to PAM sequence on the top. There is no mismatch in the position from -12 to -1 of the all off-target sites examined. Total indel frequency within the 43 bases (20 bases of target region plus upstream 10 bases and downstream 13 bases, including PAM) is shown in the second column from the left. SNV frequencies above 0.1% at each nucleotide position (-22 to -11) are separately shown with each mutant nucleotides. Mutation frequencies are highlighted as indicated at the bottom right.

mutation frequency(%)

100 40 20 10 5 1

PmCDA1, which induced indels as well as SNVs, nCas9-PmCDA1-UGI effectively induced C to T mutations within the -15 to -21 region of the four target sites with a lower frequency of indels. dCas9-PmCDA1-UGI showed fewer indels and a moderate frequency of C to T mutations. There were possible off-target mutations detected at lower than 1.5% for Target-AID throughout the analyzed off-target sites (Fig. 5 and figs. S5 to S8).

### Conclusion

We demonstrated targeted nucleotide substitution without the use of template DNA by expanding the potential of the CRISPR/Cas9 system. The Target-AID performed a very narrow range of targeted nucleotide substitution within a few bases. Its reduced cytotoxicity will be beneficial, especially for use in bacteria and other cell types that are sensitive to artificial nucleases (16). A combination of nickase Cas9(D10A) and PmCDA1 was highly efficient in yeast, whereas it induced deletions as well as point mutations in mammalian cells, suggesting that the mutation spectrum is affected by the difference of the repair pathway. The rat apolipoprotein B mRNA editing enzyme, catalytic polypeptide (rAPOBE1), another cytidine deaminase, has also been reported to serve as a programmable base editor

(BE) by fusing it to the N terminus of dCas9 (27). Consistent with our study, the use of UGI fusion and D10A nickase increased efficiency in human cells. There are also different features as well. Using Target-AID, three to five bases surrounding the -18 position upstream of the PAM sequence are substantially subjected to mutation, whereas five bases surrounding the -15 position upstream of PAM (or position 4 to 8 from the end distal to the PAM of a 20-base protospacer) are the major target for BE. This may be attributed to the difference of the enzymatic characteristics (PmCDA1 versus rAPOBEC1) or the different attachments to Cas9 (C terminus versus N terminus). Possible RNA-editing activity is also to be assessed, because APOBEC proteins have RNA-editing activity physiologically (2). Additionally, PmCDA1 exhibited temperature sensitivity that can be exploited to control its activity. Although direct comparison will be needed, these two systems may complement each other to extend the repertoire of possible editing sites. Using other CRISPR-related systems or other modifier enzymes, such as adenosine deaminase, will also broaden the editing capacity and further enrich the genome editing toolbox.

## Materials and methods

### DNA manipulation

Plasmids available from Addgene are listed in table S4. The human-optimized *Streptococcus pyogenes* Cas9 gene-containing plasmid p415-Gal-Cas9-CYC1t and the chimeric gRNA-containing plasmid p426-SNR52p-gRNA.CAN1.Y-SUP4t (13) were kindly provided by Church laboratory and obtained from Addgene (Cambridge, MA, USA). Nuclease-deficient Cas9 D10A and H840A mutations were introduced using a PCR-based method. The target gRNA sequences were replaced using a PCR method. A glycine-serine peptide linker (GGGG), a SH3 domain, and a 3xFlag tag (DYKDDHGDYKDDHIDYKDDDDK) were added at the C terminus after the SV40 nuclear localization signal (NLS) (PKKKRKV) in Cas9. The DNA coding sequences for PmCDA1, the SH3 domain, and the 3xFlag tag were synthesized (Eurofin Genomics, Tokyo, JPN). The DNA fragments were assembled and modified using standard methods. The nucleotide sequence for the entire coding region is in fig. S10.

