International Immunology, Vol. 33, No. 7, pp. 387–398
doi:10.1093/intimm/dxab019
Advance Access publication 26 April 2021

Stepwise generation of AID knock-in and conditional knockout mice from a single gene-targeting event

Kazuo Kinoshita[1,2], Munehiro Uemura[2], Takahiro Shimizu[3], Shun Kinoshita[4] and **Hiroyuki Marusawa[3]**

[1]Evolutionary Medicine, Shizuoka Graduate University of Public Health, 4-27-2 Kita-ando, Aoi-ku, Shizuoka 420-0881, Japan
[2]Shiga Medical Center Research Institute, Moriyama 524-0022, Japan
[3]Department of Gastroenterology and Hepatology, Kyoto University Graduate School of Medicine, Kyoto 606-8501, Japan
[4]Kyoto University Graduate School of Medicine Faculty of Medicine, Kyoto 606-8501, Japan

Correspondence to: K. Kinoshita; E-mail: kkinoshi@mac.com

Received 11 January 2021, *editorial decision* 19 April 2021; *accepted* 22 April 2021

Abstract

Activation-induced cytidine deaminase (AID) encoded by the *Aicda* gene initiates class-switch recombination and somatic hypermutation of immunoglobulin genes. In addition to this function, AID is also implicated in the epigenetic regulation in pluripotent stem cells and in the oncogenesis of lymphoid and non-lymphoid origins. To examine AID's role in specific cell types, we developed mouse strains of conditional knockout (Aicda-FL) and knock-in with a red fluorescent protein gene (RFP) inserted into the *Aicda* locus (Aicda-RFP). These two strains were obtained from a single targeting event in embryonic stem cells by a three-*loxP* or tri-lox strategy. Partial and complete recombination among the three *loxP* sites in the Aicda-RFP locus gave rise to Aicda-FL and AID-deficient loci (Aicda-KO), respectively, after mating Aicda-RFP mice with Cre-expressing mice driven by tissue-non-specific alkaline phosphate promoter. We confirmed RFP expression in B cells of germinal centers of intestine-associated lymphoid tissue. Mice homozygous for each allele were obtained and were checked for AID activity by class-switch and hypermutation assays. AID activity was normal for Aicda-FL but partially and completely absent for Aicda-RFP and Aicda-KO, respectively. Aicda-FL and Aicda-RFP mice would be useful for studying AID function in subpopulations of B cells and in non-lymphoid cells.

Keywords: activation-induced cytidine deaminase, antibody diversity, class switch, red fluorescent protein, somatic hypermutation

Introduction

Activation-induced cytidine deaminase (AID) is required for antibody diversification by class-switch recombination (CSR) and somatic hypermutation (SHM) (1, 2). AID is expressed predominantly in germinal center (GC) B cells, which constitutively populate mucosa-associated lymphoid tissue, represented by Peyer's patches of the small intestine (3). Although the action of AID focuses to genes encoding immunoglobulin heavy and light chains, off-target mutations by AID were reported to occur and thought to drive the development of lymphoma and leukemia (4, 5). As cytokine- or pathogen-stimulated epithelial cells (e.g. liver, stomach, intestine, bile duct, lung, mammary gland and skin) also express AID, it is also implicated in the genesis of epithelial cancers (6, 7). Another possible function of AID is epigenetic control of pluripotent stem cells by the deamination of methylated cytosines, resulting in cytosine demethylation (8). To investigate these functions of AID in various cell types, conditional knockout mice would be helpful, but have yet to be made available. A mouse strain with a fluorescent protein knocked in at the *Aicda* locus would also be beneficial for detecting AID-expressing cells *in vivo*. Therefore, in this study, we created both strains of mouse using a single kind of targeting vector, so that a conditional knockout strain (Aicda-FL) could be generated from partial Cre-mediated recombination of the knock-in locus (Aicda-RFP). We examined AID activity in these strains and a derivative complete knockout strain (Aicda-KO).

Methods

Targeting vector construction

A targeting vector (pSub-mAID-DTA) was constructed using the Red/ET recombination system (Gene Bridges, Heidelberg, Germany). A genomic region (12 175 bp; chr6:122,519,221–122,531,395 on NCBI36/mm8) containing all five exons of the mouse *Aicda* gene was

sub-cloned into a plasmid (pSub-mAID) from a bacterial artificial chromosome, RP24-68I7 (BACPAC Resources Center), harboring a 190-kb mouse genomic sequence surrounding the *Aicda* gene (Supplementary Figure 1). A DNA fragment (ColE1-amp-AID) of an ampicillin resistance gene connected to the bacterial replication origin flanked by restriction sites and a 50-bp sequence homologous to the cloning ends of the *Aicda* gene was prepared by PCR using primers AID-D-1F and AID-U-1R. *Escherichia coli* harboring RP24-68I7 and inducible Red/ET recombinase was electroporated with ColE1-amp-AID before recombinase induction. PCR screening could identify successfully recombined colonies. pSub-mAID obtained in this way was further modified with targeted insertion of a floxed chloramphenicol resistance gene [loxP-cm-loxP-AIDintC2 obtained by nested PCR, first with AID-intC-1F and AID-intC-1R (50-bp homology arm) and then with AID-intC-2F and AID-intC-2R (120-bp homology arm)] and subsequent Cre transfection to leave a single-*loxP* site after exon 3 (pSub-mAID-loxP). Exon 2 on pSub-mAID-loxP was replaced with a fragment containing the coding sequence of tandem-dimeric red fluorescent protein (tdRFP) and a neomycin resistance gene (Neo), named as e2-tdimer2-TFa, as explained in the next paragraph. pSub-mAID-KI obtained in this way was further modified with diphtheria toxin A gene (9) inserted at the *Asc*I site, which enables negative selection against embryonic stem (ES) cells with random integration. This pSub-mAID-DTA is the targeting vector used in this study.

The preparation of targeting fragment e2-tdimer2-TFa required five steps of ligation with five PCR products and excision with restriction enzymes. The tdRFP coding sequence together with the polyadenylation signal was amplified from genomic DNA of tdRFP mice (10) by PCR with e2-tdimer-3F and e2-tdimer-1R, which appended a 26-bp splice acceptor from exon 2 of the *Aicda* gene in-frame fused with the tdRFP coding sequence. This fragment was inserted into pFlox1.4, which contains two *loxP* sites and interposed cloning sites. The intronic sequence before exon 2 was extended by incorporating a 138-bp fragment of PCR product e2-1 amplified from the mouse genome using the primers e2-1F and e2-1R, giving rise to pe2-tdimer2-Flox1. This ligation step utilized a natural *Mme*I site 7 bp upstream of the exon 2 splicing acceptor. pe2-tdimer2-Flox1 was sequentially added with FRT-PGK-gb2-neo-FRT (Gene Bridges), a 463-bp downstream homology arm (D-arm-1 by PCR primers: D-arm-1F and D-arm-1R) containing *Aicda* exon 2, and a 492-bp upstream homology arm (U-arm-2 by PCR primers: U-arm-2F and U-arm-3R), resulting in pe2-tdimer2-UDa. *Sma*I and *Ssp*I digestion of this plasmid yielded e2-tdimer2-TFa.

Generation of knock-in mice

Mouse ES cells (BRUCE-4 with C57BL/6J background) were electroporated with *Swa*I-linearized pSub-mAID-DTA. Initial PCR screening for homologous recombinants using the primers MCD-R (forward primer within Neo) and R2 (reverse primer located in the chromosomal region downstream of the 3′ arm) detected 68 positive clones among 79 neomycin-resistant clones. The presence of tdRFP and *loxP* downstream of exon 3 was confirmed by PCR in 11 clones of 31 clones examined. Southern blotting using an upstream probe (PCR amplified with 5′N2-F and 5′N2-R), a downstream probe (PCR amplified with 3′N2-F and 3′N2-R) and a Neo probe (PCR amplified with NeoP-F and NeoF-R) confirmed single-copy intact homologous recombination in all of the eight ES clones examined.

One successfully targeted ES cell clone was injected into blastocysts with a Balb/c background and transplanted in pseudo-pregnant surrogate female mice of the ICR background. A born black-haired chimeric mouse (44-1) was crossed with wild-type female mice of the C57BL/6 background. Among eight F_1 offspring from two crosses, five were positive for the targeted allele. A male heterozygous F_1 mouse (44-1-4) was crossed with a female FLP recombinase transgenic mouse driven by the CAG promoter (C57BL/6J background) to remove the Neo gene. The resultant mice without Neo, designated as Aicda-RFP, were maintained by crossing with wild-type C57BL/6J mice.

To remove tdRFP but retain exons 2 and 3, Aicda-RFP mice were crossed with TNAP-Cre mice (backcrossed 11 times with C57BL/6J from the original 129 background), which express Cre under the promoter of tissue-non-specific alkaline phosphatase (TNAP). Female Aicda-RFP mice were crossed with male TNAP-Cre mice to obtain a male Aicda-RFP mouse with the TNAP-Cre transgene. Further crossing with female wild-type mice resulted in deletion of tdRFP with exons 2 and 3 intact (Aicda-FL). At the same time, a complete deletion product without exons 2 and 3 was also obtained (Aicda-KO). Mice homozygous for Aicda-RFP, Aicda-FL and Aicda-KO were viable and fertile.

Aicda-RFP, Aicda-FL, and Aicda-KO mice and TNAP-Cre mice backcrossed 20 times to the C57BL/6J background were deposited in and are available from RIKEN Bioresource Research Center (Tsukuba, Japan) under catalog numbers, RBRC10943, RBRC10944, RBRC10945 and RBRC10946, respectively. All mouse experiments were approved by the IACUC of Shiga Medical Center Research Center.

Genotyping

Discrimination of the Aicda-RFP or Aicda-FL allele from the wild type was performed using primers spanning the *loxP* site after exon 3 (AID-loxP3-F3 and AID-loxP3-R3). Owing to the insertion of *loxP*, Aicda-RFP and Aicda-FL alleles gave a 156-bp product, while the wild type gave a 59-bp product (Supplementary Figure 2). These two products can be separated by agarose gel electrophoresis or melting curve analysis by SYBR dye-supplemented real-time PCR (two-step cycles with annealing/extension at 60°C).

The presence of the tdRFP exon was checked by PCR using dimer2-120F and dimer2-211R. Combined with the above PCRs, this can discriminate Aicda-RFP from Aicda-FL.

For the genotyping of Aicda-KO, a primer pair spanning the *loxP* site left after deletion of exons 2 and 3 by the Cre reaction (AID-loxP1-F2 and AID-loxP3-R) was used, amplifying a 279-bp product. Inclusion of the third primer (AID-loxP2-R) can detect the wild-type allele, producing a 66-bp product (Supplementary Figure 2). The two products can be separated by agarose gel electrophoresis or melting curve analysis

by SYBR dye-supplemented real-time PCR (two-step cycles with annealing/extension at 60°C).

Genotyping of Aicda-RFP-KO was checked by PCR using SV40PolyA3F and AID-loxP3-R. The primer sequences are listed in Supplementary Table 1.

Fluorescent stereomicroscopy

Red fluorescence of intestinal Peyer's patches and mesenteric lymph nodes from an Aicda-RFP heterozygous mouse (30-week-old male) was observed under a stereoscopic microscope, MZ16F (Leica Microsystems, Wetzlar, Germany) and recorded with a cooled CCD camera, VB-6000 (Keyence, Osaka, Japan).

Flow cytometry and cell sorting

Peyer's patches of the small intestine were manually dissected under a stereoscopic microscope and disintegrated into a single-cell suspension between two leaves of slide glass. A total of 3×10^6 cells were stained with Alexa647-conjugated peanut hemagglutinin (PNA), FITC-anti-mouse CD95 and 7-AAD after blocking with TruStain FcX PLUS (BioLegend, San Diego, CA, USA; 156604). GC B cells (PNA^{high}, $CD95^+$) were sorted using FACSMelody (Becton Dickinson, Franklin Lakes, NJ, USA). Data were analyzed using FlowJo (Becton Dickinson). 7-AAD-positive cells were excluded from the analysis.

