

Title: *In vivo* analysis of *Aicda* gene regulation: critical balance between upstream enhancer and intronic repressor for the appropriate expression.

Authors: Le Thi Huong^{1,5}, Maki Kobayashi¹, Mikiyo Nakata¹, Go Shioi³, Hitoshi Miyachi⁴, Tasuku Honjo¹ and Hitoshi Nagaoka²

¹Department of Immunology and Genomic Medicine, Graduate School of Medicine, Kyoto University, Yoshida Sakyo-ku, Kyoto 606-8501, Japan

²Department of Molecular Pathobiochemistry, Gifu University Graduate School of Medicine, 1-1 Yanagido, Gifu 501-1194, Japan.

³Laboratory for Animal Resources and Genetic Engineering, RIKEN Center for Developmental Biology, 2-2-3 Minatojima Minami, Chuou-ku, Kobe 620-0047, Japan

⁴Reproductive Engineering Team, Institute of Virus Research, Kyoto University, Kyoto 606-8507, Japan

⁵Present address: Department of Clinical Pharmacy, Thai Nguyen University of Medicine and Pharmacy, 284 Luong Ngoc Quyen, Thai Nguyen, Viet Nam.

Corresponding Author: Hitoshi Nagaoka

e-mail: gumpb@gifu-u.ac.jp

Phone: +81-58-230-6206

Fax: +81-58-230-6209

Abstract

The *Aicda* gene, encoding activation-induced cytidine deaminase (AID) is strongly transcribed in activated B cells to diversify immunoglobulin genes, yet low-level expressions upon physiological/pathological stimuli in various cells have been reported. Mutagenic nature of AID has shown to be involved in tumor development. Here, by using a transgenic strategy with bacterial artificial chromosomes (BAC), we examined the *in vivo* function of *Aicda* regulatory elements which cluster in two regions, namely approximately 8-kb upstream of the transcription start site (region 4) and the first intron (region 2). Deletion of either of regions completely abolished the expression of *Aicda*-BAC-reporters, indicating critical roles of these elements. Furthermore, we found that the selective deletion of two C/EBP binding sites in region 4 inactivates its enhancer activity in spite of the presence of the intact NF- κ B-, STAT6- and Smad-binding-sites. On the other hand, the selective deletion of E2F- and c-Myb-binding-sites in region 2 increased the frequency of the cells with active *Aicda* promoter in germinal center B cells, indicating that E2F and c-Myb function as silencer *in vivo*. Interestingly, the silencer deletion did not cause the ectopic activation of *Aicda* promoter, indicating that specific stimulation of the enhancer is required for *Aicda* activation. In summary, the precise regulation of the *Aicda* promoter appears to dependent on a coordinated balance of activities between the enhancer and silencer elements.

Introduction

Activation-induced cytidine deaminase (AID) encoded by *Aicda* is the molecule central to the antigen-induced immunoglobulin (Ig) diversification, i.e. somatic hypermutation (SHM) and class switch recombination (CSR) in activated B cells [1,2]. AID triggers induction of nicking cleavage at the *Ig* loci [3,4]. Although the target specificity of AID is limited to *Ig* genes, in which SHM and CSR take place, AID can target many other genes including proto-oncogenes with much lower frequencies [5-7]. Such unphysiological off-target attack by AID was suggested to be involved in tumorigenesis of not only B cells but also other cell lineages [8]. In support of this notion, artificial over-expression of AID in transgenic mice caused tumors in non-B cells such as T lymphoma, lung tumor and hepatoma. [9,10]. Besides, *Aicda* knockout resulted in a drastic delay of tumor development in animal models of plasmacytoma, associated with the suppression of *Ig-Myc* translocation [11,12]. AID is thus suspected to be an endogenous mutagenic enzyme.

Because of this self-mutating activity of AID, it was expected that the *Aicda* regulation should be strict enough to avoid any leaky expression. Actually, a number of reports demonstrated that the high-level expression of AID is virtually restricted to germinal center B cells *in vivo* [13-15]. But accumulating evidence indicates that AID can be expressed in cells other than activated B cells when the cells are exposed to strong stimulation. Particularly, induction of AID expression in gastric epithelial cells by *Helicobacter pylori* infection has provided insights into mechanisms of

pathogen-associated tumorigenesis. Namely, AID might play a role of the mutagen in the inflammation-associated tumor development [16]. In accord with this idea, several reports showed that some other tumorigenic pathogens, such as EB virus, hepatitis C virus and HTLV-1, could induce/enhance AID expression [17-20]. The NF- κ B pathway appears to be involved in AID induction in *H. pylori* infected gastric epithelial cells, suggesting that inflammatory signals associated with infection of the pathogen may be involved in AID induction [16].

Under the physiological conditions AID is also expressed in immature B cells and some activated T cells at the level of one to two orders of magnitude lower than that of germinal center B cells although the roles of such low amounts of AID are unclear [14,15,21,22]. *Aicda* is transiently expressed in a fraction of activated T cells, which contain an IL-10 producing subset of T cells, suggesting that *Aicda* inductive stimuli correlate with the development of this T cell subset [15]. In the case of immature B cells, AID was suggested to be involved in removing B cells with self-reactive Igs on their surface [23,24]. These results suggest that a marginal level of AID might have some additional physiological roles other than Ig diversification although the precise nature of such function is still largely unclear. Taken together, AID can be expressed at various levels in different lineages and conditions, which might be involved in the physiological or pathological cellular responses.

Four well-conserved regions (region 1, 2, 3 and 4) in and around *Aicda* were found by the genome sequence conservation between human and mouse [14,25,26]. The importance of the downstream region (region 3) was demonstrated using *Aicda*

bacterial-artificial-chromosome (BAC) constructs although the regulatory function of region 3 was not clear by the *in vitro* reporter assay [14,26]. The several studies for understanding *Aicda* regulation have shown that the *Aicda* promoter (region 1) is, in general, not particularly specific to B cells, and promotes transcription in various types of cell [25,26]. The HoxC4 binding site located in the promoter region was reported to contribute the induction of *Aicda* transcription by increasing the level of HoxC4 [27,28]. Besides those regulatory elements proximal to the promoter, *Aicda* contains major regulatory elements in region 2 and region 4 [25,26]. Region 2 located in the first intron contains the binding sites for B cell specific transcription factors, such as Pax5 and E proteins [25,26,29]. Moreover, E2F and c-Myb binding sites in region 2 are shown to have a strong silencing activity to the promoter. Region 4 located about 8-kb upstream of the transcription initiation site contains stimulation-responsive elements for STAT6, NF- κ B, Smad and C/EBP. The *in vitro* reporter-assay of mutations in these elements led us to propose a balanced regulation model of the *Aicda* promoter, in which the coordinated action of B-cell-specific and stimulation-responding enhancers counteracts the silencers to derepress the *Aicda* promoter.

To understand the physiological mechanisms for *Aicda* regulation, it is important to verify the *in vivo* activity of these regulatory elements identified by *in vitro* experiments. For this purpose, we used BAC DNA to generate several lines of transgenic mice and evaluated *in vivo* function of *Aicda* regulatory elements designated by the previous *in vitro* reporter assay [26]. The present study clearly demonstrates that *Aicda* is regulated *in vivo* by the balance between the positive and negative transcriptional factors which we

identified in the *in vitro* experiments.

Methods

Generation of Transgenic Mice.