For CHO transfection plasmid, human codon-optimized PmCDA1 was synthesized on pcDNA3.1 backbone (Genscript, NJ, USA). Human-optimized Cas9 coding sequence was inserted downstream of CMV promoter, followed by SV40 NLS, dead SH3 domain, 3xFlag, PmCDA1, and UGI by PCR sewing and Gibson assembly method. Neomycin-resistant gene was linked via 2A peptide (EGRGSL-LTCGDVEENPGP) at the end of the coding region. The nucleotide sequence for the entire coding sequence is in fig. S11.

### Targeted mutagenesis

*Saccharomyces cerevisiae* BY4741 (*MATa his3Δ0 leu2Δ0 met15Δ0 ura3Δ0*) (Open Biosystems, Huntsville, AL, USA) was used in this study. Yeast culture and transformation were performed

according to standard methods. For the mutational analysis, the yeast cells were grown at 30°C in synthetic complete media with an indicated carbon source and an appropriate auxotrophic compound complemented by the plasmids. The cells were first inoculated in a 2% glucose medium and grown to saturation, diluted 16-fold into a 2% raffinose medium, and grown to saturation again. The cells were then diluted 16-fold or 32-fold into 2% raffinose medium plus 0.2 or 0.02% galactose medium for induction. To calculate the mutation rate per generation, cell culture was diluted 32-fold every 24 hours. After the indicated time, the cell culture was sampled and serially diluted 10-fold in water and spotted or spread onto a synthetic medium plate supplemented with the appropriate dropout mix and the appropriate drug. For CAN1 mutational analysis, arginine was omitted and 60 mg/l canavanine was added. For LYP1 mutational analysis, lysine was omitted and 100 mg/l of S-aminoethyl-L-cysteine was added. The mutation frequency of ADE1 was counted by colony color because the loss of function of *ade1* results in the accumulation of red pigment when subjected to an adenine-limited condition. The plates were incubated at 28°C for 2 to 3 days and the colonies were counted. Sample size was determined based on countable colony number in a single spot (up to 40) or dish (up to 400). The *t* test statistical analysis was done using Excel software (Microsoft, WA, USA). The plate images were acquired using an Image Quant LAS 4000 (GE Healthcare Japan, Tokyo, Japan) or a Cybershot DSC-WX100 digital camera (Sony, Tokyo, Japan) and were processed using Adobe CS2 software. For the sequencing analysis, the average-sized colonies were randomly picked and directly amplified by PCR and were analyzed by Sanger sequencing using a 3130xL Genetic Analyzer (Applied Biosystems, CA, USA) according to the manufacturer's instructions.

### Viability test

Yeast cells were precultured in synthetic complete media with 2% raffinose minus leucine and uracil for overnight, then diluted by 32-fold and supplemented with 0.02% galactose to induce protein expression. Cells were harvested at each time point. Ten-fold serial dilution of harvested culture was spotted onto synthetic complete media minus leucine and uracil and observed for colony formation for 2 days.

### Mutation spectra analysis

Yeast strains harboring galactose-inducible modifier gene plasmid (Cas9, nCas9, dCas9, dCas9-CDAI, nCas9-CDAI, or CDAI) containing LEU2 marker and gRNA-expressing plasmid containing Ura3 marker were precultured in 5 ml SC-Leu-Ura+Ade media containing 2% galactose overnight at 30°C with rotation. Each cell sample was washed once with SC-Leu-Ura+Ade media containing 2% raffinose, resuspended in the same raffinose media at 0.1 OD600 nm, and incubated overnight at 30°C with rotation. For the modifier induction, cells were resuspended into SC-Leu-