AID western blot

Male mice were sacrificed and splenocytes were isolated. After hemolysis with ACK buffer (150 mM ammonium chloride, 1 mM potassium bicarbonate, 0.1 mM EDTA, pH 7.3), cells were counted with trypan blue exclusion. Frozen cell stocks were prepared with 1×10^7 cells suspended in 1 ml of Bambanker medium (Nippon Genetics, Tokyo, Japan) and frozen at –80°C. A total of 1×10^7 thawed cells were suspended in 6 ml of RPMI-Spleen [RPMI1640 (Nacalai Tesque, Kyoto, Japan; 30264-85) supplemented with 10% fetal bovine serum, 1× non-essential amino acid (Invitrogen, Carlsbad, CA, USA; 11140-050), 1 mM sodium pyruvate (Invitrogen, 11360-070), 0.5 mM 2-mercaptoethanol, 1× penicillin/streptomycin (Nacalai Tesque, 26253-84)] and 50 μg ml^{-1} lipopolysaccharide (LPS; Sigma, St Louis, MO, USA; L6511). After incubation at 37°C with a 5% CO_2 atmosphere for 3 days, 1×10^7 cells were lysed with 330 μl 1× phosphate-buffered saline, 0.5% (v/v) Nonidet P-40 substitute (Nacalai Tesque; 18558-54), 1× complete protease inhibitor cocktail (Roche, Basel, Switzerland). Supernatants after centrifugation at 16 000 × *g* for 20 min were subjected to sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE) using a 5–20% gradient gel (Nacalai Tesque). Cell lysate equivalent to 3.8×10^5 cells per lane was loaded under reducing conditions. Semidry-blotted membranes were probed with either anti-AID monoclonal antibody (MAID-2) or anti-β-tubulin antibody (TU27), using SNAP i.d. 2.0 apparatus (Merck, Darmstadt, Germany). After incubation with secondary antibodies labeled with fluorescence dyes, membranes were scanned with an Odyssey scanner (Li-cor, Lincoln, NE, USA).

Reverse transcription–PCR

Total RNA was isolated from *in vitro* culture of splenocytes from frozen stocks and stimulated with 50 μg ml^{-1} LPS, LPS plus mouse IL-4 (Peprotech, Rocky Hill, NJ, USA; 214-14) or LPS plus human TGF-β1 (Peprotech; 100-21) for 2 days using a QuickGene RC-S2 kit (Kurabo, Osaka, Japan). First-strand cDNA was synthesized from 500 ng of total RNA using an iScript advanced kit (Bio-Rad). Quantitative PCR was performed by Light Cycler 96 (Roche) using QuantiTect enzyme mix (Qiagen, Düsseldorf, Germany) and three sets of intron-spanning primers and probes: (i) mAID-TM1 (exons 2–3), (ii) mAID-TM2 (exons 4–5) and (iii) mAID-TM4 (exons 3–4). Quantitative reverse transcription (RT)–PCR spanning *Aicda* exons 1 and 2 was performed as above except using Superscript IV (Thermo Fisher Scientific, Waltham, MA, USA) and KAPA SYBR FAST qPCR Master Mix (Roche) by TP980 (Takara Bio, Kusatsu, Japan). RFP transcripts were quantified using mAID-e1-2F and dimer2-67R. The values are presented in arbitrary units after calibration with 18S ribosomal RNA levels. Sequences of primers and probes are listed in Supplementary Table 1.

In vitro *CSR assay*

Male mice with various genotypes aged from 33 to 84 weeks were sacrificed and their splenocytes were isolated. After hemolysis, the cells were counted with trypan blue exclusion. A total of 2×10^6 cells were suspended in 2 ml of RPMI-Spleen with 25 μg ml^{-1} LPS and 15 ng ml^{-1} mouse IL-4. After incubation at 37°C with a 5% CO_2 atmosphere for 3 days, the cells were stained with APC-conjugated anti-mouse IgG1 antibody (Becton Dickinson, 560089) after Fc receptor blocking with TruStain FcX PLUS (BioLegend). Stained cells were analyzed by FACSMelody (Becton Dickinson).

Enzyme-linked immunosorbent assay

The serum concentrations of all isotypes of immunoglobulins except IgD and IgE were measured by enzyme-linked immunosorbent assay (ELISA). The antibodies used for measuring of IgM, IgG1, IgG2b, IgG3 and IgA are listed in Supplementary Table 2. A mouse IgG2c ELISA Kit (Bethyl Laboratories Inc., Montgomery, TX, USA) was used to quantify IgG2c.

SHM analysis

Genomic DNA of GC B cells from Peyer's patches was purified with QIAamp DNA Mini kit (Qiagen) and the intronic sequence downstream of the rearranged V genes was amplified using previously reported primers (11) and PrimeSTAR Max polymerase (Takara Bio). After agarose electrophoresis, the shortest one (approximately 1250 bp) of the four bands corresponding to four different J_H (1–4) usages was excised and cloned into pCR4Blunt-TOPO (Invitrogen) before TempliPhi amplification (GE Healthcare, Chicago, IL, USA) and Sanger sequencing from M13 forward and M13 reverse primers using BigDye version 1.1 and ABI3130xl sequencer (Applied Biosystems, Foster City, CA, USA). The assignment of V, D and J segments was performed by Ig-BLAST (NCBI), revealing the combination of V genes

of the J558 family, various D and J_H4. Sequences apparently duplicated by PCR were excluded from the analysis. Using the intronic 1106-bp sequence downstream of J_H4 of the C57BL/6J genome (chr12:113,427,408–113,428,513 on GRCm38/mm10) as a reference, single-base substitutions were cataloged using R software. The mutational signature was represented after adjustment by trinucleotide frequency of the sequenced region, in accordance with the literature (12, 13).

Results

Gene-targeting strategy

The purpose of this study is to generate both mice with conditional knockout of the *Aicda* locus (Aicda-FL) and mice with knock-in of a fluorescent protein (Aicda-RFP) from a single targeting event in ES cells (Fig. 1). The targeting vector pSub-mAID-DTA was designed so that the Neo can be deleted by FLP recombinase, leaving an exon encoding tdRFP knocked in upstream of exon 2 of the *Aicda* locus (Aicda-RFP). As the tdRFP exon is flanked by *loxP* sites, it can be converted to a single-*loxP* site (Aicda-FL) by mating with Cre-expressing mice. This *loxP* site together with another *loxP* site downstream of exon 3 flank a functionally critical element, which can be deleted by Cre expression, resulting in a knockout state (Aicda-KO). To obtain Aicda-FL from the Aicda-RFP configuration, complete recombination, which gives rise to the Aicda-KO locus, must be avoided.

Following gene targeting in ES cells and the production of chimeric mice, gene-targeted mice were crossed with FLP-expressing mice to eliminate the Neo cassette and obtain Aicda-RFP mice. When a male TNAP-Cre mouse that expresses Cre in primordial germ cells was crossed with a female Aicda-RFP homozygote, offspring inheriting the TNAP-Cre allele included not only the Aicda-KO configuration, but also Aicda-RFP, Aicda-RFP-KO (partial recombination between the second and third *loxP*s) and a mosaic of all other sorts of recombination products (Table 1). Because DNA for genotyping was extracted from auricles, observed mosaicism did not necessarily parallel with that of germ cells. One male offspring possessed both Aicda-RFP and TNAP-Cre alleles, which was mated with a female wild-type mouse. Among the resultant 20 offspring in three pregnancies, two mice with the Aicda-FL locus were obtained together with eight heterozygous Aicda-KO mice (Table 1).

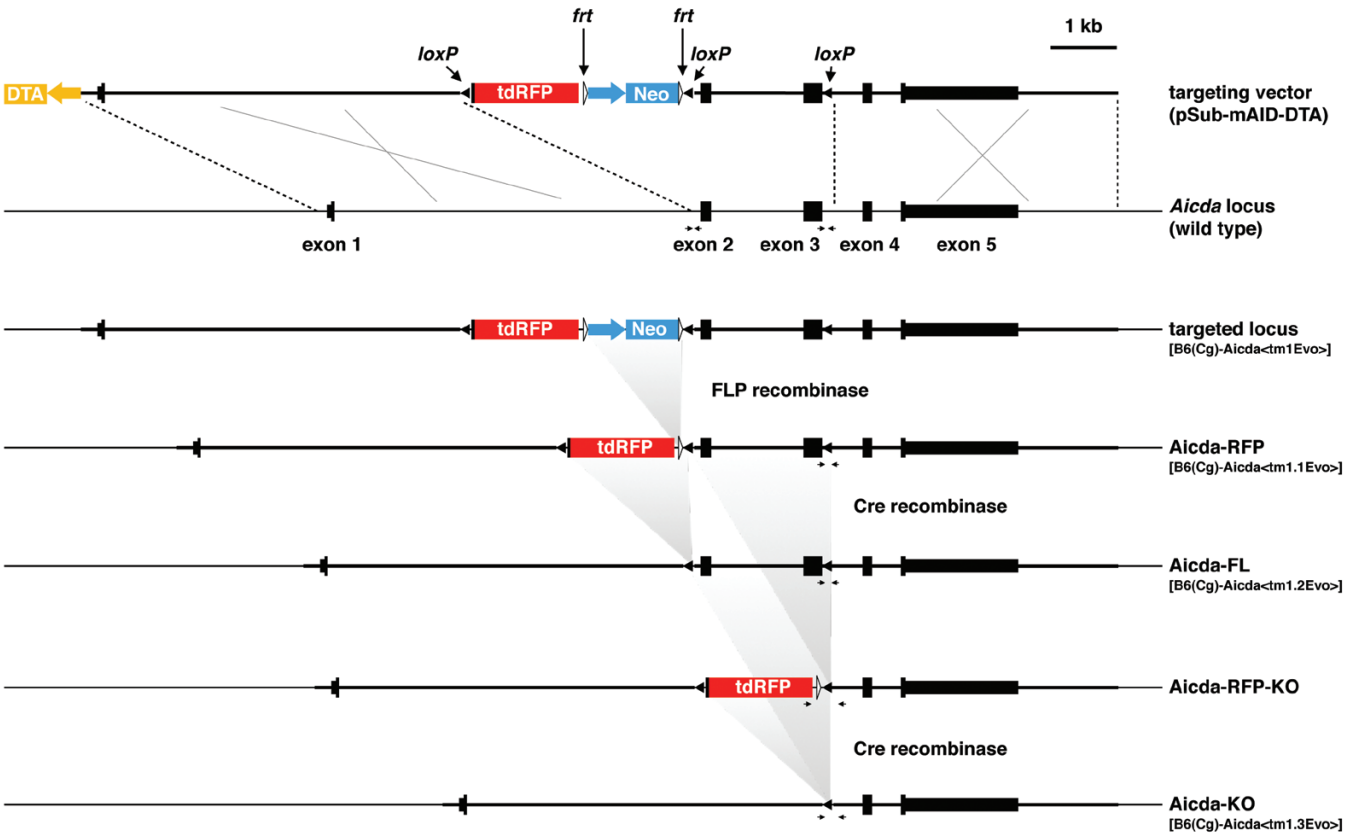


Fig. 1. Gene-targeting scheme and recombination products. In the targeted locus, *loxP*-flanked element containing tandem-dimer RFP (tdRFP) fused with splice acceptor of exon 2 and *frt*-flanked Neo cassette was inserted before exon 2 of the *Aicda* gene. The third *loxP* was introduced after exon 3. ES cells with a C57BL/6J background (BRUCE-4) were electroporated with the targeting vector. Targeted mice were crossed with FLP-transgenic mice (C57BL/6J background) to obtain Aicda-RFP mice. Aicda-RFP mice (three *loxP*s) were crossed with male TNAP-Cre mice (C57BL/6J background) that express Cre in primordial germ cells to obtain 2-*loxP*s partial recombination products of Aicda-FL and Aicda-RFP-KO (deletion between the second and third *loxP*s) and single-*loxP* Aicda-KO allele. Horizontal arrows represent positions of primers for genotyping. DTA, diphtheria toxin A; Neo, neomycin resistance gene.