To generate BAC transgenic mice, we obtained a BAC clone (RP24-68I7, which is 190-kb in length and harboring the entire *Aicda* gene from BACPACCHORI (<http://bacpac.chori.org/>) [30]. To generate the Aid-cre-cd2 construct an expression cassette of cre IRES hCD2 whose intra cellular domain were deleted, was inserted into the exon 2 of mouse *Aicda* on the BAC by using homologous recombination in bacteria as described previously [30] (Fig. 1A). Transgenic constructs with series of deletions (Fig.1A) were derived from the Aid-cre-cd2 construct. All BAC modification was done by using (Red/ET recombination kit, Gene Bridges, Germany) according to the manufacture's instruction. In brief, we constructed vectors carrying two homology-arms (either 80-bp synthesized oligonucleotide or approximately 0.5-kb PCR product) adjacent to the DNA region that should be deleted/replaced. The homology-arms are ligated with donor DNA containing the PGK-neo cassettes floxed by Frt sequences. DNA fragments for targeting were prepared by PCR, then transfected to *E. coli* carrying the BAC with pRedET plasmid. Recombinants were selected by kanamycin, and then PGK-neo cassette was removed by transfecting the FLPe plasmid. When further modification was introduced into Aid-cre-cd2 BAC construct, to avoid unwanted deletion between the first Frt site and newly introduced Frts, Frt in the second targeting-construct was placed in reverse orientation to avoid deleting off of the intervening region; then the clones with reversion were excluded after the transfection of FLPe plasmid.

The primers used to amplify the homology arms were as follows:

The dR2-ac-cre transgene:

5'arm-dR2-Fwd: 5'-ATATGCAGCTCATGATGAGCCA-3'

5'arm-dR2-Rv: 5'-AGGGGTGAGGGAGAGTGGGAGG-3'

3'arm-dR2-Fwd: 5'-GTGCTGGATGGAGGTGCCCAGT-3'

3'arm-dR2-Rv: 5'-GTTATAATAATAAGCACAGGTA-3'

The dCEBP-ac-cre transgene:

5'arm-dCEBP-Fwd: 5'-GAGCGTAGAGGTCAGTGGACAA-3'

5'arm-dCEBP-Rv: 5'-CAACTTCTGCATTCCCCGATTTT-3'

3'arm-dCEBP-Fwd: 5'-AGCTGGTGTCCAAAATAGTGA-3'

3'arm-dCEBP-Rv: 5'-CTCTTACAGATGCCGTACACAT-3'

The dR4-ac-cre transgene:

5' homology arm (80bp):

5'-TCACCACACTCAGTTTGAAGCCTTTTAATTTTCATAGTATTCCATTTGAAAA
TTCTTTTTTCTCACTATAGGGCTCGAGGAA-3'

3' homology arm (80bp):

5'-CCACACCCTCTCTTACAGATGCCGTACACATGCCATAGGAACAAATA
CAAAACCCTTGAGAATTCGATATACGAAGTTAT-3'

The dME-ac-cre transgene was generated with two steps. First, we amplified the fragment containing the binding motifs of c-Myb with primers 5'arm-ME-Fwd: 5'-CTTTTCCTTCTCCTCCTTCTCC-3' and 5'arm-ME-Rv: CCTCTCCTCGGGATGTCCCTCC, then cloned it to pGEM-T easy vector (Promega).

The whole plasmid was PCR-amplified by primers, myb-mut-F

5'-NNGGATCCACATCCTGAGCCCTCAAAAAGCA-3' and myb-mut-R
5'-NNGGATCCTTGTCTAGCATGTGTGAGGTCTTC-3'. The PCR product was
digested by *Bam*HI and self-ligated. The other homology-arm that is adjacent region of
the E2F site was amplified with primers, 3'arm-ME-Fwd:
5'-GCTCCTAGCTAGAGTTGAGGGG-3' and 3'arm-ME-Rv:
5'-TCTAATCCAGCCAGTTTTAAAC-3'. Two homology arms were ligated with both
end of the PGK-Neo cassette and used for the homologous recombination in *E. coli*
harboring Aid-cre-cd2. Nucleotide sequences of modified area in final constructs were
checked by sub-cloning and sequencing.

Transgenic mice were generated by injection of supercoiled BAC DNA into CD-1 or
C57BL/6 fertilized eggs by the Laboratory for Animal Resources and Genetic
Engineering, RIKEN (accession numbers: CDB0481T, Aid-cre-cd2; CDB0482T,
dCEBP-ac-cre; CDB0483T, dR4-ac-cre; CDB0484T, dR2-ac-cre:
<http://www.cdb.riken.jp/arg/TG%20mutant%20mice%20list.html>) and Reproductive
Engineering Team, Institute of Virus Research, Kyoto University. Founder mice had been
backcrossed to C57BL/6 for more than 5 generations. All mice were bred and maintained
in specific pathogen-free conditions at Kyoto University and Riken Center for
Developmental Biology Animal Experiment. All mouse protocols were approved by the
Animal Research Committee, the Institutional Animal Care at RIKEN and Graduate
School of Medicine, Kyoto University (Permit Number: 10055).

Rosa-tdRFP [31] mice were bred and maintained in specific pathogen-free
conditions at Institute of Laboratory Animals Graduate School of Medicine, Kyoto

University. Double mutant animals were obtained by intercrossing these animals.

Southern blot and PCR analysis.

DNA from tail biopsies was used for PCR and Southern blot analysis. For Southern blotting, DNA was digested with *HindIII*, then separated on a 0.8% agarose gel and transferred to a nylon membrane (Roche). Blots were air-dried and DNA was crosslinked using UV light. DIG-labeled DNA probes indicated in Fig. 1 were used for hybridization in ULTRAhyb Buffer (Ambion; Applied Biosystems). All procedures for the DIG application system (Roche) were performed according to the manufacturer's recommendations.

PCR primers used for confirmation of generated BAC constructs are followings:

dR2-Fwd	5'-ATATGCAGCTCATGATGAGCCA-3',	dR2-Rv	5'-
	GTTATAATAATAAGCACAGGTA-3';		dR4-Fwd
	5'-TGAATTGGTTCACCTCCCCCTA-3',		dR4-Rv
	5'-CAGCTTACAGGAAACTTCCCA-3'.		

Antibody staining, detection of hCD2 expression and cell sorting

Single cell suspensions from tissues in FACS buffer (PBS supplemented with 4% FBS; 1 mM HEPES; 0.6% sodium citrate) were stained with the following monoclonal antibodies conjugated with FITC, PE, PE-Cy7, APC, APC-Cy7, Biotin or Pacific-Blue: specific for B220 (RA3-6B2), CD93 (AA4.1), IgM (R6-60.2), IgD (11-26), CD38 (90), CD3 (500A2), CD4 (GK1.5), CD8 (53-6.7), CD44 (IM7), CD62L (MEL-14), and hCD2 (RPA-2.10) (BD Biosciences or eBioscience). Biotinylated reagents were detected with

either streptavidin-APC-Cy7 or streptavidin-APC (BD Biosciences). Stained cells were acquired by FACSCantoII (BD Biosciences) or FACS Aria for analysis and cell sorting, respectively. Acquired data were analyzed by using FlowJo software (Tree Star, Inc.). Alive lymphocyte population was gated according to forward and side scatters, and 7-amino-actinomycin D staining.

Quantitative Real-time PCR

Total RNA was extracted by TRIzol reagent (Invitrogen Life Technologies), and cDNA were synthesized with TaqMan® Reverse Transcription Reagents (Applied Biosystems). Gene expression was assessed by real-time PCR by using iQ Sybr green supermix (Bio-Rad Laboratories, Inc.). Data were normalized to the *Gapdh* expression. Primers used were: *hCD2*-forward 5'-GACCACCAGCCTGAGTGCAA-3', *hCD2*-reverse 5'-GCTCCTCATCATTTCTCCGAC-3'; *Gapdh*-forward 5'-TGTGTCCGTCGTGGATCTGA-3', *Gapdh*-reverse 5'-CCTGCTTCACCACCTTCTTGAT-3'.

Results

Construction of Aicda BAC transgenic mice

To examine the *in vivo* function of the regulatory elements of *Aicda*, we used the BAC transgenic strategy. The BAC clone RP24-68I7 that covers the 190-kb region of *Aicda* and its flanking regions was successfully used to establish a faithful monitoring system for the *Aicda* expression [15]. A basic construct, Aid-cre-cd2 harboring a bicistronic expression cassette for human CD2 and cre in the AID coding region was previously generated (Fig. 1A) [30]. The *Aicda* BAC contains three regulatory regions (region 2, region 3 and region 4) in addition to the promoter (region 1) as assessed by the *in vitro* culture system [14,25,26]. The locations and binding motifs of region 2 and region 4 are described above (Fig. 1).