Ura+Ade media containing 2% galactose at 0.1 OD600 nm and incubated overnight at 30°C with rotation. Each sample was then spread on SC-Leu-Ura-Arg+Ade and SC-Leu-Ura-Arg+Ade containing 60 μg/ml canavanine plates. From each plate, colonies were scraped and pooled with sterile water and genomic DNA was extracted using YeaStar Genomic DNA Kit (Zymo Research). Target genomic locus was amplified by PCR with primers containing common sequence ends and reamplified by common Illumina adapter primers containing index tags for sample multiplexing. PCR bands were size-selected using E-Gel SizeSelect 2% Agarose Gel (Thermo Fisher Scientific) and quantified by KAPA Library Quantification Kits for Illumina (KAPA Biosystems). Sequencing libraries were multiplexed at equal molar ratio and analyzed by Illumina MiSeq Reagent Kit V2 (250×250 bp paired-end sequencing). Sequence reads were aligned to the primer sequences and the target genomic regions using BLAST+ program (short-read option; E-value cutoff of 1e-3) to demultiplex the read samples and to identify mutations and indels. Fractions of mutations, base insertions and deletions were separately calculated for every base position of target genomic region and subtracted by those of the PmCDA1 dataset derived from corresponding screening condition to normalize sequencing errors.

### Whole-genome sequencing

Yeast strains harboring each expression construct (dCas9, nCas9, Cas9, PmCDA1, dCas9-PmCDA1, and nCas9-PmCDA1) with gRNA targeting *Can1* were preincubated in 2% raffinose medium overnight, diluted by 32-fold into 0.02% galactose medium for induction and grown overnight. Cells were spread onto canavanine-containing plate medium to isolate single colonies. Each three independent colonies were inoculated into YPDA medium containing 300 μg/ml canavanine. Overnight culture was collected and genomic DNA was extracted using Wizard Genomic DNA purification Kit (Promega, WI, USA) following manufacturer's instructions. Genomic DNA was then fragmented by sonication using Bioruptor UCD-200 TS Sonication System (Diagenote, NJ, USA) to obtain fragments with size distribution centering at 500 to 900 bp. Genomic DNA library was prepared by using NEBNext Ultra DNA Library Prep Kit for Illumina (New England Biolabs, MA, USA) and labeled by Dual Index Primers Set 1. Size selection of the library was done using Agencourt AMPure XP (Beckman Coulter, CA, USA) to obtain tagged fragments with length ranging from 600 to 800 bp. Size distribution was evaluated by the Agilent 2100 Bioanalyzer system (Agilent Technologies, CA, USA). DNA was quantified using Qubit HS dsDNA HS Assay Kit and fluorometer (Thermo Fisher Scientific, MA, USA). Sequencing was performed using MiSeq sequencing system (Illumina, CA, USA) and MiSeq Reagent Kit v3 to obtain 2 × 300 bp read length, expecting approximately 30-fold coverage for the genome size. The sequence reads were paired and overlapping reads within a read pair were merged then trimmed based on a quality limit of 0.001

with a maximum of 2 ambiguities. Reads were mapped to *Saccharomyces cerevisiae* S288C reference genome by the following setting (Masking mode = no masking, Mismatch cost = 2, Insertion cost = 3, Deletion cost = 3, Length fraction = 0.5, Similarity fraction = 0.9, Global alignment = No, Auto-detect paired distances = Yes, Nonspecific match handling = ignore). Local realignment was done with default settings (Realign unaligned ends = Yes, Muti-pass realignment = 2). InDels and Structural Variants were detected with the following settings (*P*-value threshold = 0.001, Maximum number of mismatches = 3, Filter variants = Yes, Maximum number of reads = 10). Local realignment was redone using the Variant track of InDels and Structural Variants and Force realignment to guidance-variant track. The variant calling was performed with the following settings (Ignore positions with coverage = 1000, Ploidy = 1, Required variant probability = 90, Ignore broken pairs = No, Ignore Nonspecific matches = Region length 50, Minimum coverage = 10, Minimum count = 2, Minimum frequency = 67%, Base quality filter = No, Read detection filter = Yes, Relative read direction filter = 1%, Significance = 1%, Read position filter = 1%, Remove pyro-error variants = No). Compare variant within group was done with threshold of 1% using the Variant Detection files of all three clones of 6 different modifiers (dCas9, nCas9, Cas9, PmCDA1, dCas9-PmCDA1, and nCas9-PmCDA1). Mutations appeared in both of two groups (d/n/wtCas9-only and PmCDA1-only) were assigned as common mutations.