RFP expression in Aicda-RFP mice

AID is expressed constitutively in GCs of intestine-associated lymphoid tissue such as Peyer's patches and mesenteric lymph nodes. Peyer's patches of Aicda-RFP mice contained four to six cell clusters with RFP expression (Fig. 2A and B). Similar RFP-positive clusters were also observed in mesenteric lymph nodes (Fig. 2C). RFP expression was recognized in other lymphoid organs including the spleen and lymph nodes, reflecting physiological sites of AID expression. However, it was not detectable in the skin, brain, thymus, testes and the surface of other abdominal and thoracic organs, for which history of AID expression could be detected by a fate-mapping approach using *Aicda* promoter-driven Cre transgenic and LacZ reporter mice (14). This suggested that AID expression in those tissues is weak and transient in limited cell populations.

Table 1. Genotype of pups after mating of Aicda-RFP and TNAP-Cre mice

Genotype	♀ Aicda-RFP/RFP × ♂ TNAP-Cre	♀ Wild type × ♂ Aicda-RFP+ TNAP-Cre
Wild type	0 (0)	9 (3)
Aicda-RFP	3 (1)	0 (0)
Aicda-RFP-KO	1 (1)	0 (0)
Aicda-FL	0 (0)	2 (0)
Aicda-KO	4 (4)	8 (5)
Aicda-[FL, KO]	0 (0)	1 (1)
Aicda-[RFP, FL, RFP-KO, KO]	1 (1)	0 (0)
Aicda-[RFP-KO, KO]	1 (1)	0 (0)

The number of pups with the indicated genotype is shown after two kinds of crossing: female Aicda-RFP homozygote with male TNAP-Cre transgenic mouse and female wild type with male Aicda-RFP heterozygote with TNAP-Cre transgene. Both types of partial recombination products Aicda-FL and Aicda-RFP-KO were obtained. Multiple genotypes within brackets indicate chimerism in auricular DNA. Number in parentheses is the number of pups possessing the TNAP-Cre transgene.

Flow cytometric analysis of Peyer's patches of Aicda-RFP mice confirmed that 97% of GC B cells defined as PNA^{high}, $CD95^{+}$ were positive for RFP (Fig. 3). This observation indicated that RFP expression recapitulated that of AID, and that RFP-positive structures within Peyer's patches and lymph nodes were GCs. Notably, the percentages of GC B cells in Peyer's patches of homozygous Aicda-RFP (43%) and Aicda-KO (38%) mice were higher than that of the control wild-type mice (6%). This suggested AID insufficiency in Aicda-RFP homozygotes, because previous AID-deficient mice developed a phenotype with enlarged GCs (1, 15).

The intensity of RFP fluorescence may reflect *Aicda* promoter activity. This hypothesis arose from an observation of compound heterozygotes of Aicda-RFP and Aicda-KO (RFP/KO). The RFP intensity of GC cells sorted from Peyer's patches of RFP/KO

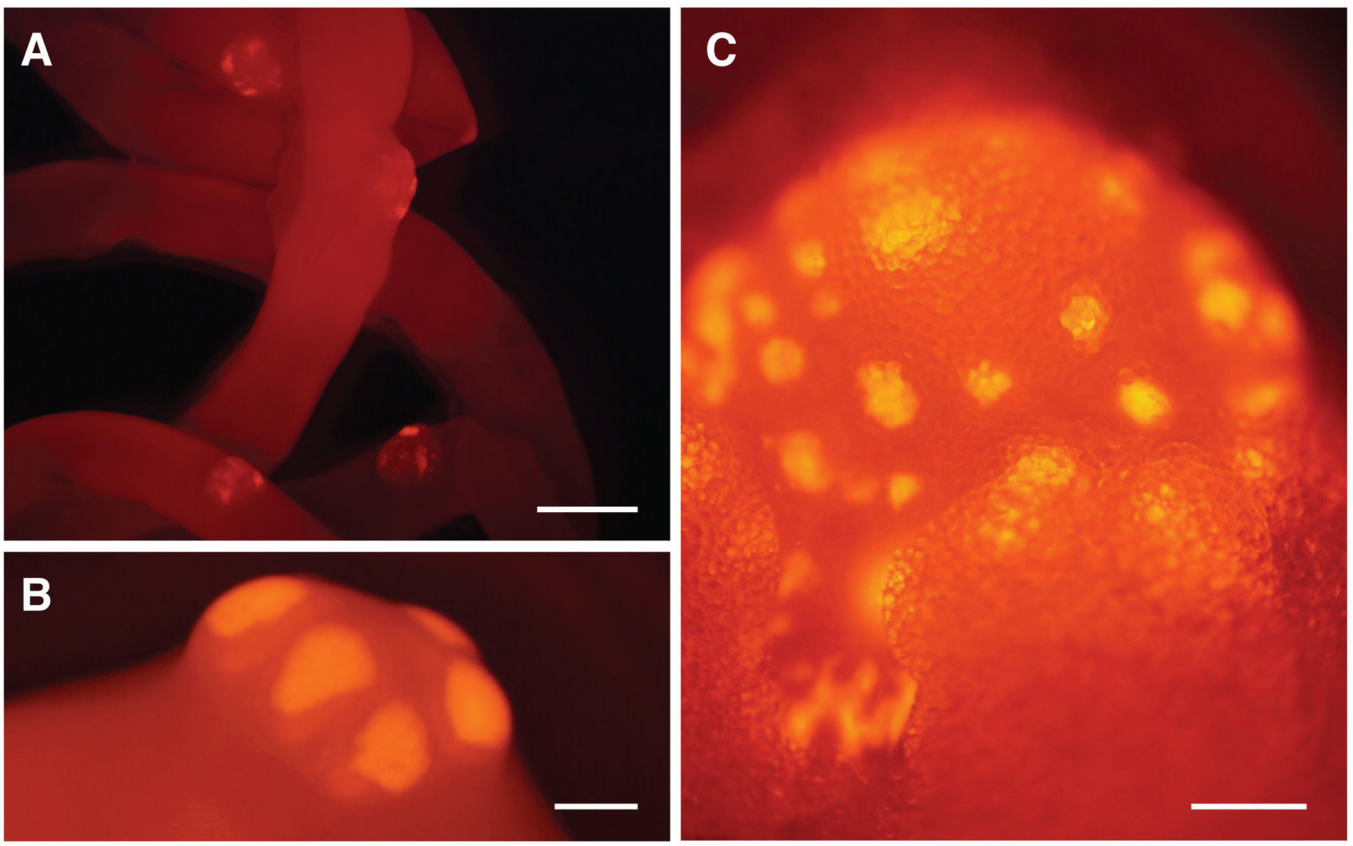


Fig. 2. RFP expression in GCs in Aicda-RFP heterozygous mouse. RFP fluorescence was observed in GCs. (A) Small intestine. (B) Peyer's patch. (C) Mesenteric lymph node. Bar: (A) 2 mm, (B) 500 μm, (C) 1 mm.

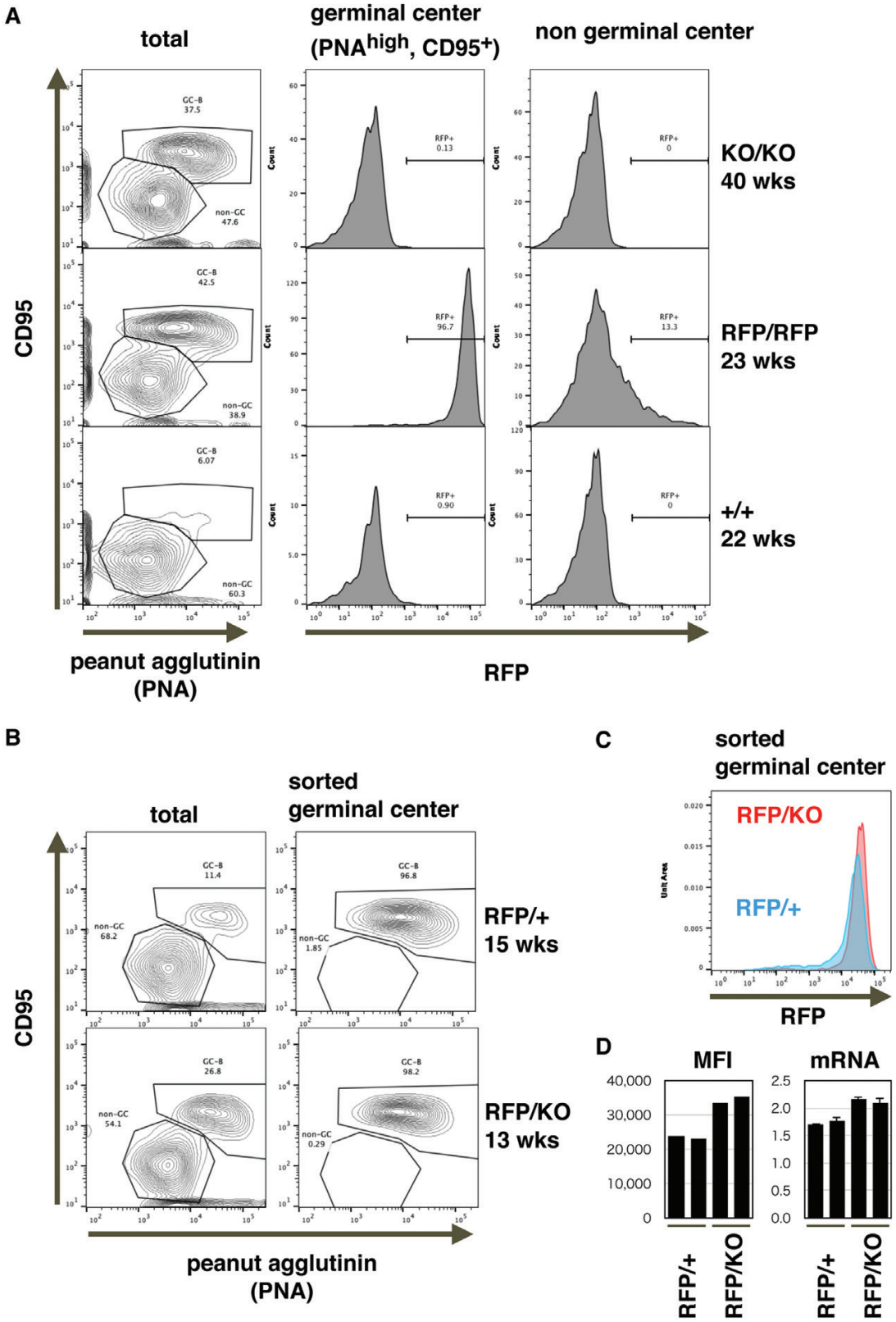


Fig. 3. Flow cytometric analysis of Peyer's patches from Aicda-RFP and Aicda-KO homozygotes and compound heterozygotes. (A) Cells from Peyer's patches of mice with indicated genotypes and age were stained with FITC-labeled anti-CD95 (Fas) and Alexa647-labeled peanut agglutinin (PNA) and were analyzed by flow cytometry. GC cells with PNA^{high}, CD95^{+} were gated for a histogram plot showing RFP expression levels. Non-GC cells were analyzed in comparison. (B) Flow cytometric profiles of Peyer's patches and sorted GC cells from an Aicda-RFP heterozygote (RFP/+) and compound heterozygote (RFP/KO). (C) Histogram of RFP fluorescence intensity of sorted GC cells from RFP/+ (blue) and RFP/KO (red) mice. (D) Comparison of RFP expression levels in sorted GC cells between RFP/+ and RFP/KO. Left and right panels, respectively, show mean fluorescent intensity (MFI) measured by flow cytometry and mRNA amount by quantitative RT-PCR after calibration with the ribosomal 18S amount. For the mRNA quantitation, data represent the mean of triplicate assay with error bars showing standard errors.

mice was 1.5 times higher than that of Aicda-heterozygous mice (RFP/+) (Fig. 3B–D, left). In parallel with this, the level of RFP transcripts was slightly higher in RFP/KO GC cells than in RFP/+ GC cells (Fig. 3D, right). Of note, insertion of the RFP-coding sequence and two *loxP* and single *frt* sites does not delete any nucleotides but accompanies 140-bp duplication of the region containing the exon 2 splice acceptor, making a possibility that observed transcriptional upregulation is an artifact.