To examine the *in vivo* function of the regulatory elements identified *in vitro*, deletion mutations were introduced to the basic construct, giving rise to four additional constructs (Fig. 1A and Supplementary Fig. 1). The entire region 2 was deleted in the dR2-ac-cre construct. In dME-ac-cre, the silencer elements, c-Myb and E2F binding sites, in region 2 were deleted. The deletion of the whole region 4 was introduced in dR4-ac-cre, while dCEBP-ac-cre contains a restrictive deletion of the two C/EBP sites. We generated transgenic mice carrying these BAC constructs and crossed with the rosa-tdRFP reporter mouse [31], thereby *Aicda* promoter activation can be visualized by the expression of hCD2 and RFP for the ongoing and past expression, respectively (Fig. 1B). In the

previous report [15], cre-loxP mediated irreversible expression of reporters at the *rosa26* locus allowed more sensitive detection than direct staining of hCD2. In fact, it visualized the transient and/or marginal activation of the *Aicda* promoter in immature B cells and memory T cells [15]. All the mice used in this paper were crossed with the *rosa26*-tdRFP reporter strain unless otherwise described.

Positive regulation of Aicda expression by region 2 and region 4

We chose 3 or 5 lines of transgenic mice for each construct (Table 1), which carry a single copy of the transgene while approximately 4 copies were detected in the transgenic mouse published previously (Fig.1C) [15]. About one percent of splenic B cells were RFP+ in all these *Aid*-cre-*cd2* lines, indicating that the basic BAC transgene was expressed similarly in the animal (Table 1). In fact, germinal center B cells, in which *Aicda* is actively transcribed, strongly expressed hCD2 and RFP with double positive cells reaching about 35% of the B220+CD38- gated population (Fig. 2). Next, germinal center B cells in Peyer's patch of the mice carrying manipulated BAC reporter constructs were also examined. The deletion of region 2 or region 4 completely abolished the *Aicda* expression on the transgene (Fig. 2). Similar inactivation was observed in five each of independent founder lines of *dR2*-ac-cre and *dR4*-ac-cre (Table 1). We conclude that region 2 and region 4 are essential for normal *Aicda* expression in activated B cells *in vivo*.

As described above, region 4 contains the binding motifs for NF- κ B, STAT6 and Smad3/4 that are involved in the signaling pathway of CD40, IL-4 and TGF- β ,

respectively [32-34]. In addition to these elements, the importance of two C/EBP sites was demonstrated in the *in vitro* reporter assay [26]. Consistently, dCEBP-ac-cre in which two C/EBP sites including their intervening sequence in region 4 were selectively deleted (Supplementary Fig. 1), did not express the *Aicda* reporters, indicating that the C/EBP sites are in fact indispensable for the enhancer activity of region 4 in activated germinal center B cells (Fig. 2 and Table 1).

Suppressive activity of E2F and c-Myb motifs

In contrast to the enhancer function of the C/EBP sites in region 4, E2F and c-Myb binding sites showed the silencer activity *in vitro* [26]. To determine if the E2F and c-Myb sites are responsible for silencing *Aicda* in germinal center B cells, the transgenic lines with dME-ac-cre construct were examined. In contrast to the dR2-ac-cre transgenic mice, we observed the augmented expression of both reporters in germinal center B cells in Peyer's patch of the dME-ac-cre transgenic mice (Fig. 2). The percentage of hCD2 and RFP positive cells in their germinal center B cells was almost double of those in the *Aid-cre-cd2* transgenic mice. The mean fluorescent intensity of hCD2⁺ B cells was similar between *Aid-cre-cd2* and dME-ac-cre (Fig. 2), suggesting that the efficiency of *Aicda* expression is not so much different once it is activated. In addition, the percentages of RFP positive B cells in various lymphoid organs were significantly higher in dME-ac-cre than those in *Aid-cre-cd2* transgenic mice (Fig. 2 and Table 1). The results suggest that the *Aicda* promoter of dME-ac-cre was more readily activated than that of *Aid-cre-cd2*. To examine whether the increase in the percentage of hCD2 and RFP

positive cells correlate with the amount of the mRNA level, we performed RT-qPCR after sorting of germinal center B cells as well as other developmental stages of B cells. Consistent with FACS analysis, the mRNA levels of the reporter gene were significantly higher in these sorted cells of dME-ac-cre transgenic mice than those of Aid-cre-cd2 (Fig. 3). Taken together, these results indicate that E2F and c-Myb elements negatively modulate the activity of the *Aicda* promoter in the germinal center B cells *in vivo*.

To determine if the deletion of the silencer elements E2F and c-Myb upregulates the modest AID expression in immature B cells, we examined the reporter expression in developing B cells in dME-ac-cre mice. We did not detect robust expression of hCD2 or RFP in immature stage of B cells in dME-ac-cre mice (Fig. 4AB). However, the level of hCD2 mRNA expression was augmented in sorted preB and transitional B cells ($p < 0.05$) compared with those of *Aicda*-cre-cd2 (Fig. 3B). The RT-qPCR detection seems to be sensitive than BAC reporters. The results support the notion that the E2F and c-Myb elements in region 2 negatively regulate the *Aicda* promoter in immature B cells.

The germinal centers in Peyer's patch spontaneously develop by continuous stimulation from gut flora [35]. It is important to examine if the increased induction of RFP positive cells in dME mutation are dependent on chronic stimuli or not. For this purpose, we immunized naïve mice by intra-peritoneal injection of SRBC and analyzed germinal center B cells on day 7, the early phase of the immune response [36]. Germinal center B cells in the spleen were gated as B220⁺CD38⁻ cells [37]. As shown in Figure 5, the hCD2 and RFP positive cells were observed in spleen germinal center B cells with a much higher percentage in dME-ac-cre mice as observed in the Peyer's patch. Therefore,

the silencing activity of c-Myb and E2F elements negatively regulates the *Aicda* promoter activation at the initiation stage of B cell activation.

Deletion of the silencer elements in the region 2 did not cause the enhancement of *Aicda* expression in T cells.

Since the silencer activity of region 2 is not limited in B cells [26,38], it is conceivable that the silencer may be responsible for repressing *Aicda* in other cell lineages. The *Aicda* gene is weakly expressed in chronically activated T cells that coincide with the development of the IL-10 producing T cell subpopulation [15]. We therefore examined the reporter expression in T cells of dME-ac-cre mice, but failed to observe a significant fraction of RFP⁺ cells (Fig. 6). Unlike the previous multiple copy transgenic mice, the RFP⁺ T cells were undetectable in the present single-copy transgenic mice carrying the basic construct, *Aid-cre-cd2* (Table 1). Since the dME-ac-cre construct efficiently expressed in germinal center B cells, inability of the expression is not due to the functional inactivation of transgene itself. Although the c-Myb and E2F have strong suppressing effects on the *Aicda* promoter, the *Aicda* expression at the marginal level in T cells appears to require additional positive regulation [26].

Discussion

AID is the molecule central for the antigen-induced alteration of immunoglobulin genes in activated B cells while its mutagenic activity can induce tumorigenesis in various types of cells [8]. AID is also suggested to play some roles in immature B cell selection and T cell activation although its expression level was marginal as compared with activated B cells [15,23,24]. It is thus clear that the *Aicda* promoter is finely regulated by the cell lineage and the cellular environment. Understanding how the *Aicda* gene is repressed in most of the cells, moderately expressed by some environmental stimuli, and fully expressed in activated B cells, is critical for elucidation of the physiological and pathological function of AID. However, the *Aicda* expression under the physiological condition has not been revealed in previous *in vitro* or *ex vivo* cellular studies [14,25-29,39].