### CHO cell experiments

CHO-K1 adherent cells (ECACC 85051005) were used in this study. Cells were cultured in Ham's F12 medium (Life Technologies, Carlsbad, CA, USA) supplemented with 10% fetal bovine serum (Biosera, Nuaille, France) and 100 µg/mL Penicilline-Streptomycin (Life Technologies, Carlsbad, CA, USA) at 37°C with a humidified 5% CO<sub>2</sub> atmosphere. For the transfection, cells were plated at 0.5×10<sup>5</sup> cells per well of 24-well plates and cultured for 1 day. The cells were transfected with 1.5 µg plasmid and 2 µL Lipofectamine 2000 (Life Technologies, Carlsbad, CA, USA) per well according to manufacturer's instructions. Five hours after transfection, the medium was changed to Ham's F12 medium containing 0.125 mg/mL G418 (InvivoGen, San Diego, CA, USA), and the cells were incubated for 7 days. For pulse incubation at 25°C, the cells were first transfected and incubated at 37°C for 1 day. Cells were then transferred to 25°C and incubated for 1 day, followed by 37°C incubation for 2 days. This process was repeated twice. To calculate the mutation frequency of the HPRT gene, the cells were released from the well using trypsin-EDTA (Life Technologies, Carlsbad, CA, USA) and serially diluted and spread to obtain countable colony density onto Ham's F12 medium containing G418 or G418 plus 5 g/mL 6-TG (Tokyo Chemical Industry, Tokyo, Japan). After 7 days, the number of resistant colonies was counted. The mutation frequency was calculated as the rate of 6TG res-

sistant colonies over G418 resistant colonies. For sequencing analysis, cells were trypsinized and pelleted by centrifugation. Genomic DNA was extracted from the pellets using the NucleoSpin Tissue XS kit (Macherey-Nagel, Düren, Germany) according to the manufacturer's instructions. PCR fragments including HPRT target site was amplified from the genomic DNA using a forward primer GGCTACA-TAGAGGGATCCTGTGTCA and a reverse primer ACAGTAGCTCTTCAGTCTGATAAAA. PCR products were cloned into *E.coli* vector and analyzed by Sanger sequencing.

### Deep sequencing of target and off-target region of CHO cells

Target region-containing fragment (~1.5 kb) was first PCR-amplified using 1st primer pairs from the extracted genomic DNA. The second nested PCR was performed to obtain adapter-added amplicon (~0.3 kb) by using the first PCR products as template and 2nd primer pairs containing adapter sequences. The primers used in this study are listed in table S5. The amplicon was labeled using NEBNext Multiplex Oligos for Illumina (Index Primers Set 1 and Dual Index Primers Set 1) (New England Biolabs, MA, USA). Deep sequencing was performed using MiSeq sequencing system (Illumina, CA, USA) to obtain paired 300 bp read length and more than 100,000 reads per sample on average.

Data analysis was done using CLC Genomics Workbench 7.0 (CLC bio, Aarhus, Denmark). The sequence reads were paired and trimmed based on a quality limit of 0.05 with a maximum of two ambiguities and then overlapping reads within a read pair were merged. Reads were mapped to each reference sequence obtained from CHoGenome.org ([www.chogenome.org/](http://www.chogenome.org/)) by the following setting (Masking mode = no masking, Mismatch cost = 2, Insertion cost = 3, Deletion cost = 3, Length fraction = 0.2, Similarity fraction = 0.5, Global alignment = No, Auto-detect paired distances = Yes, Nonspecific match handling = ignore). The samples containing low-quality reads (Efem1 on-target, Efem1 off-target 5, MGAT1 off-target 1 and EMX1 off-target 4) were reanalyzed by the following setting instead. The sequence reads were paired and overlapping reads within a read pair were merged then trimmed based on a quality limit of 0.01 with a maximum of 2 ambiguities. Reads were mapped with the following settings (Masking mode = no masking, Mismatch cost = 2, Insertion cost = 3, Deletion cost = 3, Length fraction = 0.6, Similarity fraction = 0.9, Global alignment = No, Auto-detect paired distances = Yes, Nonspecific match handling = ignore). The variant calling was performed with the following settings (Ignore positions with coverage = 1,000,000, Ignore broken pairs = No, Ignore Nonspecific matches = Region length 50, Minimum coverage = 10, Minimum count = 2, Minimum frequency = 0.1%, Base quality filter = No, Read detection filter = Yes, Relative read direction filter = 1%, Significance = 1%, Read position filter = 1%, Remove pyro-error variants = No). Output file was rearranged using Excel (Microsoft, WA, USA).