AID expression from Aicda-RFP and Aicda-FL alleles

To confirm the AID deficiency in Aicda-RFP homozygotes, splenocytes cultured from frozen stocks were stimulated with LPS for 3 days and cell lysates were subjected to western blotting using anti-AID antibody. As anticipated, samples from Aicda-RFP and Aicda-KO homozygotes were devoid of 25-kDa bands, while the Aicda-FL homozygote expressed AID at a level comparable to that of the wild-type control (Fig. 4A; Supplementary Figure 3). The presence of two bands was previously reported (16); although the reason for this is unknown, it might reflect splice variants reported for human AID (17, 18).

To detect AID expression with higher sensitivity than that of western blotting, total RNA was purified from splenocytes stimulated with LPS for 2 days. Quantitative PCR using primers between exons 1 and 2 and those between exons 2 and 3 identified *Aicda* transcripts in Aicda-RFP homozygotes at a level 16–17 times lower than that of the wild type (Fig. 4B, left upper two panels). The reduction in Aicda-RFP was somewhat milder with other primer pairs: 3.6-fold less when PCR was performed between exons 3 and 4, and 8.6-fold less when between exons 4 and 5 (Fig. 4B, left lower two panels). Although exact estimation of the reduction was difficult, low level transcription appeared to occur in stimulated B cells from Aicda-RFP homozygous mice, despite the presence of two polyadenylation sites after the tdRFP coding sequence. When splenocytes were stimulated with LPS and IL-4 and with LPS and TGF-β1, Aicda-RFP homozygotes expressed 17- and 20-fold less Aicda transcripts compared to wild-type splenocytes, respectively (Fig. 4B, top panels). Direct sequencing of RT-PCR products spanning exons 1-2 from Aicda-RFP homozygotes revealed a nucleotide sequence identical to those from wild-type mice, indicating that skipping of RFP-coding exons occurred (data not shown).

Evaluation of CSR

To examine the CSR function of B cells from Aicda-RFP, Aicda-FL and Aicda-KO homozygous mice, freshly isolated splenocytes were stimulated *in vitro* with LPS and IL-4 or TGF-β1 for 3 days. The induction of IgG1 was measured by flow cytometry (Fig. 5A, top). Aicda-KO and Aicda-RFP homozygotes showed only background levels of positivity for IgG1, while the Aicda-FL homozygote showed IgG1 induction as strong as that of the wild-type control and the Aicda-KO heterozygote. This indicated that Aicda-KO and Aicda-RFP loci are defective for AID, while the Aicda-FL locus can express AID at a level comparable to that of the wild-type locus. The CSR defect of Aicda-RFP homozygotes was observed as well for IgG2b after LPS stimulation with and without TGF-β1 (Fig. 5A, bottom).

To examine the CSR potency *in vivo*, serum immunoglobulin concentrations were examined by isotype-specific ELISA (Fig. 5B). As reported previously (1), Aicda-KO homozygotes had no IgG1, IgG2b, IgG2c, IgG3 or IgA in the serum, while they had a higher concentration of IgM than the wild-type mice. The levels of each isotype in Aicda-FL homozygotes were similar to those in the wild type. The Aicda-RFP

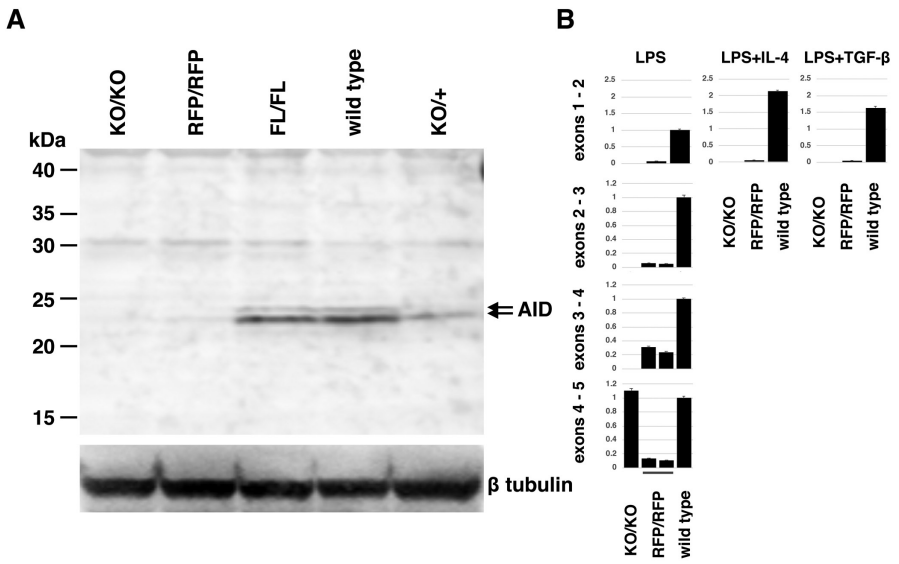


Fig. 4. Expression of AID in Aicda-RFP and Aicda-FL homozygotes. (A) Western blots of AID protein in splenocytes of indicated genotypes stimulated *in vitro* with LPS for 3 days. β-Tubulin was probed to indicate equal loading among wells. (B) Quantitative RT-PCR for AID transcripts in splenocytes stimulated with LPS (and IL-4 or TGF-β1) for 2 days. The PCR primers are located within indicated exons. For Aicda-RFP homozygotes, two individuals were analyzed except a primer pair spanning exons 1 and 2. The results were calibrated with the amount of ribosomal 18S RNA, and error bar indicates standard error.

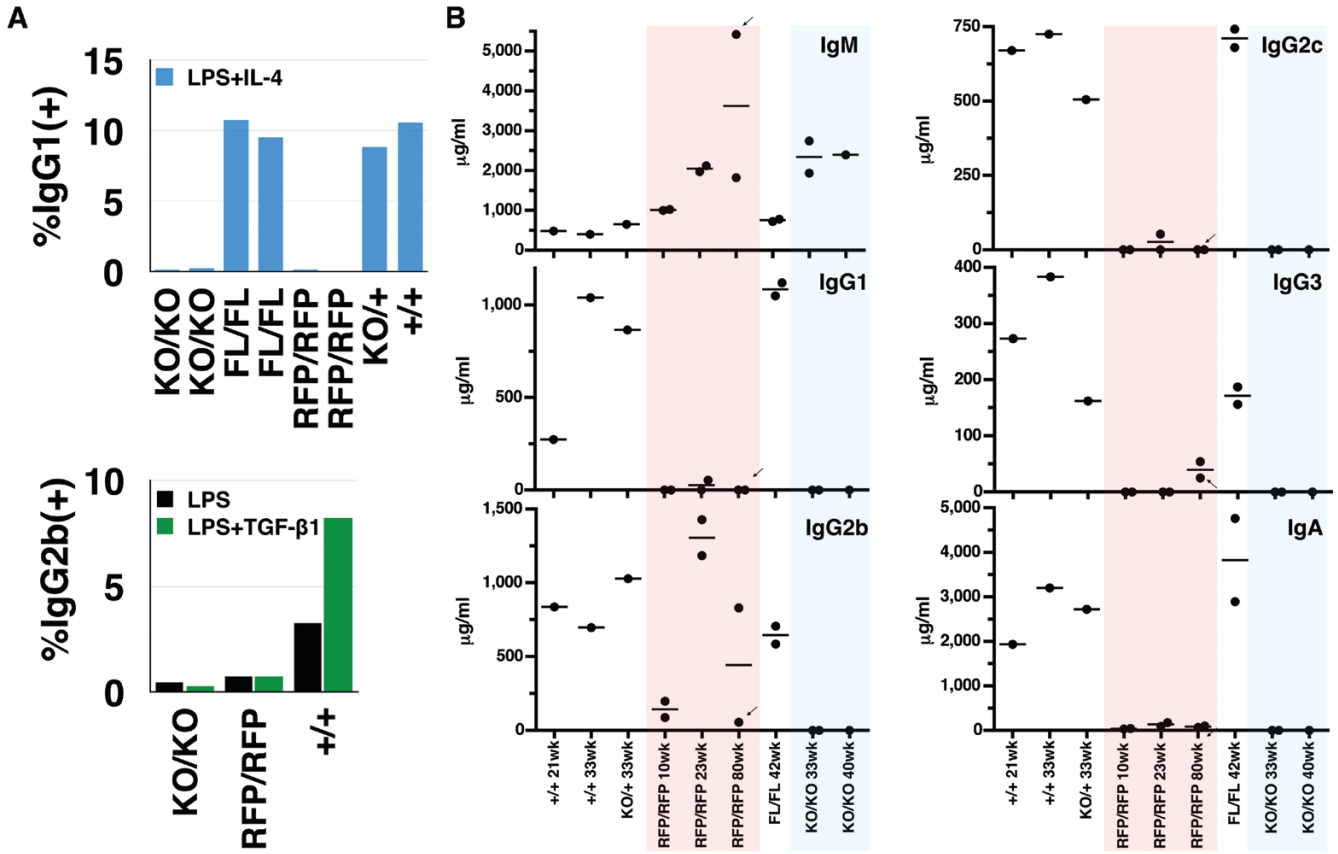


Fig. 5. CSR assay for Aicda-RFP, Aicda-FL and Aicda-KO homozygotes. (A) *In vitro* CSR was analyzed by culturing splenocytes from mice with the indicated genotype. After 3 days of stimulation with LPS and IL-4, cells were analyzed for surface IgG1 expression by flow cytometry. For IgG2b expression, splenocytes were stimulated with LPS (and TGF-β1) for 3 days. Results from individual mice are represented by independent columns. (B) Serum concentrations of immunoglobulins of the indicated isotypes were measured by ELISA. Results from individual mice of the indicated genotype and age in weeks are represented as dots with bars for the average. Dots indicated by arrows in the Aicda-RFP 80wk group are derived from the same individual.

homozygote showed an age-dependent increase of IgM concentration, which was higher than that in the control at ages older than 23 weeks. The concentrations of secondary isotypes were very low, except for IgG2b, which showed considerable variation among individuals.

Evaluation of SHM

To assess the degree of SHM in Aicda-RFP homozygotes, two Aicda-RFP mice at 23 weeks of age were sacrificed to collect Peyer's patches. GC B cells (PNA^{high}, $CD95^{+}$) within the collected Peyer's patches were sorted and genomic DNA was purified to analyze SHM in the DNA region downstream of the rearranged V genes (Fig. 6A). The forward primer anneals to multiple V genes of the J558 family, which are most frequently used in C57BL/6 strain, but not in Balb/c (19). The mutation distribution and frequency were represented by histograms and a bar chart, revealing a reduced frequency of SHM in Aicda-RFP homozygotes compared with that in the wild type, but a higher frequency than in Aicda-KO homozygotes (Fig. 6B and C). This indicated the production of a small amount of AID from the Aicda-RFP locus, which was detectable by RT-PCR but not by western blotting.

Wild-type B cells examined in parallel yielded 265 mutations, all of which are considered to be induced by the action of AID, as the Aicda-KO sample contained no mutations. Mutational signatures have been deduced from a catalog of somatic mutations in many cancer genomes, which are postulated to reflect the underlying mechanisms of mutagenesis. As a future reference, making use of these 265 mutations, the AID signature was depicted in a trinucleotide context as reported (Fig. 6D) (12, 13).