Critical roles of region 2 in B cell specific AID expression

In the present study, we have demonstrated the important functions of region 2 and region 4 for the induction of *Aicda* in activated B cells *in vivo*. **Although we did not evaluate the activity of region 3** [14], **all of the constructs that we used in this study contain region 3. Therefore, it is clear that region 3 cannot functionally compensate the lack of region 2 or region 4.** The region 2 contains binding sites of B cell-specific transcription factors, i.e. Pax5 and E proteins, and the present *in vivo* result is consistent with the previous *in vitro* report that has shown the enhancer activity in this region [26,29]. In addition to the B cell specific enhancers, E2F and c-Myb sites in region 2

functioned as a silencer in the *in vitro* reporter assay. The silencers counteracted against positive regulatory elements of region 2 and region 4 in B cells as well as non-B cells [26]. To elucidate the mechanism for lineage-specific and activation-specific induction of AID, it was important to evaluate the silencer function of E2F and c-Myb sites under the physiological conditions. As expected, these elements were demonstrated to function as elements that disturb the *Aicda* promoter activation in *in vivo* activated B cells. The percentage of hCD2-positive B cells in germinal centers was drastically increased by the deletion of E2A and c-Myb sites (dME-ac-cre), suggesting that the elements reduce the frequency of the activated *Aicda* promoter within the gated population. On the other hand, when hCD2 positive germinal center B cells are compared between *Aid-cre-cd2* and *dME-ac-cre*, the mean fluorescent intensities of hCD2 staining were not significantly different. Since the fluorescent intensity of hCD2 should correlate with the transcription efficiency of the reporter gene in individual cells, the observation indicates that the E2F and c-Myb sites do not affect the transcription efficiency of the active reporter gene. These results may suggest that E2F and c-Myb negatively regulate the activation step of the *Aicda* promoter, but do not affect the transcription efficiency of the already-activated promoter.

Usually, *Aicda* is not strongly expressed in other cells than activated B cells; therefore *Aicda* silencing should be stably maintained in other cells, otherwise the exposure to various environmental stimuli could induce this dangerous internal mutagen even at a minimal amount. In this context, the E2F proteins are known to associate with the pRB protein to function as a constitutive as well tentative gene repressor [40].

Because the silencers in region 2 can function not only in B cells but also in other lineages of cells [26], the silencers are expected to contribute B cell-specific *Aicda* suppression. We actually expected to observe ectopic expression of *Aicda* reporters in the absence of the silencers (dME-ac-cre mice). Unexpectedly, however, we did not observe either ectopic expression or enhanced expression of the reporter in immature B cells and developing T cells, which modestly activate the *Aicda* promoter under the normal condition [15]. These results suggest that the removal of silencers is not enough to induce the ectopic *Aicda* expression. Probably, the activity of B cell specific enhancers (Pax5 and E proteins) may not be sufficient to induce the promoter activation in most of cells including immature B cells and activated T cells, indicating that the coordinated balance between enhancers and silencers is critical for the lineage- and activation-specific expression of AID. An alternative possibility that other unidentified silencers are required for maintaining the lineage specificity is less likely because previous extensive deletion analysis did not find other obvious silencer elements in the AID gene and its flanking regions [26].

The previous study has clearly shown that the BAC transgenic system using the cre recombinase coupled with floxed RFP can detect the low level expression of *Aicda* [15]. The present study however could not visualize the *Aicda* expression records in immature B cells and activated T cells. Since the *Aicda* expression in these cells has been confirmed by other methodologies by different groups, it is unlikely that previous observation was artifact [14,15,21,22]. Rather, it seems that the sensitivity of the system in the current study was lower than those reported by Qin et al. (2011). The difference is likely due to

the copy number of transgenes. Although the previous transgenic mice carried about 4 copies, the current study chose animals only with a single copy integration.

Positive regulation by enhancers in region 4

The role of region 4 in *Aicda* regulation was first demonstrated by *in vitro* analysis of Tran et al. [26]. The current study convincingly demonstrated *in vivo* the critical role of the enhancer elements of the region 4 which includes the binding sites for NF- κ B, STAT6, Smad3/4 and C/EBP. Although the dCEBP-ac-cre selectively deleted the tandem C/EBP binding elements and left others intact, the expression of *Aicda* reporters completely disappeared. This suggests the indispensable role of C/EBP protein in the *Aicda* regulation as previously shown *in vitro*. NF- κ B, STAT6 and Smad proteins are major transcription factors involved in the CD40, IL-4, and TGF- β signaling pathways, respectively.

In contrast to those well-known factors that contribute to the induction of CSR, the involvement of C/EBP proteins in CSR has not been well appreciated before. Among many C/EBP proteins, the binding of C/EBP β to the region 4 has been suggested by chromatin immunoprecipitation assay [26]. Since C/EBP β can physically interact with NF- κ B and modulate its activity [41-43], they can functionally cooperate with each other, and are often involved in the signaling pathway mediated by toll-like receptors and cytokines, which are related with inflammatory reactions [41,43-46].

For example, in the case of the COX2 promoter that contains a tandem array of NF- κ B and C/EBP binding sites, a mutation of the C/EBP binding site abolished the

NF- κ B-dependent promoter activation in renal medullary interstitial cells [43]. The results suggest that the C/EBP pathway dominantly regulates the gene expression in the tonicity-induced COX2 expression. The NF- κ B and C/EBP binding sites are located close to each other in region 4 although, in contrast to the *COX2* gene, the distance is almost 100 bp. CD40 signaling through NF- κ B is well established to be required for AID induction. In this context, the critical role of C/EBP β in *Aicda* expression *in vivo* would be reasonable. It is interesting to elucidate the molecular mechanisms by which C/EBP and NF- κ B would interact and cooperatively regulate *Aicda* expression. In summary the activation of the AID expression requires not only B cell specific enhancers in region 2 but also stimulation-dependent enhancers in region 4.

Acknowledgements

References

1. Muramatsu M, Kinoshita K, Fagarasan S, Yamada S, Shinkai Y et al. (2000) Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. *Cell* 102: 553-563.
2. Revy P, Muto T, Levy Y, Geissmann F, Plebani A et al. (2000) Activation-induced cytidine deaminase (AID) deficiency causes the autosomal recessive form of the Hyper-IgM syndrome (HIGM2). *Cell* 102: 565-575.
3. Nagaoka H, Muramatsu M, Yamamura N, Kinoshita K, Honjo T (2002) Activation-induced deaminase (AID)-directed hypermutation in the immunoglobulin Smu region: implication of AID involvement in a common step of class switch recombination and somatic hypermutation. *J Exp Med* 195: 529-534.
4. Petersen S, Casellas R, Reina-San-Martin B, Chen HT, Difilippantonio MJ et al. (2001) AID is required to initiate Nbs1/gamma-H2AX focus formation and mutations at sites of class switching. *Nature* 414: 660-665.
5. Liu M, Duke JL, Richter DJ, Vinuesa CG, Goodnow CC et al. (2008) Two levels of protection for the B cell genome during somatic hypermutation. *Nature* 451: 841-845.
6. Kato L, Begum NA, Burroughs AM, Doi T, Kawai J et al. (2012) Nonimmunoglobulin target loci of activation-induced cytidine deaminase (AID) share unique features with immunoglobulin genes. *Proc Natl Acad Sci U S A* 109: 2479-2484.
7. Kotani A, Okazaki IM, Muramatsu M, Kinoshita K, Begum NA et al. (2005) A target selection of somatic hypermutations is regulated similarly between T and B cells upon activation-induced cytidine deaminase expression. *Proc Natl Acad Sci U S A* 102: 4506-4511.
8. Honjo T, Kobayashi M, Begum N, Kotani A, Sabouri S et al. (2012) The AID Dilemma: Infection, or Cancer? *Adv Cancer Res* 113: 1-44.
9. Okazaki IM, Hiai H, Kakazu N, Yamada S, Muramatsu M et al. (2003) Constitutive expression of AID leads to tumorigenesis. *J Exp Med* 197: 1173-1181.
10. Takai A, Toyoshima T, Uemura M, Kitawaki Y, Marusawa H et al. (2009) A novel mouse model of hepatocarcinogenesis triggered by AID causing deleterious p53 mutations. *Oncogene* 28: 469-478.