### REFERENCES AND NOTES

- M. Muramatsu *et al.*, Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. *Cell* **102**, 553–563 (2000). doi: [10.1016/S0092-8674\(00\)00078-7](https://doi.org/10.1016/S0092-8674(00)00078-7); pmid: [11007474](https://pubmed.ncbi.nlm.nih.gov/11007474/)
- B. A. Knisbacher, D. Gerber, E. Y. Levanon, DNA Editing by APOBECs: A Genomic Preserver and Transformer. *Trends Genet.* **32**, 16–28 (2016). doi: [10.1016/j.tig.2015.10.005](https://doi.org/10.1016/j.tig.2015.10.005); pmid: [26608778](https://pubmed.ncbi.nlm.nih.gov/26608778/)
- F. W. Alt, Y. Zhang, F. L. Meng, C. Guo, B. Schwer, Mechanisms of programmed DNA lesions and genomic instability in the immune system. *Cell* **152**, 417–429 (2013). doi: [10.1016/j.cell.2013.01.007](https://doi.org/10.1016/j.cell.2013.01.007); pmid: [23374339](https://pubmed.ncbi.nlm.nih.gov/23374339/)
- J. Chaudhuri *et al.*, Transcription-targeted DNA deamination by the AID antibody diversification enzyme. *Nature* **422**, 726–730 (2003). doi: [10.1038/nature01574](https://doi.org/10.1038/nature01574); pmid: [12692563](https://pubmed.ncbi.nlm.nih.gov/12692563/)
- R. Bransteitter, P. Pham, M. D. Scharff, M. F. Goodman, Activation-induced cytidine deaminase deaminates deoxycytidine on single-stranded DNA but requires the action of RNase. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 4102–4107 (2003). doi: [10.1073/pnas.0730835100](https://doi.org/10.1073/pnas.0730835100); pmid: [12651944](https://pubmed.ncbi.nlm.nih.gov/12651944/)
- C. Rada *et al.*, Immunoglobulin isotype switching is inhibited and somatic hypermutation perturbed in UNG-deficient mice. *Curr. Biol.* **12**, 1748–1755 (2002). doi: [10.1016/S0960-9822\(02\)01215-0](https://doi.org/10.1016/S0960-9822(02)01215-0); pmid: [12401169](https://pubmed.ncbi.nlm.nih.gov/12401169/)
- M. Liu, D. G. Schatz, Balancing AID and DNA repair during somatic hypermutation. *Trends Immunol.* **30**, 173–181 (2009). doi: [10.1016/j.it.2009.01.007](https://doi.org/10.1016/j.it.2009.01.007); pmid: [19303358](https://pubmed.ncbi.nlm.nih.gov/19303358/)
- J. van der Oost, E. R. Westra, R. N. Jackson, B. Wiedenheft, Unravelling the structural and mechanistic basis of CRISPR-Cas systems. *Nat. Rev. Microbiol.* **12**, 479–492 (2014). doi: [10.1038/nrmicro3279](https://doi.org/10.1038/nrmicro3279); pmid: [24909109](https://pubmed.ncbi.nlm.nih.gov/24909109/)
- P. D. Hsu, E. S. Lander, F. Zhang, Development and applications of CRISPR-Cas9 for genome engineering. *Cell* **157**, 1262–1278 (2014). doi: [10.1016/j.cell.2014.05.010](https://doi.org/10.1016/j.cell.2014.05.010); pmid: [24906146](https://pubmed.ncbi.nlm.nih.gov/24906146/)
- P. Mali, K. M. Esvelt, G. M. Church, Cas9 as a versatile tool for engineering biology. *Nat. Methods* **10**, 957–963 (2013). doi: [10.1038/nmeth.2649](https://doi.org/10.1038/nmeth.2649); pmid: [24076990](https://pubmed.ncbi.nlm.nih.gov/24076990/)
- M. Jinek *et al.*, A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816–821 (2012). doi: [10.1126/science.1225829](https://doi.org/10.1126/science.1225829); pmid: [22745249](https://pubmed.ncbi.nlm.nih.gov/22745249/)