Discussion

AID is expressed in GC B cells in lymphoid organs and *in vitro*-stimulated B cells. Recent studies have indicated that AID is also expressed, albeit at lower levels, in leukemic cells and cancers of epithelial origin (4–6). Indeed, cell lines derived from cancers of the stomach, liver, lung and skin, and normal skin keratinocytes express AID in response to stimulation with TNF-α, IL-1β, TGF-β or LPS, all of which activate the transcription factor NF-κB (14, 20–25). Using transgenic mice that constitutively express AID in epithelial cells and AID-deficient mice, our and other laboratories provided supportive evidence that AID expression is involved in epithelial

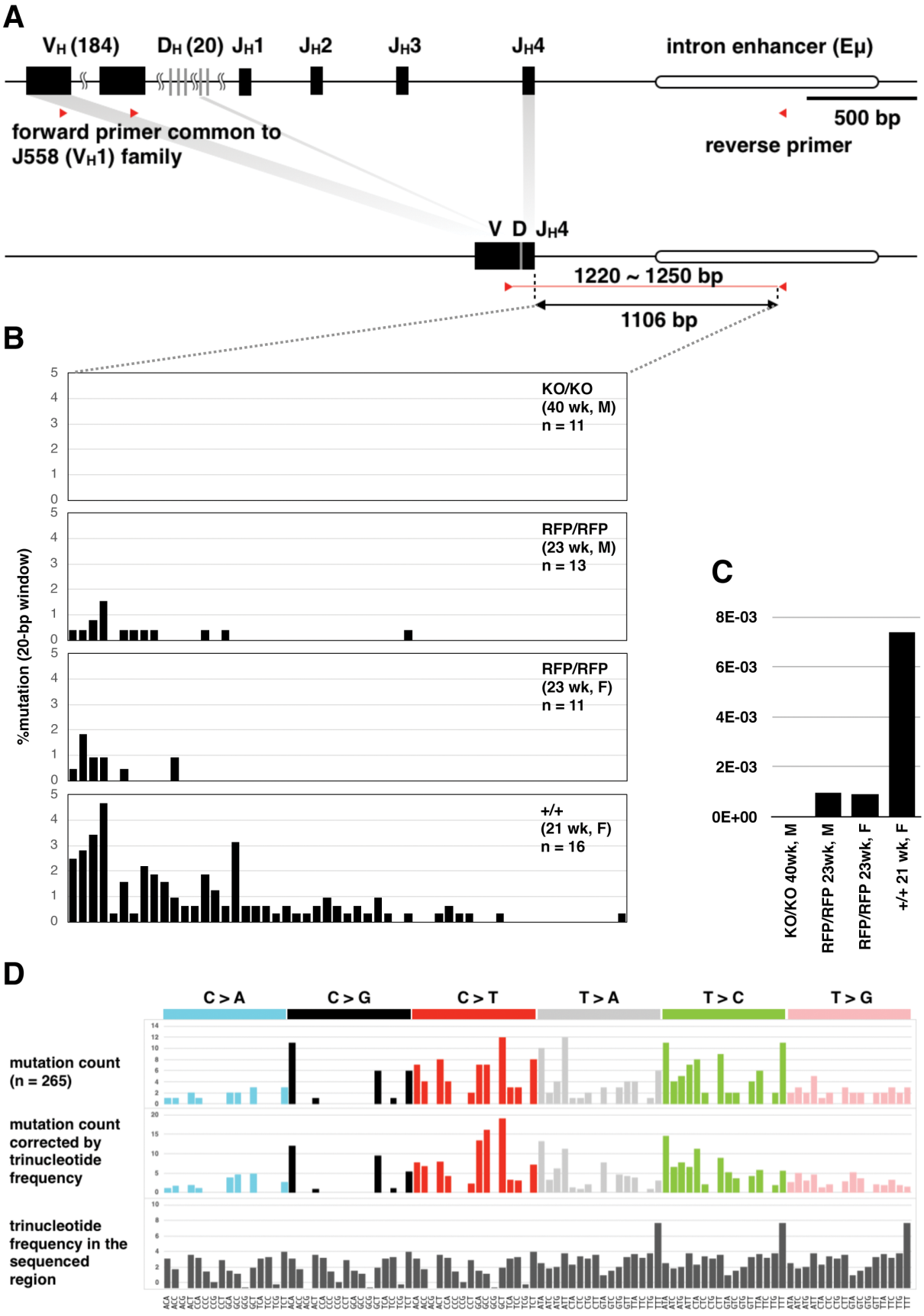


Fig. 6. SHM assay for Aicda-RFP and Aicda-KO homozygotes and wild-type mice. SHM in J_H4 downstream intron for GC B cells in Peyer's patches was analyzed. (A) Scheme of variable region genes of mouse Ig heavy chain gene before and after VDJ recombination. Positions of PCR primers are indicated (red triangles). (B) Distribution of mutations detected for mice with the indicated genotype. Each panel represents

cancer (14, 25–29). There are also reports suggesting the involvement of AID in human cancer (14, 21–24, 30–38). The use of AID-deficient mice to examine the role of AID during carcinogenesis is complicated by the fact that AID deficiency brings about immunological disturbance due to the lack of CSR and SHM of antibodies. Therefore, it is critical to evaluate conditional AID knockout mice that lack AID in specific cell compartments (e.g. epithelia).

In this study, we report the generation of conditional AID knockout mice (Aicda-FL) that can lose catalytically important exons 2 and 3 when tissue-specific Cre recombinase is provided. This Aicda-FL was derived from Aicda-RFP mice, which have an additional exon coding tdRFP. The conversion from Aicda-RFP (three *loxP*s) to Aicda-FL (two *loxP*s) requires partial Cre-mediated recombination between the first and second *loxP*s without involvement of the third *loxP*. This strategy called 'three-*loxP*' or 'tri-lox' is sometimes used to make conditional knockouts with floxed Neo removed. Partial recombination creating 2-*loxP*s from 3-*loxP*s requires moderate Cre activity in ES cells (39), or 16-cell-stage morulae (40). Alternatively, it is accomplished by mating a germline-expressing Cre transgenic strain of mice such as CMV-Cre, EIIa-Cre and MeuCre40. The frequency of desirable partial recombination in mice is often so difficult to predict that there was a case where more than 100 offspring had to be screened to obtain the correctly recombined allele (41–44). Currently, the FLP-*frt* recombination system is preferentially used to reliably remove the Neo cassette in the production of conditional knockout mice.

We used a male TNAP-Cre mouse (45) as a mating partner of a female Aicda-RFP homozygote, and successfully obtained a partially recombined allele, named Aicda-FL. Another partially recombined allele of Aicda-RFP-KO was also obtained. This indicates that TNAP-Cre mice can be used to convert 3-*loxP*s to the 2-*loxP*s configuration.

Aicda-KO homozygotes (complete recombination) showed exactly the same phenotype as previously reported AID-deficient mice (1), namely (i) lack of CSR, (ii) lack of SHM and (iii) enlarged GCs. Therefore, Aicda-FL mice must be a useful model to study the function of AID in specific compartments and within a specific period during development and immune reaction. Additional applications may include the study of AID in carcinogenesis, the epigenetic role of AID during gametogenesis, the restriction of AID during infection or sensitization to allergens and the role of AID in homeostasis of gut microbiota, where AID is most abundantly and constitutively expressed (3, 15, 46).

Aicda-RFP exhibited a hypomorphic allele, expressing about one-tenth of the wild-type level, judging from transcript abundance. Homozygous mice had IgG2b in the serum. The concentration appeared to increase with age from 10 to 80 weeks, showing substantial inter-individual variation. This is reminiscent of human hyper-IgM syndrome type 2 patients, where disproportional reduction of serum immunoglobulin isotypes with substantial inter-patient variation was reported even for individuals with identical hypomorphic AID mutations (47). In contrast to IgG1, the IgG2b switching of splenocytes from RFP-homozygotes stimulated with LPS (and TGF-β1) was faintly higher than the background level of the Aicda-KO homozygote, which may be related to occasional IgG2b accumulation in the serum of, especially aged, Aicda-RFP homozygotes. IgG1 switching might require a lower amount of AID than that of IgG2b.

SHM frequency in Aicda-RFP homozygotes was reduced to 13% of the wild-type level. Because the SHM assay adopted in this paper reflected *in vivo* selection of B cells for affinity-enhanced Igs, the actual SHM efficiency in Aicda-RFP homozygotes may be much lower than the calculated value. This may explain the discrepancy of observed degrees of defect between CSR and SHM in Aicda-RFP homozygotes.

Green fluorescent protein (GFP) knock-in into the *Aicda* locus was previously reported (48), in which GFP was fused to the carboxyl terminus of AID. Although this fusion protein may faithfully reflect the subcellular localization and amount of AID protein, the fluorescent intensity appears to be weak, resulting in poor separation from the negative cell population even with flow cytometry. In contrast, GC B cells from Aicda-RFP mice express a very bright tandem-dimeric form of RFP, which is easily detectable upon live imaging (Fig. 2) and suitable for quantitative study, as exemplified in Fig. 3(D).

AID belongs to the AID/APOBEC cytidine deaminase family, which comprises 11 genes in humans: *AICDA*, *APOBEC1*, *APOBEC2*, *APOBEC3A*, *APOBEC3B*, *APOBEC3C*, *APOBEC3D*, *APOBEC3F*, *APOBEC3G*, *APOBEC3H* and *APOBEC4*. Comprehensive analysis of mutations in human cancer deciphered elemental mutational patterns called 'mutational signatures', which may reflect underlying unique mutagenic mechanisms (12). They described the ABOBEC signature (Signature 2), which fits well with the observed mutations induced by APOBEC3A, APOBEC3B and APOBEC3C in yeast, where the 5′ base of mutated C was often T, while it was A or G for AID (49). To the best of authors' knowledge, AID-induced mutations in B lymphocytes have not been analyzed in the trinucleotide context, while a precise analysis in the penta-nucleotide context was reported (50). This study is the first to experimentally determine the trinucleotide AID signature in mammalian B cells (Fig. 6D). It was not concentrated on C-to-T transition at cytosines following A or G, but rather distributed to diverse trinucleotide contexts, despite some degree of preference. This pattern is most similar to Signature 3 (SBS3), which was reported to be associated with BRCA1/2 deficiency (12), or flat signatures such as SBS5 and SBS40 (13). In a study where signature analysis was performed for murine embryonic fibroblasts overexpressing AID, the trinucleotide signature differed from the one presented here, but focused on C-to-T transitions at cytosines following A or G (51), which is compatible with the yeast study (49). Recently, AID-induced mutations were extensively studied in

data from one mouse with the indicated genotype and age. Mutation frequencies for a 20-bp window are plotted. Numbers (*n*) of plasmid clones sequenced are indicated. (C) Mutation frequencies for the entire sequenced region (1106 bp) are plotted. (D) Mutational signature by AID. Frequencies of 96 trinucleotide patterns of single-base replacement for 265 mutations observed in the J_H4 downstream intron from the wild-type mice are plotted according to Alexandrov *et al.* (12, 13). Top and middle panels show mutation counts and those calibrated with trinucleotide frequency in the sequenced region (bottom), respectively.

conditions involving deficiency in base-excision repair (BER) and/or mismatch repair (MMR); mutations were also enriched at C-to-T transitions after A or G (52). This pattern may reflect the native AID preference before efficient repair by BER and MMR systems.

In summary, we generated RFP knock-in and conditional knockout mice from a single targeting event in ES cells, by a three-*loxP* strategy. Both strains of mice (Aicda-RFP and Aicda-FL) would be useful for functional analysis of AID in various body compartments and developmental stages.

Acknowledgements

We express our great gratitude to Andrea Nagy for providing the TNAP-Cre mice, Takeshi Yagi and Masanobu Oshima for the DTA gene cassette, and Jörg Fehling and Sidonia Fagarasan for the tdRFP mice. We thank Naoko Tomikawa for technical assistance. We also thank Hiroyuki Nishimura and Masamichi Muramatsu for critical reading of the manuscript and discussions.