11. Ramiro AR, Jankovic M, Eisenreich T, Difilippantonio S, Chen-Kiang S et al. (2004) AID is required for c-myc/IgH chromosome translocations in vivo. *Cell* 118: 431-438.
12. Takizawa M, Tolarova H, Li Z, Dubois W, Lim S et al. (2008) AID expression levels determine the extent of cMyc oncogenic translocations and the incidence of B cell tumor development. *J Exp Med* 205: 1949-1957.
13. Muramatsu M, Sankaranand VS, Anant S, Sugai M, Kinoshita K 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: 18470-18476.
14. Crouch EE, Li Z, Takizawa M, Fichtner-Feigl S, Gourzi P et al. (2007) Regulation of AID expression in the immune response. *J Exp Med* 204: 1145-1156.
15. Qin H, Suzuki K, Nakata M, Chikuma S, Izumi N et al. (2011) Activation-Induced Cytidine Deaminase Expression in CD4 T Cells is Associated with a Unique IL-10-Producing Subset that Increases with Age. *PLoS One* 6: e29141.
16. Matsumoto Y, Marusawa H, Kinoshita K, Endo Y, Kou T et al. (2007) Helicobacter pylori infection triggers aberrant expression of activation-induced cytidine deaminase in gastric epithelium. *Nat Med* 13: 470-476.
17. Epeldegui M, Hung YP, McQuay A, Ambinder RF, Martinez-Maza O (2007) Infection of human B cells with Epstein-Barr virus results in the expression of somatic hypermutation-inducing molecules and in the accrual of oncogene mutations. *Mol Immunol* 44: 934-942.
18. Fan J, Ma G, Nosaka K, Tanabe J, Satou Y et al. (2010) APOBEC3G generates nonsense mutations in human T-cell leukemia virus type 1 proviral genomes in vivo. *J Virol* 84: 7278-7287.
19. Machida K, Cheng KT, Sung VM, Shimodaira S, Lindsay KL et al. (2004) Hepatitis C virus induces a mutator phenotype: enhanced mutations of immunoglobulin and protooncogenes. *Proc Natl Acad Sci U S A* 101: 4262-4267.
20. Ishikawa C, Nakachi S, Senba M, Sugai M, Mori N (2011) Activation of AID by human T-cell leukemia virus Tax oncoprotein and the possible role of its constitutive expression in ATL genesis. *Carcinogenesis* 32: 110-119.
21. Mao C, Jiang L, Melo-Jorge M, Puthenveetil M, Zhang X et al. (2004) T cell-independent somatic hypermutation in murine B cells with an immature

- phenotype. *Immunity* 20: 133-144.
22. Ueda Y, Liao D, Yang K, Patel A, Kelsoe G (2007) T-independent activation-induced cytidine deaminase expression, class-switch recombination, and antibody production by immature/transitional 1 B cells. *J Immunol* 178: 3593-3601.
 23. Kuraoka M, Holl TM, Liao D, Womble M, Cain DW et al. (2011) Activation-induced cytidine deaminase mediates central tolerance in B cells. *Proc Natl Acad Sci USA* 108: 11560-11565.
 24. Meyers G, Ng YS, Bannock JM, Lavoie A, Walter JE et al. (2011) Activation-induced cytidine deaminase (AID) is required for B-cell tolerance in humans. *Proc Natl Acad Sci U S A* 108: 11554-11559.
 25. Yadav A, Olaru A, Saltis M, Setren A, Cerny J et al. (2006) Identification of a ubiquitously active promoter of the murine activation-induced cytidine deaminase (AICDA) gene. *Mol Immunol* 43: 529-541.
 26. Tran TH, Nakata M, Suzuki K, Begum NA, Shinkura R et al. (2010) B cell-specific and stimulation-responsive enhancers derepress *Aicda* by overcoming the effects of silencers. *Nat Immunol* 11: 148-154.
 27. Park SR, Zan H, Pal Z, Zhang J, Al-Qahtani A et al. (2009) *HoxC4* binds to the promoter of the cytidine deaminase AID gene to induce AID expression, class-switch DNA recombination and somatic hypermutation. *Nat Immunol* 10: 540-550.
 28. Park SR, Kim PH, Lee KS, Lee SH, Seo GY et al. (2012) APRIL stimulates NF-kappaB-mediated *HoxC4* induction for AID expression in mouse B cells. *Cytokine*
 29. Sayegh CE, Quong MW, Agata Y, Murre C (2003) E-proteins directly regulate expression of activation-induced deaminase in mature B cells. *Nat Immunol* 4: 586-593.
 30. Kwon K, Hutter C, Sun Q, Bilic I, Cobaleda C et al. (2008) Instructive role of the transcription factor E2A in early B lymphopoiesis and germinal center B cell development. *Immunity* 28: 751-762.
 31. Luche H, Weber O, Nageswara Rao T, Blum C, Fehling HJ (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: 43-53.
 32. Lalmanach-Girard AC, Chiles TC, Parker DC, Rothstein TL (1993) T

- cell-dependent induction of NF-kappa B in B cells. *J Exp Med* 177: 1215-1219.
33. Takeda K, Kishimoto T, Akira S (1997) STAT6: its role in interleukin 4-mediated biological functions. *J Mol Med* 75: 317-326.
 34. Heldin CH, Miyazono K, ten Dijke P (1997) TGF-beta signalling from cell membrane to nucleus through SMAD proteins. *Nature* 390: 465-471.
 35. Fagarasan S, Muramatsu M, Suzuki K, Nagaoka H, Hiai H et al. (2002) Critical roles of activation-induced cytidine deaminase in the homeostasis of gut flora. *Science* 298: 1424-1427.
 36. Jacob J, Kassir R, Kelsoe G (1991) In situ studies of the primary immune response to (4-hydroxy-3-nitrophenyl)acetyl. I. The architecture and dynamics of responding cell populations. *J Exp Med* 173: 1165-1175.
 37. Takahashi Y, Ohta H, Takemori T (2001) Fas is required for clonal selection in germinal centers and the subsequent establishment of the memory B cell repertoire. *Immunity* 14: 181-192.
 38. Nagaoka H, Tran TH, Kobayashi M, Aida M, Honjo T (2010) Preventing AID, a physiological mutator, from deleterious activation: regulation of the genomic instability that is associated with antibody diversity. *Int Immunol* 22: 227-235.
 39. Gonda H, Sugai M, Nambu Y, Katakai T, Agata Y et al. (2003) The balance between Pax5 and Id2 activities is the key to AID gene expression. *J Exp Med* 198: 1427-1437.
 40. Frolov MV, Dyson NJ (2004) Molecular mechanisms of E2F-dependent activation and pRB-mediated repression. *J Cell Sci* 117: 2173-2181.
 41. Matsusaka T, Fujikawa K, Nishio Y, Mukaida N, Matsushima K et al. (1993) Transcription factors NF-IL6 and NF-kappa B synergistically activate transcription of the inflammatory cytokines, interleukin 6 and interleukin 8. *Proc Natl Acad Sci U S A* 90: 10193-10197.
 42. LeClair KP, Blonar MA, Sharp PA (1992) The p50 subunit of NF-kappa B associates with the NF-IL6 transcription factor. *Proc Natl Acad Sci U S A* 89: 8145-8149.
 43. Chen J, Zhao M, Rao R, Inoue H, Hao CM (2005) C/EBP{beta} and its binding element are required for NF{kappa}B-induced COX2 expression following hypertonic stress. *J Biol Chem* 280: 16354-16359.
 44. Takeuchi O, Akira S (2010) Pattern recognition receptors and inflammation. *Cell*

140: 805-820.

45. Ramji DP, Foka P (2002) CCAAT/enhancer-binding proteins: structure, function and regulation. *Biochem J* 365: 561-575.
46. Poli V (1998) The role of C/EBP isoforms in the control of inflammatory and native immunity functions. *J Biol Chem* 273: 29279-29282.