- J. E. Duerber *et al.*, Synthetic protein scaffolds provide modular control over metabolic flux. *Nat. Biotechnol.* **27**, 753–759 (2009). doi: [10.1038/nbt.1557](https://doi.org/10.1038/nbt.1557); pmid: [19648908](https://pubmed.ncbi.nlm.nih.gov/19648908/)
- J. E. DiCarlo *et al.*, Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas systems. *Nucleic Acids Res.* **41**, 4336–4343 (2013). doi: [10.1093/nar/gkt135](https://doi.org/10.1093/nar/gkt135); pmid: [23460208](https://pubmed.ncbi.nlm.nih.gov/23460208/)
- S. Q. Tsai *et al.*, Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing. *Nat. Biotechnol.* **32**, 569–576 (2014). doi: [10.1038/nbt.2908](https://doi.org/10.1038/nbt.2908); pmid: [24770325](https://pubmed.ncbi.nlm.nih.gov/24770325/)
- R. B. Vercoe *et al.*, Cytotoxic chromosomal targeting by CRISPR/Cas systems can reshape bacterial genomes and expel or remodel pathogenicity islands. *PLOS Genet.* **9**, e1003454 (2013). doi: [10.1371/journal.pgen.1003454](https://doi.org/10.1371/journal.pgen.1003454); pmid: [23637624](https://pubmed.ncbi.nlm.nih.gov/23637624/)
- W. Jiang, D. Bikard, D. Cox, F. Zhang, L. A. Marraffini, RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat. Biotechnol.* **31**, 233–239 (2013). doi: [10.1038/nbt.2508](https://doi.org/10.1038/nbt.2508); pmid: [23360965](https://pubmed.ncbi.nlm.nih.gov/23360965/)
- B. Zhao, Z. Xie, H. Shen, Z. Wang, Role of DNA polymerase eta in the bypass of abasic sites in yeast cells. *Nucleic Acids Res.* **32**, 3984–3994 (2004). doi: [10.1093/nar/gkh710](https://doi.org/10.1093/nar/gkh710); pmid: [15284331](https://pubmed.ncbi.nlm.nih.gov/15284331/)
- A. J. Rattray, J. N. Strathern, Error-prone DNA polymerases: When making a mistake is the only way to get ahead. *Annu. Rev. Genet.* **37**, 31–66 (2003). doi: [10.1146/annurev.genet.37.042203.132748](https://doi.org/10.1146/annurev.genet.37.042203.132748); pmid: [14616055](https://pubmed.ncbi.nlm.nih.gov/14616055/)
- A. G. Lada *et al.*, Genome-wide mutation avalanches induced in diploid yeast cells by a base analog or an APOBEC deaminase. *PLOS Genet.* **9**, e1003736 (2013). doi: [10.1371/journal.pgen.1003736](https://doi.org/10.1371/journal.pgen.1003736); pmid: [24039593](https://pubmed.ncbi.nlm.nih.gov/24039593/)
- I. B. Rogozin *et al.*, Evolution and diversification of lamprey antigen receptors: Evidence for involvement of an AID-APOBEC family cytosine deaminase. *Nat. Immunol.* **8**, 647–656 (2007). doi: [10.1038/ni1463](https://doi.org/10.1038/ni1463); pmid: [17468760](https://pubmed.ncbi.nlm.nih.gov/17468760/)
- I. M. Slaymaker *et al.*, Rationally engineered Cas9 nucleases with improved specificity. *Science* **351**, 84–88 (2016). doi: [10.1126/science.1252277](https://doi.org/10.1126/science.1252277); pmid: [26628643](https://pubmed.ncbi.nlm.nih.gov/26628643/)
- Z. G. Wang, D. G. Smith, D. W. Mosbaugh, Overproduction and characterization of the uracil-DNA glycosylase inhibitor of bacteriophage PBS2. *Gene* **99**, 31–37 (1991). doi: [10.1016/0378-1119\(91\)90030-F](https://doi.org/10.1016/0378-1119(91)90030-F); pmid: [1902430](https://pubmed.ncbi.nlm.nih.gov/1902430/)