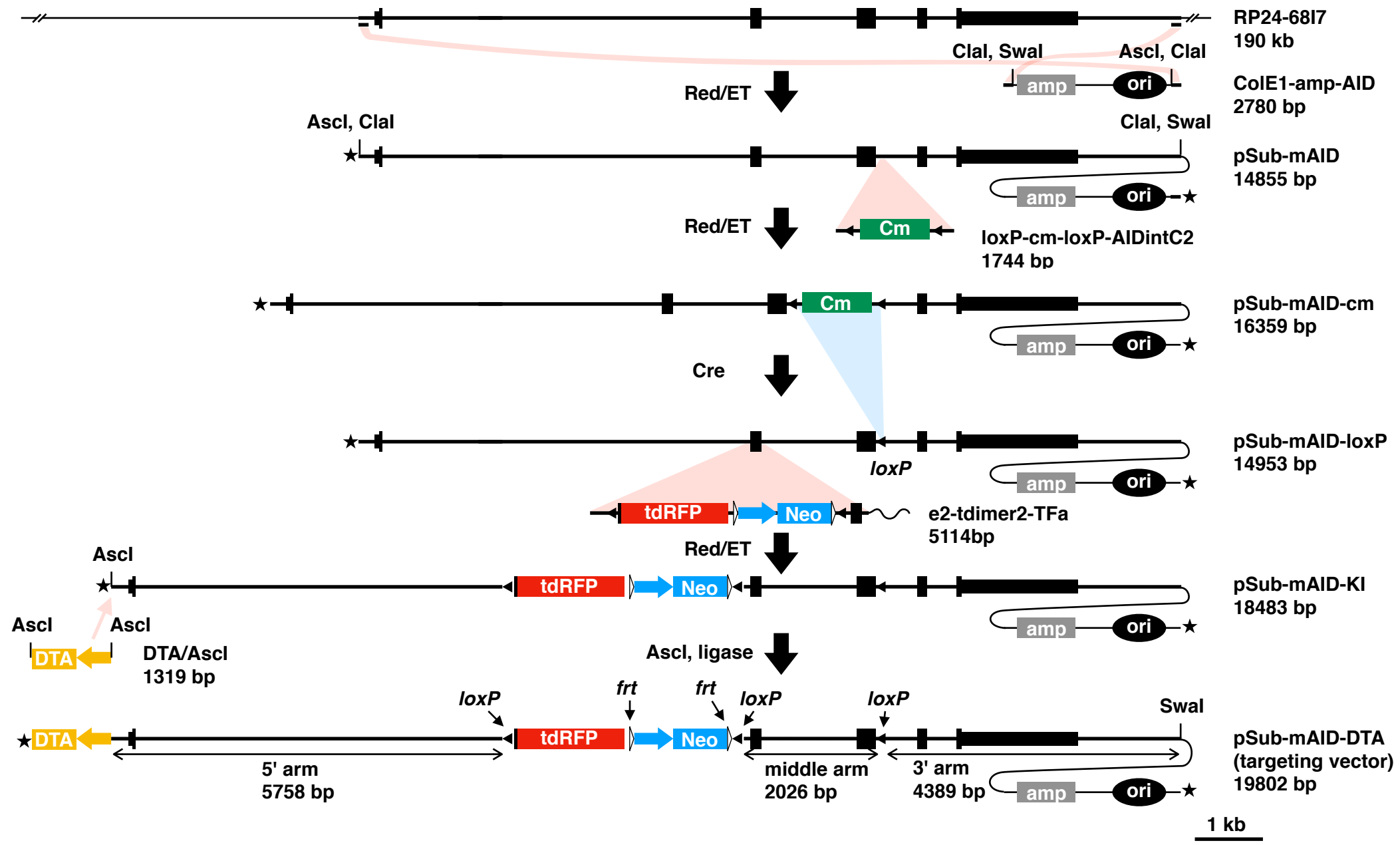
Conflicts of interest statement: the authors declared no conflicts of interest.

References

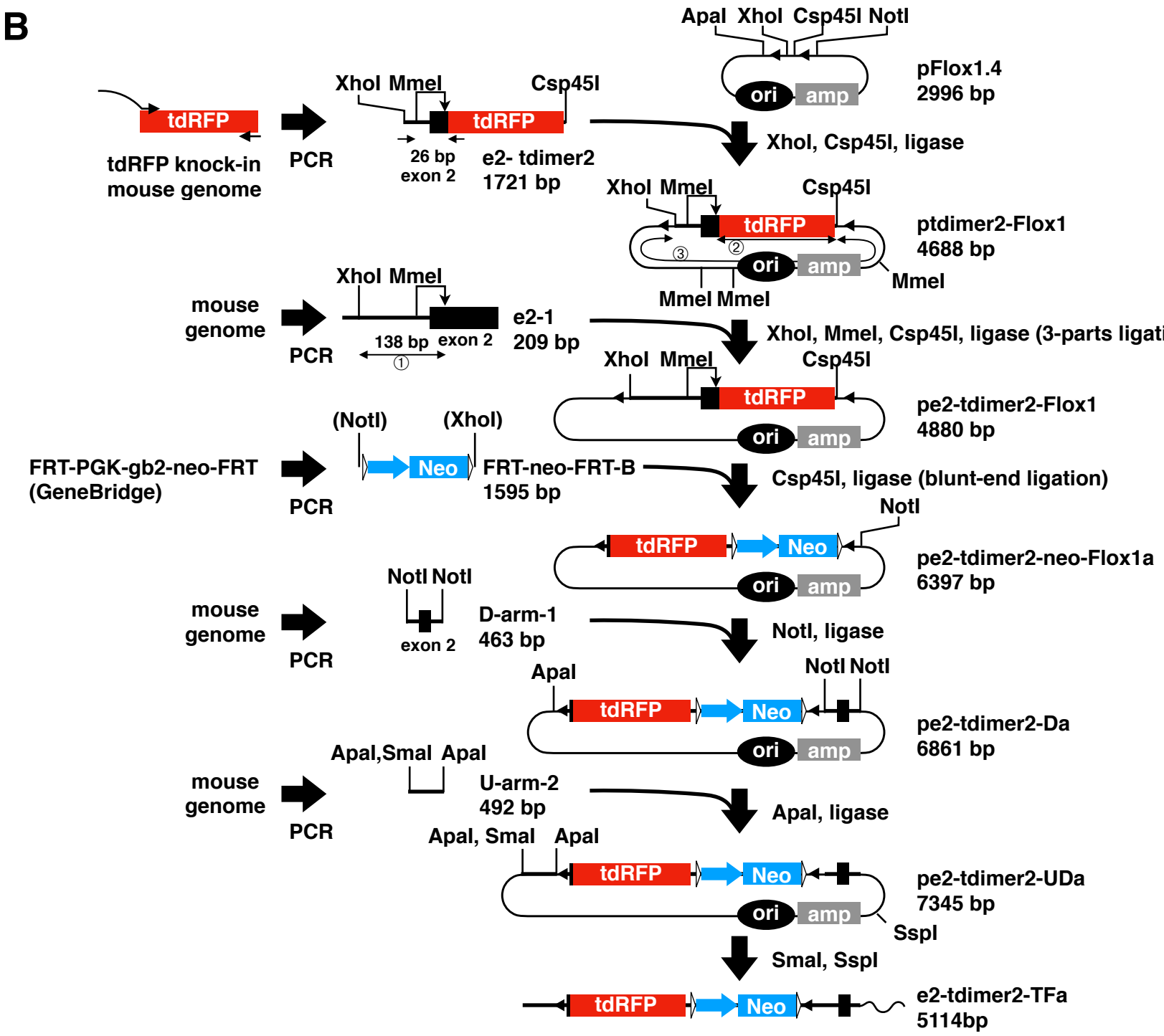
1 Muramatsu, M., Kinoshita, K., Fagarasan, S. *et al.* 2000. Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. *Cell* 102:553.
2 Revy, P., Muto, T., Levy, Y. *et al.* 2000. Activation-induced cytidine deaminase (AID) deficiency causes the autosomal recessive form of the hyper-IgM syndrome (HIGM2). *Cell* 102:565.
3 Muramatsu, M., Sankaranand, V. S., Anant, S. *et al.* 1999. Specific expression of activation-induced cytidine deaminase (AID), a novel member of the RNA-editing deaminase family in germinal center B cells. *J. Biol. Chem.* 274:18470.
4 Kinoshita, K. and Nonaka, T. 2006. The dark side of activation-induced cytidine deaminase: relationship with leukemia and beyond. *Int. J. Hematol.* 83:201.
5 Greaves, M. 2018. A causal mechanism for childhood acute lymphoblastic leukaemia. *Nat. Rev. Cancer* 18:471.
6 Marusawa, H., Takai, A. and Chiba, T. 2011. Role of activation-induced cytidine deaminase in inflammation-associated cancer development. *Adv. Immunol.* 111:109.
7 Honjo, T., Kobayashi, M., Begum, N. *et al.* 2012. The AID dilemma: infection, or cancer? *Adv. Cancer Res.* 113:1.
8 Kumar, R., DiMenna, L. J., Chaudhuri, J. *et al.* 2014. Biological function of activation-induced cytidine deaminase (AID). *Biomed. J.* 37:269.
9 Yagi, T., Ikawa, Y., Yoshida, K. *et al.* 1990. Homologous recombination at c-fyn locus of mouse embryonic stem cells with use of diphtheria toxin A-fragment gene in negative selection. *Proc. Natl Acad. Sci. USA* 87:9918.
10 Luche, H., Weber, O., Nageswara Rao, T. *et al.* 2007. Faithful activation of an extra-bright red fluorescent protein in "knock-in" Cre-reporter mice ideally suited for lineage tracing studies. *Eur. J. Immunol.* 37:43.
11 Jiang, C., Zhao, M. L. and Diaz, M. 2009. Activation-induced deaminase heterozygous MRL/lpr mice are delayed in the production of high-affinity pathogenic antibodies and in the development of lupus nephritis. *Immunology* 126:102.
12 Alexandrov, L. B., Nik-Zainal, S., Wedge, D. C. *et al.*; Australian Pancreatic Cancer Genome Initiative; ICGC Breast Cancer Consortium; ICGC MMML-Seq Consortium; ICGC PedBrain. 2013. Signatures of mutational processes in human cancer. *Nature* 500:415.
13 Alexandrov, L. B., Kim, J., Haradhvala, N. J. *et al.*; PCAWG Mutational Signatures Working Group; PCAWG Consortium. 2020. The repertoire of mutational signatures in human cancer. *Nature* 578:94.
14 Nonaka, T., Toda, Y., Hiai, H. *et al.* 2016. Involvement of activation-induced cytidine deaminase in skin cancer development. *J. Clin. Invest.* 126:1367.
15 Fagarasan, S., Muramatsu, M., Suzuki, K. *et al.* 2002. Critical roles of activation-induced cytidine deaminase in the homeostasis of gut flora. *Science* 298:1424.
16 Schrader, C. E., Linehan, E. K., Mochegova, S. N. *et al.* 2005. Inducible DNA breaks in Ig S regions are dependent on AID and UNG. *J. Exp. Med.* 202:561.
17 Noguchi, E., Shibasaki, M., Inudou, M. *et al.* 2001. Association between a new polymorphism in the activation-induced cytidine deaminase gene and atopic asthma and the regulation of total serum IgE levels. *J. Allergy Clin. Immunol.* 108:382.
18 Zaprazna, K., Reblova, K., Svobodova, V. *et al.* 2019. Activation-induced deaminase and its splice variants associate with trisomy 12 in chronic lymphocytic leukemia. *Ann. Hematol.* 98:423.
19 Collins, A. M., Wang, Y., Roskin, K. M. *et al.* 2015. The mouse antibody heavy chain repertoire is germline-focused and highly variable between inbred strains. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 370:20140236.
20 Kou, T., Marusawa, H., Kinoshita, K. *et al.* 2007. Expression of activation-induced cytidine deaminase in human hepatocytes during hepatocarcinogenesis. *Int. J. Cancer* 120:469.
21 Endo, Y., Marusawa, H., Kinoshita, K. *et al.* 2007. Expression of activation-induced cytidine deaminase in human hepatocytes via NF-kappaB signaling. *Oncogene* 26:5587.
22 Matsumoto, Y., Marusawa, H., Kinoshita, K. *et al.* 2007. *Helicobacter pylori* infection triggers aberrant expression of activation-induced cytidine deaminase in gastric epithelium. *Nat. Med.* 13:470.
23 Endo, Y., Marusawa, H., Kou, T. *et al.* 2008. Activation-induced cytidine deaminase links between inflammation and the development of colitis-associated colorectal cancers. *Gastroenterology* 135:889.
24 Komori, J., Marusawa, H., Machimoto, T. *et al.* 2008. Activation-induced cytidine deaminase links bile duct inflammation to human cholangiocarcinoma. *Hepatology* 47:888.
25 Kitamura, J., Uemura, M., Kurozumi, M. *et al.* 2015. Chronic lung injury by constitutive expression of activation-induced cytidine deaminase leads to focal mucous cell metaplasia and cancer. *PLoS ONE* 10:e0117986.
26 Okazaki, I. M., Hiai, H., Kakazu, N. *et al.* 2003. Constitutive expression of AID leads to tumorigenesis. *J. Exp. Med.* 197:1173.
27 Rucci, F., Cattaneo, L., Marrella, V. *et al.* 2006. Tissue-specific sensitivity to AID expression in transgenic mouse models. *Gene* 377:150.
28 Takai, A., Toyoshima, T., Uemura, M. *et al.* 2009. A novel mouse model of hepatocarcinogenesis triggered by AID causing deleterious p53 mutations. *Oncogene* 28:469.
29 Takai, A., Marusawa, H., Minaki, Y. *et al.* 2012. Targeting activation-induced cytidine deaminase prevents colon cancer development despite persistent colonic inflammation. *Oncogene* 31:1733.
30 Babbage, G., Ottensmeier, C. H., Blaydes, J. *et al.* 2006. Immunoglobulin heavy chain locus events and expression of activation-induced cytidine deaminase in epithelial breast cancer cell lines. *Cancer Res.* 66:3996.
31 Lin, C., Yang, L., Tanasa, B. *et al.* 2009. Nuclear receptor-induced chromosomal proximity and DNA breaks underlie specific translocations in cancer. *Cell* 139:1069.
32 Shinmura, K., Igarashi, H., Goto, M. *et al.* 2011. Aberrant expression and mutation-inducing activity of AID in human lung cancer. *Ann. Surg. Oncol.* 18:2084.
33 Nakanishi, Y., Kondo, S., Wakisaka, N. *et al.* 2013. Role of activation-induced cytidine deaminase in the development of oral squamous cell carcinoma. *PLoS ONE* 8:e62066.
34 Okura, R., Yoshioka, H., Yoshioka, M. *et al.* 2014. Expression of AID in malignant melanoma with BRAF(V600E) mutation. *Exp. Dermatol.* 23:347.
35 El Kadi, N., Wang, L., Davis, A. *et al.* 2018. The EGFR T790M mutation is acquired through AICDA-mediated deamination of 5-methylcytosine following TKI treatment in lung cancer. *Cancer Res.* 78:6728.