Figure legends

Figure 1. Generation of BAC transgenic mice.

(A) The *Aicda* gene on the BAC clone (RP24-68I7) was modified as described in the text. Open boxes with numbers denote the first to third exons. The expression cassette for cre and human CD2 are inserted in the second exon as indicated. Regulatory elements in region 2 or region 4 that were deleted or replaced by Frt sites (closed triangle) are drawn in enlarged scale with the number indicating relative position (base pair) from the transcription initiation site. The binding elements are marked by different symbols with the name of the factors. A thick line indicated the probe that was used for Southern blotting; the *HindIII* restriction sites adjacent to the probe were marked by H. pA, poly A signal.

(B) Aid-cre-cd2 TG mice crossed with *rosa*-tdRFP mice shows characteristic feature of cellular phenotype. By stimulation that turns on the *Aicda* promoter (*pAicda*), the BAC transgene expresses cre recombinase and hCD2 markers. Cre recombinates loxP sites in the indicator gene, *rosa*-tdRFP, thereby irreversibly expresses RFP. The RFP expression is maintained after the decline of *pAicda* activity.

(C) Southern blot analysis with the probe defined in (A) is shown. Genomic DNA from wild type (lane wt), Aid-cre-cd2 (lane 1), dR2-ac-cre (lane 2), dME-ac-cre (lane 3), dR4-ac-cre (lane 4), dCEBP-ac-cre (lane 5) and *Aicda*-cre (lane 6) [15] mice were digested by *HindIII*; the 3.2 kb and 3.5 kb bands correspond to endogenous and transgenic loci, respectively.

Figure 2. Expression of dual-reporters, hCD2 and RFP, in *in vivo* activated B cells of each BAC transgenic strain crossed with *rosa-tdRFP*.

Expression of hCD2 and RFP in B cells (B220+) or germinal center (GC) B cells (B220+CD38-) in Peyer's Patch (mice at 8 weeks of age) from *Aid-cre-cd2* line1; *dME-ac-cre* line1; *dR2-ac-cre* line1; *dR4-ac-cre* line1; *dCEBP-ac-cre* line1 mice. All crossed with *rosa-tdRFP* reporter strain. Mean fluorescent intensity (MFI) of hCD2 staining of the gated population is indicated. Percentage within the gate is indicated. FACS results are representative of three independent experiments with three mice each.

Figure 3. Correlation between the increase of RFP+ B cells and of the hCD2 mRNA in *dME-ac-cre* mice in germinal center B cells.

(A) Representative of sorting and gating strategies from germinal center of Peyer's patches (PP), bone marrow (BM) and spleen (SPL) in *Aid-cre-RFP* mice (8-weeks). Cells were stained by indicated markers for FACS analysis. Sorting gates were indicated. (B) Sorted cells were analyzed by RT-qPCR for expression of hCD2. Eight-week-old of *Aid-cre-cd2* line1 and *dME-ac-cre* line1 mice were used. The values were normalized by *Gapdh* expression. Data represents means \pm s.d. (bars) of three independent experiments. An asterisk indicates mild statistic difference between *Aid-cre-cd2* and *dME-ac-cre* by two-tailed Student's t-test (preB, $p = 0.042$; T1, $p = 0.02$; T2, $p = 0.014$; GC, $p = 0.024$). T1, transitional B cell; T2, transitional B cell 2; MZ, marginal zone B cell; FO, follicular B cell; GC, germinal center B cell; n.d., not detectable.

Figure 4. FACS analysis of hCD2 and RFP expression in Aid-cre-cd2 and dME-ac-cre in B lineage lymphocytes.

(A) Spleen (SPL) and (B) Bone marrow (BM) B cells of Aid-cre-cd2 line 1 and dME-ac-cre line1 mice (8-week) were stained by indicated marker for FACS analysis. Numbers shown in each gate indicated percent of cell. FACS results are representative of three independent experiments with three mice each.

Figure 5. FACS analysis of hCD2 and RFP expression in immunized mice.

Aid-cre-cd2 line1 and dME-ac-cre line1 (12-weeks) mice were immunized with SRBC and sacrificed on day 7 after immunization. The percentage of hCD2 and RFP positive cells among B cells or GC B cell was indicated. SPL, spleen. FACS results are representative of two independent experiments.

Figure 6. Lack of hCD2 and RFP expression in peripheral T cells of dME-ac-cre mice.

FACS analysis of hCD2 and RFP expression in naïve CD3⁺ or CD4⁺CD3⁺ T cells isolated from (A) spleen (SPL), (B) mesenteric lymph nodes (mLN), (C) Peyer's patches (PP) and CD4⁺CD8⁺ T cells isolated from thymus (D). Data are representative of three independent experiments. Shown are the data from a mouse 24 weeks old. Numbers in the panels show the percentage of gated cells of each gate.

Table 1

Constructs (line #)	Percentage of RFP+ lymphocytes (mean \pm SEM)					
	Spleen		mLN		Peyer's patch	
	B(%)	T(%)	B(%)	T(%)	B(%)	T(%)
Aid-cre-cd2-#1	1.0 \pm 0.05	0	0.9 \pm 0.2	0	2.8 \pm 0.3	0
-#2	1.1 \pm 0.2	0	0.8 \pm 0.2	0	2.9 \pm 0.7	0
-#3	1.3 \pm 0.4	0	0.8 \pm 0.1	0	3.1 \pm 0.2	0.03 \pm 0.02
dR2-ac-cre-#1	0	0	0	0	0	0
-#2	0	0	0	0	0	0
-#3	0	0	0	0	0	0
-#4	0	0	0	0	0	0
-#5	0	0	0	0	0	0
dR4-ac-cre-#1	0	0	0	0	0.005 \pm 0.001	0
-#2	0	0	0	0	0	0
-#3	0	0	0	0	0	0
-#4	0	0	0	0	0	0
-#5	0	0	0	0	0	0
dCEBP-ac-cre-#1	0	0	0	0	0	0
-#2	0	0	0	0	0	0
-#3	0	0	0	0	0	0
-#4	0	0	0	0	0	0
-#5	0	0	0	0	0	0
dME-ac-cre-#1	1.9 \pm 0.1	0	1.2 \pm 0.2	0.01 \pm 0.01	4.2 \pm 0.8	0.06 \pm 0.01
-#2	1.9 \pm 0.2	0.07 \pm 0.06	1.2 \pm 0.1	0.02 \pm 0.01	4.4 \pm 0.8	0.06 \pm 0.06

All animals were crossed with *rosa-tdRFP*.

Three mice from each line were analyzed.

Percentages of RFP+ B cells in mice carrying dME-ac-cre transgenes were significantly higher in those in mice carrying Aid-cre-cd2 transgenes. $P = 0.0067$ (Spleen), $P=0.0034$ (mLN), $P=0.0021$ (Peyer's patch) by two-tailed unpaired Student's t test.

mLN, mesenteric lymph node.