23. J. S. Lee, T. B. Kallehauge, L. E. Pedersen, H. F. Kildegaard, Site-specific integration in CHO cells mediated by CRISPR/Cas9 and homology-directed DNA repair pathway. *Sci. Rep.* **5**, 8572 (2015). doi: [10.1038/srep08572](https://doi.org/10.1038/srep08572); pmid: [25712033](https://pubmed.ncbi.nlm.nih.gov/25712033/)
24. J. Shin *et al.*, Efficient CRISPR/Cas9-mediated multiplex genome editing in CHO cells via high-level sgRNA-Cas9 complex. *Biotechnol. Bioprocess Eng.* **20**, 825–833 (2015). doi: [10.1007/s12257-015-0233-7](https://doi.org/10.1007/s12257-015-0233-7)
25. S. Q. Tsai *et al.*, GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nat. Biotechnol.* **33**, 187–197 (2015). doi: [10.1038/nbt.3117](https://doi.org/10.1038/nbt.3117); pmid: [25513782](https://pubmed.ncbi.nlm.nih.gov/25513782/)
26. M. Stemmer, T. Thumberger, M. Del Sol Keyer, J. Wittbrodt, J. L. Mateo, CCTop: An intuitive, flexible and reliable

- CRISPR/Cas9 target prediction tool. *PLOS ONE* **10**, e0124633 (2015). doi: [10.1371/journal.pone.0124633](https://doi.org/10.1371/journal.pone.0124633); pmid: [25909470](https://pubmed.ncbi.nlm.nih.gov/25909470/)
27. A. C. Komor, Y. B. Kim, M. S. Packer, J. A. Zuris, D. R. Liu, Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* **533**, 420–424 (2016). doi: [10.1038/nature17946](https://doi.org/10.1038/nature17946); pmid: [27096365](https://pubmed.ncbi.nlm.nih.gov/27096365/)

#### ACKNOWLEDGMENTS

We thank M. Yajima for helpful discussion. This work was supported by a Special Coordination Fund for Promoting Science and Technology, Creation of Innovative Centers for Advanced Interdisciplinary Research Areas (Innovative Bioproduction Kobe) from the Ministry of Education, Culture, Sports and Technology (MEXT) of Japan. This work was also partly supported by the commission for Development of Artificial Gene Synthesis Technology for Creating Innovative Biomaterial from the

Ministry of Economy, Trade and Industry (METI), Japan, by JSPS KAKENHI grant numbers 26119710 and 16K14654, and by Cross-ministerial Strategic Innovation Promotion Program (SIP); 'Technologies for creating next-generation agriculture, forestry and fisheries.' Patents have been filed related to this work.

#### SUPPLEMENTARY MATERIALS

[www.sciencemag.org/content/353/6305/aaf8729/suppl/DC1](http://www.sciencemag.org/content/353/6305/aaf8729/suppl/DC1)  
Materials and Methods  
Figs. S1 to S11  
Tables S1 to S9  
References

13 April 2016; accepted 25 July 2016  
Published online 4 August 2016  
[10.1126/science.aaf8729](https://doi.org/10.1126/science.aaf8729)