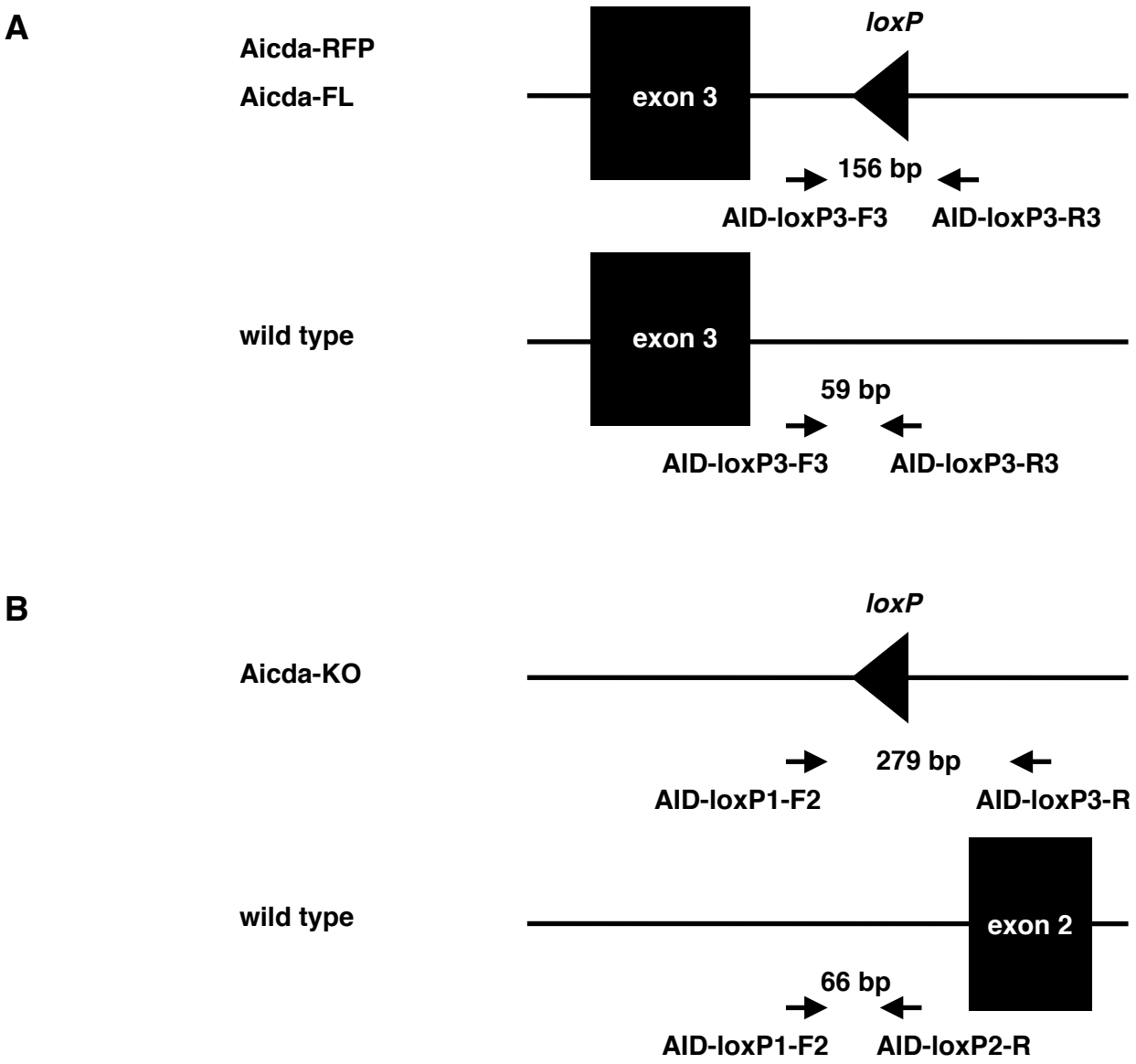
36 Li, H., Li, Q., Ma, Z. *et al*. 2019. AID modulates carcinogenesis network via DNA demethylation in bladder urothelial cell carcinoma. *Cell Death Dis*. 10:251.
37 Li, L., Su, N., Cui, M. *et al*. 2019. Activation-induced cytidine deaminase expression in colorectal cancer. *Int. J. Clin. Exp. Pathol*. 12:4119.
38 Che, Z., Fan, J., Zhou, Z. *et al*. 2020. Activation-induced cytidine deaminase expression facilitates the malignant phenotype and epithelial-to-mesenchymal transition in clear cell renal cell carcinoma. *DNA Cell Biol*. 39:1299.
39 Nagy, A. 2000. Cre recombinase: the universal reagent for genome tailoring. *Genesis* 26:99.
40 Kaartinen, V. and Nagy, A. 2001. Removal of the floxed neo gene from a conditional knockout allele by the adenoviral Cre recombinase *in vivo*. *Genesis* 31:126.
41 Wiebel, F. F., Rennekampff, V., Vintersten, K. *et al*. 2002. Generation of mice carrying conditional knockout alleles for the transcription factor SRF. *Genesis* 32:124.
42 Holzenberger, M., Lenzner, C., Leneuve, P. *et al*. 2000. Cre-mediated germline mosaicism: a method allowing rapid generation of several alleles of a target gene. *Nucleic Acids Res*. 28:E92.
43 Xu, X., Li, C., Garrett-Beal, L. *et al*. 2001. Direct removal in the mouse of a floxed neo gene from a three-loxP conditional knockout allele by two novel approaches. *Genesis* 30:1.
44 Leneuve, P., Colnot, S., Hamard, G. *et al*. 2003. Cre-mediated germline mosaicism: a new transgenic mouse for the selective removal of residual markers from tri-lox conditional alleles. *Nucleic Acids Res*. 31:e21.
45 Lomelí, H., Ramos-Mejía, V., Gertsenstein, M. *et al*. 2000. Targeted insertion of Cre recombinase into the TNAP gene: excision in primordial germ cells. *Genesis* 26:116.
46 Fagarasan, S., Kinoshita, K., Muramatsu, M. *et al*. 2001. *In situ* class switching and differentiation to IgA-producing cells in the gut lamina propria. *Nature* 413:639.
47 Quartier, P., Bustamante, J., Sanal, O. *et al*. 2004. Clinical, immunologic and genetic analysis of 29 patients with autosomal recessive hyper-IgM syndrome due to activation-induced cytidine deaminase deficiency. *Clin. Immunol*. 110:22.
48 Crouch, E. E., Li, Z., Takizawa, M. *et al*. 2007. Regulation of AID expression in the immune response. *J. Exp. Med*. 204:1145.
49 Taylor, B. J., Nik-Zainal, S., Wu, Y. L. *et al*. 2013. DNA deaminases induce break-associated mutation showers with implication of APOBEC3B and 3A in breast cancer kataegis. *Elife* 2:e00534.
50 Yaari, G., Vander Heiden, J. A., Uduman, M. *et al*. 2013. Models of somatic hypermutation targeting and substitution based on synonymous mutations from high-throughput immunoglobulin sequencing data. *Front. Immunol*. 4:358.
51 Olivier, M., Weninger, A., Ardin, M. *et al*. 2014. Modelling mutational landscapes of human cancers *in vitro*. *Sci. Rep*. 4:4482.
52 Álvarez-Prado, Á. F., Pérez-Durán, P., Pérez-García, A. *et al*. 2018. A broad atlas of somatic hypermutation allows prediction of activation-induced deaminase targets. *J. Exp. Med*. 215:761.

A

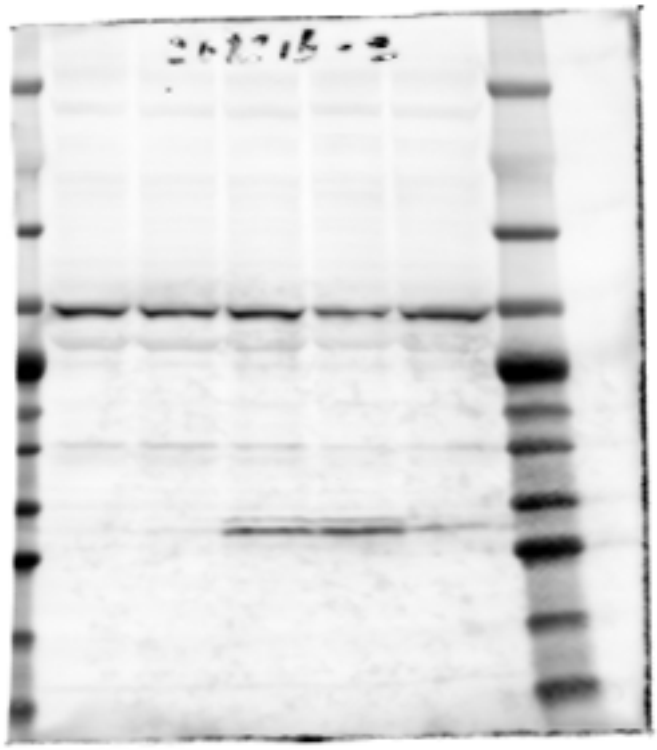

RP24-68I7
190 kb
ClaI, SwaI
AscI, ClaI
amp
ori
ColE1-amp-AID
2780 bp
Red/ET
AscI, ClaI
ClaI, SwaI
pSub-mAID
14855 bp
Red/ET
Cm
loxP-cm-loxP-AIDintC2
1744 bp
Cm
pSub-mAID-cm
16359 bp
Cre
loxP
pSub-mAID-loxP
14953 bp
tdRFP
Neo
e2-tdimer2-TFa
5114bp
Red/ET
pSub-mAID-KI
18483 bp
AscI
AscI
AscI
DTA
DTA/AscI
1319 bp
AscI, ligase
loxP
frt
frt
loxP
loxP
SwaI
DTA
5' arm
5758 bp
middle arm
2026 bp
3' arm
4389 bp
pSub-mAID-DTA
(targeting vector)
19802 bp
1 kb