Figure 1

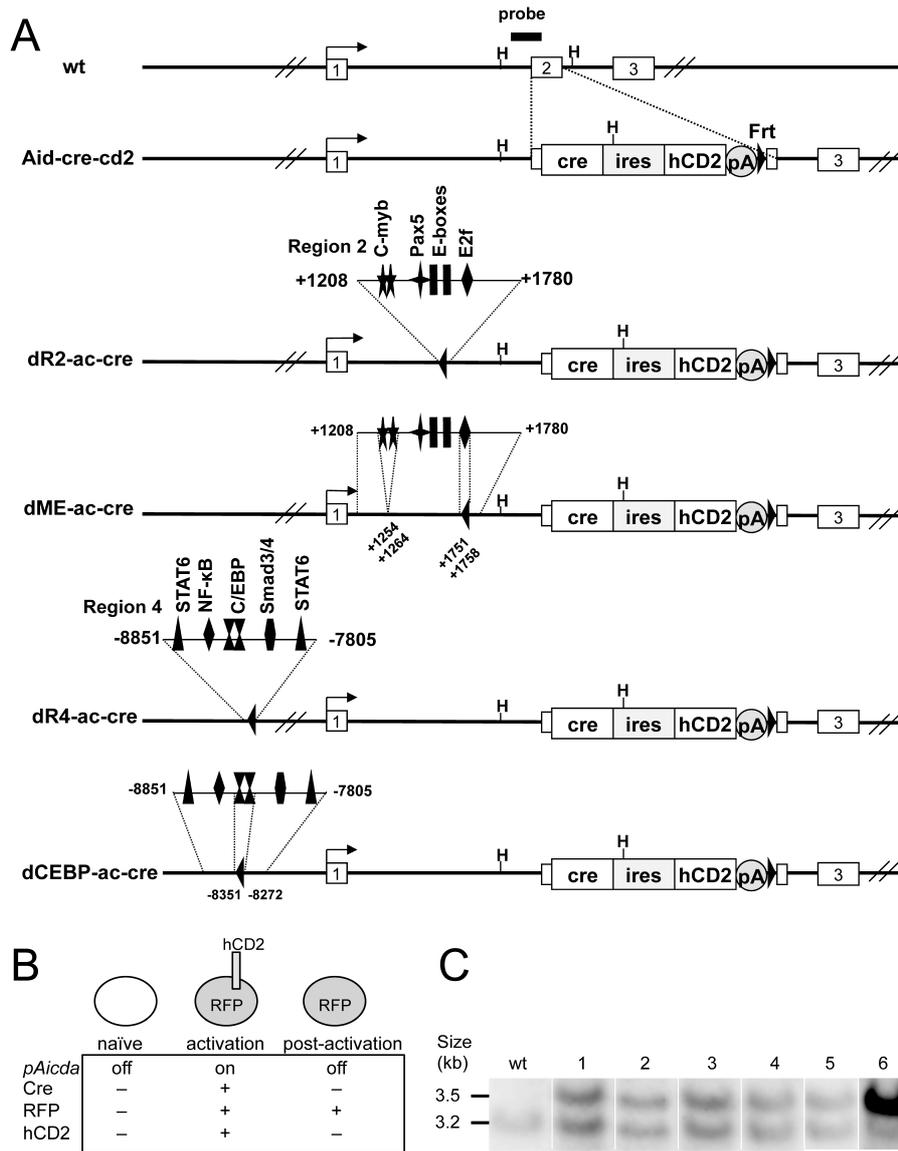


Figure 2

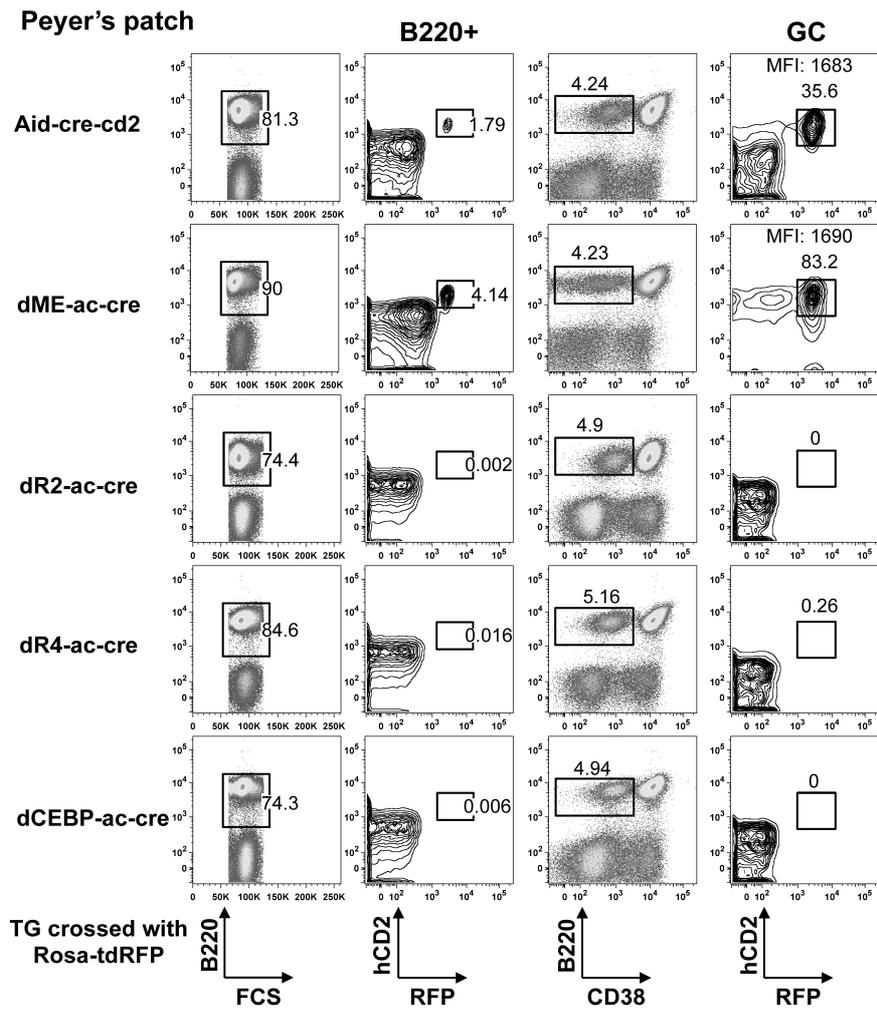


Figure 3

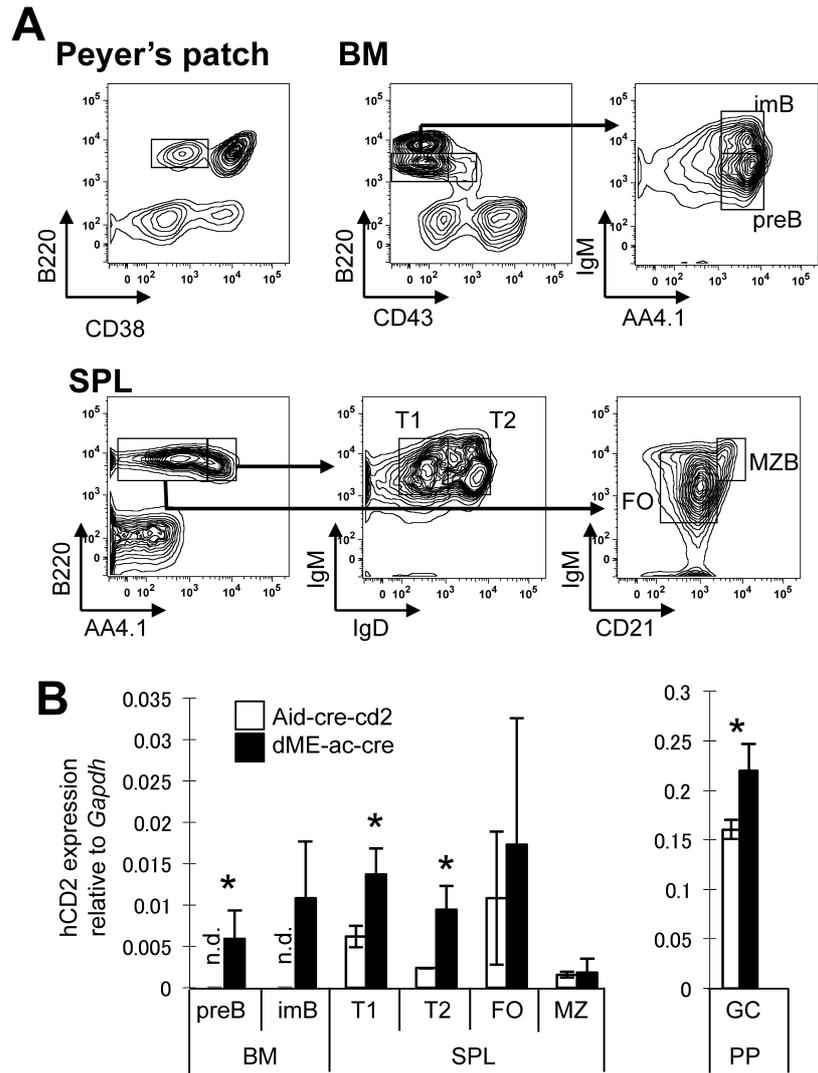