B

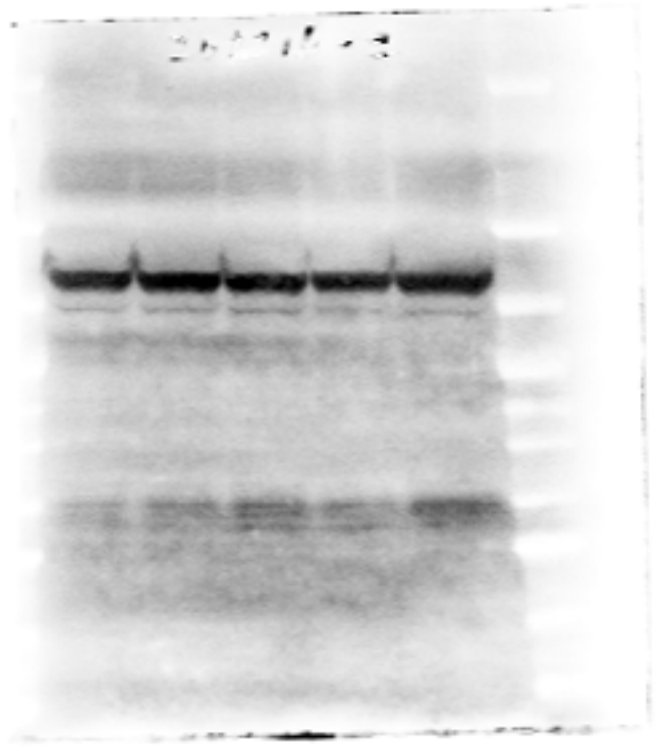

ApaI XhoI Csp45I NotI
pFlox1.4
2996 bp
ori
amp
tdRFP knock-in
mouse genome
PCR
XhoI MmeI
Csp45I
tdRFP
26 bp
exon 2
e2- tdimer2
1721 bp
XhoI, Csp45I, ligase
XhoI MmeI
Csp45I
tdRFP
ptdimer2-Flox1
4688 bp
MmeI
MmeI MmeI
mouse
genome
PCR
XhoI MmeI
138 bp exon 2
e2-1
209 bp
XhoI, MmeI, Csp45I, ligase (3-parts ligation)
XhoI MmeI
Csp45I
tdRFP
pe2-tdimer2-Flox1
4880 bp
(NotI)
(XhoI)
FRT-PGK-gb2-neo-FRT
(GeneBridge)
PCR
Neo
FRT-neo-FRT-B
1595 bp
Csp45I, ligase (blunt-end ligation)
NotI
tdRFP
Neo
pe2-tdimer2-neo-Flox1a
6397 bp
NotI NotI
mouse
genome
PCR
exon 2
D-arm-1
463 bp
NotI, ligase
ApaI
NotI NotI
tdRFP
Neo
pe2-tdimer2-Da
6861 bp
ApaI,SmaI ApaI
mouse
genome
PCR
U-arm-2
492 bp
ApaI, ligase
ApaI, SmaI ApaI
tdRFP
Neo
pe2-tdimer2-UDa
7345 bp
SspI
SmaI, SspI
tdRFP
Neo
e2-tdimer2-TFa
5114bp


A
Aicda-RFP
Aicda-FL
loxP
exon 3
156 bp
AID-loxP3-F3
AID-loxP3-R3
wild type
exon 3
59 bp
AID-loxP3-F3
AID-loxP3-R3
B
Aicda-KO
loxP
279 bp
AID-loxP1-F2
AID-loxP3-R
wild type
exon 2
66 bp
AID-loxP1-F2
AID-loxP2-R

whole blot image for MAID-2 (AID)



whole blot image for TU27 (β-tublin)



primer name	sequence	note
AID-D-1F	TGTTGTGCACTGTTGTTGCACAGTGAGCCATCTTGCCAGCTC ATATTCCATCGATATTTAAATTTAATTAAGCTCTCCTGAGTA GGACAAATC	Red and back portions match Aicda gene and ColE1-amp sequence, respectively. Light blue contains restriction sites for cloning purpose.
AID-U-1R	GCAGTTCTGCTCATGCTTAATGTCCAACTTCCTTCGAGGCTGT TCCTCTAATCGATGGCGCGCCGTTTAAACTCACAGCTTGTCT GTAAGCGGATG	Red and back portions match Aicda gene and ColE1-amp sequence, respectively. Light blue contains restriction sites for cloning purpose.
AID-intC-1F	ACACTGGAGAGAGCGGTCTGAGTTGCCACTCAGAGTGAGTGTC AGCGGGGCTGCAGAATTAACCCTCACTAAAGGGCG	Red and back portions match Aicda gene and loxP-gb2-cm-loxP sequence, respectively.
AID-intC-1R	GATATCCAGGACGAACTGAAGGTCTTTAAGTAGCACCCCACCC CCAGTTTAAGCTTAATACGACTCACTATAGGGCTC	Red and back portions match Aicda gene and loxP-gb2-cm-loxP sequence, respectively.
AID-intC-2F	TGAGGGGCTGCGGAGACTGCACCGCGCTGGGGTCCAGATCGG GATCATGACCTTCAAAGGTGAGACTTGCACACTGGAGA	Red and orange portions match Aicda gene and end of the 1st PCR product, respectively.
AID-intC-2R	CGGAGAATTGCAAGTTGAAGGTCAGCCTGGGATAGCGAGATCC TACCTCAAAAAATAAAGTAATACTTTTGATATCCAGG	Red and orange portions match Aicda gene and end of the 1st PCR product, respectively.
e2-tdimer-3F	ACTCTCGAGTCCAACTCACAGCCTTCTGATGAAGCAAGCCTCCT	Red and black portion matches Aicda and tdRFP gene, respectively. Light blue contains restriction sites for cloning purpose.
e2-tdimer-1R	ACCTTCGAAGAAGTTATACGCGTTAAGATACA	Black portion matches tdRFP gene. Light blue contains restriction sites for cloning purpose.
e2-1F	TGCCTCGAGCGCTCAGCTACCTTGCCTATG	Red portion matches Aicda gene. Light blue contains restriction sites for cloning purpose.
e2-1R	GTAGGTCTCATGCCGTC	Red portion matches Aicda gene.
neo-FRT-1F	CCATTCGAAAATTAACCCTCACTAAAGGGCG	Black portion matches FRT-PGK-gb2-neo-FRT. Light blue contains restriction sites.
neo-FRT-1R	TGGTTCGAATAATACGACTCACTATAGGGCTC	Black portion matches FRT-PGK-gb2-neo-FRT. Light blue contains restriction sites.
D-arm-1F	CGGGCGGCCGCCGCTCAGCTACCTTGCCTATG	Red portion matches Aicda gene. Light blue contains restriction sites for cloning purpose.
D-arm-1R	GTTGCGGCCGCTGTTTGTTTTTCCAGACAGGG	Red portion matches Aicda gene. Light blue contains restriction sites for cloning purpose.
U-arm-2F	AATGGGCCCGGGCAGAAGGTTAAAGGTGTC	Red portion matches Aicda gene. Light blue contains restriction sites for cloning purpose.
U-arm-3R	GCTGGGCCCGATATCTTCAGCACCGAGCTGCATCC	Red portion matches Aicda gene. Light blue contains restriction sites for cloning purpose.
MCD-R	CTTCCTCGTGCTTTA CGGTATC	
R2	CCCCTTCAAACGCTGTAATTAACATC	
5'N2-F	AAAGAGGTCGTAGCAACTCCTTACT	
5'N2-R	ATATTCTAAACCTCATAGCCAAACG	
3'N2-F	ATGGCTATGAAATGTTGAAAGTAGC	
3'N2-R	AGAGTAGATTTCACGCTCTGAAGAC	
NeoP-F	GAACAAGATGGATTGCACGCAGGTTCTCCG	
NeoP-R	GTAGCCAACGCTATGTCCTGATAG	
AID-loxP3-F3	ACTCAGAGTGAGTGTCAGCG	genotyping
AID-loxP3-R3	ACTGAAGGTCTTTAAGTGCACCC	genotyping
AID-loxP1-F2	CATGATGATCAAAGTAGGCCTG	genotyping
AID-loxP2-R	CCATAGGCAAGGTAGCTGA	genotyping
AID-loxP3-R	ATACTTGTAATTCCGCCACTCAGAA	genotyping
dimer2-120F	CACCCAGACCGCCAAGCTG	genotyping
dimer2-211R	CGGGGTGCTTCACGTACGC	genotyping
SV40PolyA3F	AACTCATCAATGTATCTTAACGC	genotyping
VHJ558-F	GCCTGACATCTGAGGACTCTGC	Immunology 126:102, 2008
mIgHJ4down-2R	CCTCTCCAGTTTCGGCTGAATCC	Immunology 126:102, 2008
mAIDm58F	GCAAGCTTCCTTTGGCCTA	mAID exon 1 forward
mAID82R	TTCACCACGTAGCAGAGGTA	mAID exon 2 reverse
mAID93F	CGTGGTGAAGAGGAGAGATAGTG	mAID-TM1 (mAID exon 2 forward)
mAID203R	CAGTCTGAGATGTAGCGTAGGAA	mAID-TM1 (mAID exon 3 reverse)
mAID117F-P	CACCTCCTGCTCACTGGACTTCGGC	mAID-TM1: FAM(5') and TAMRA(3') (mAID exon 2 forward)
mAID523F	ATTTTTACTGCTGGAATACATTT	mAID-TM2 (mAID exon 4 forward)
mAID690R	ACATTCCAGGAGGTTGCTT	mAID-TM2 (mAID exon 5 reverse)
mAID601F-P	TCCGGCTAACCAGACAACTTCGGCGCAT	mAID-TM2: FAM(5') and TAMRA(3') (mAID exon 5 forward)
mAID162F	CTGCCACGTGGAATTGTTG	mAID-TM4 (mAID exon 3 forward)
mAID337R	GCGGTCTTCACAGAAGTAGAG	mAID-TM4 (mAID exon 4 reverse)
mAID258R-P	GCCACGTGCCGGGCACAGTCATAGCAC	mAID-TM4: FAM(5') and BHQ1(3') (mAID exon 3 reverse)
mAID-e1-2F	GCAAGCTTCCTTTGGCCTAAG	RT-PCR for RFP transcripts
dimer2-67R	GAACTCGTGGCCGTTCAC	RT-PCR for RFP transcripts
18S-2F	CCTGAGAAACGGCTACCACAT	RT-PCR for 18S ribosomal RNA
18S-2R	AATTACAGGGCCTCGAA	RT-PCR for 18S ribosomal RNA

antigen	label	isotype	company	code	usage
mouse IgG1	APC	rat IgG1	Becton Dickinson	560089	flow cytometry
mouse IgG2b	APC	goat Igs	Southern Biotechnology	1090-11	flow cytometry
mouse CD95	FITC	hamster IgG2	Becton Dickinson	554257	flow cytometry
N/A		7-AAD	Biolegend	420403	flow cytometry
N/A	Alexa647	PNA	Invitrogen	L-32460	flow cytometry
mouse IgM		goat Igs	Southern Biotechnology	1020-01	ELISA capture
mouse IgG1		goat Igs	Southern Biotechnology	1070-01	ELISA capture
mouse IgG2b		goat Igs	Southern Biotechnology	1090-01	ELISA capture
mouse IgG3		goat Igs	Southern Biotechnology	1100-01	ELISA capture
mouse IgA		goat Igs	Southern Biotechnology	1040-01	ELISA capture
mouse IgM	ALP	goat Igs	Southern Biotechnology	1020-04	ELISA detection
mouse IgG1	ALP	goat Igs	Southern Biotechnology	1070-04	ELISA detection
mouse IgG2b	ALP	goat Igs	Southern Biotechnology	1090-04	ELISA detection
mouse IgG3	ALP	goat Igs	Southern Biotechnology	1100-04	ELISA detection
mouse IgA	ALP	goat Igs	Invitrogen	62-6722	ELISA detection
N/A		mouse IgM	Sigma-Aldrich	M3795	ELISA standard
N/A		mouse IgG1	Sigma-Aldrich	M9035	ELISA standard
N/A		mouse IgG2b	Sigma-Aldrich	M8894	ELISA standard
N/A		mouse IgG3	Sigma-Aldrich	I3784	ELISA standard
N/A		mouse IgA	Sigma-Aldrich	M1421	ELISA standard
mouse AID		rat IgG2a	my laboratory	MAID-2	western blotting
beta tubulin		mouse IgG1	Biolegend	903401	western blotting
rat IgG	Alexa680	goat Igs	Invitrogen	A21096	western blotting
mouse IgG	IRDye800CW	goat Igs	Li-cor	926-32210	western blotting