Figure 4

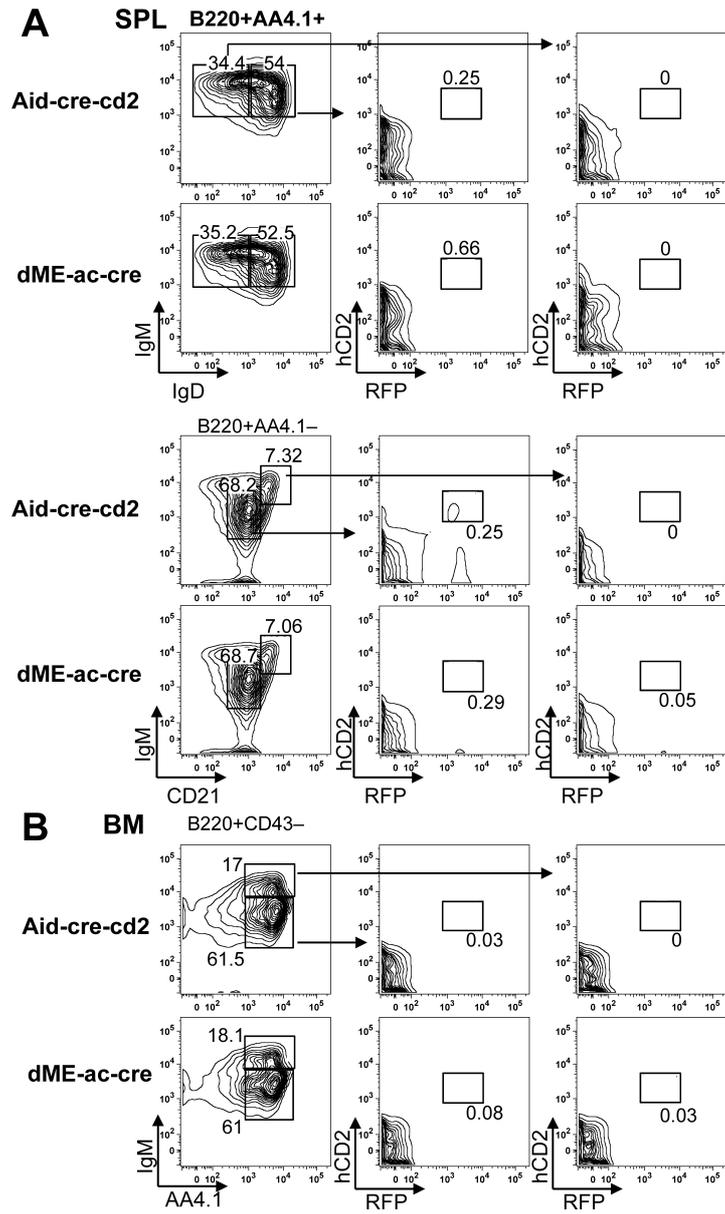


Figure 5

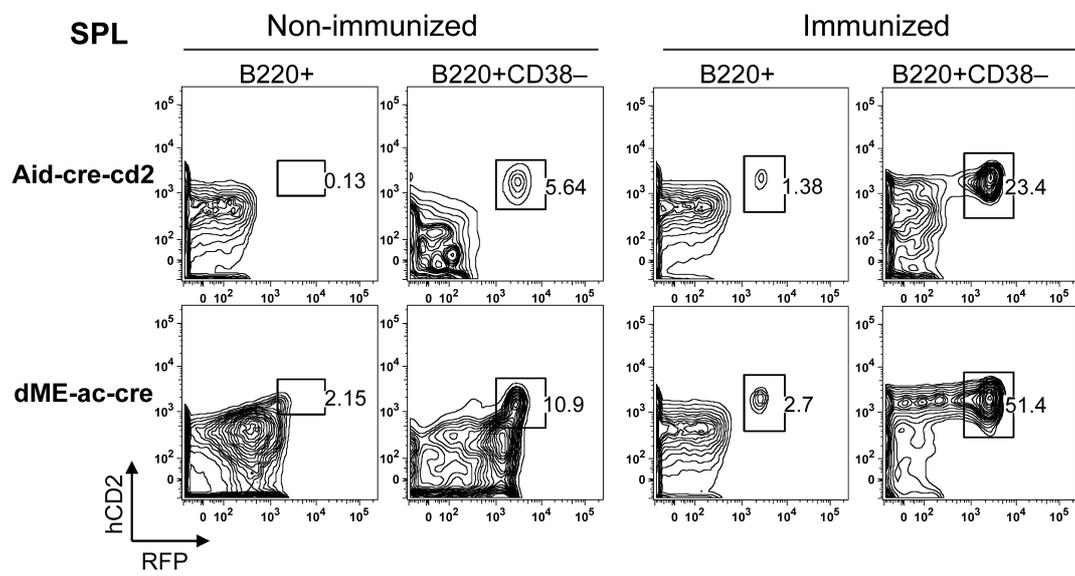
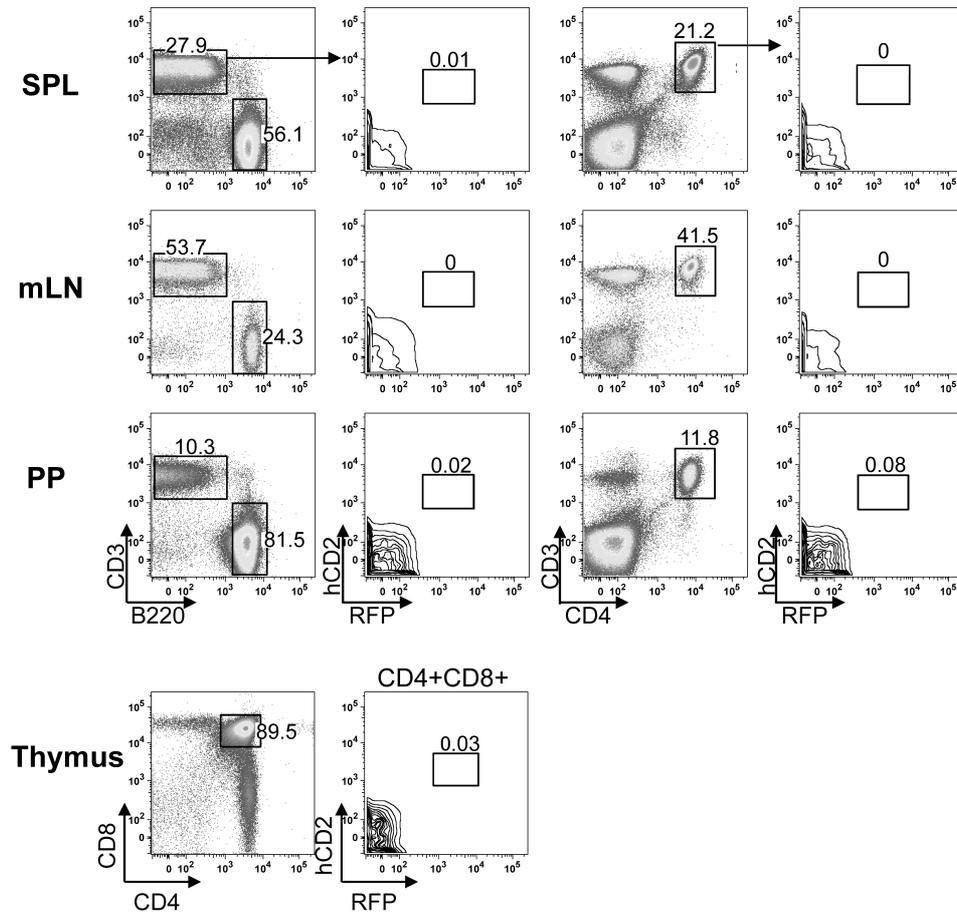
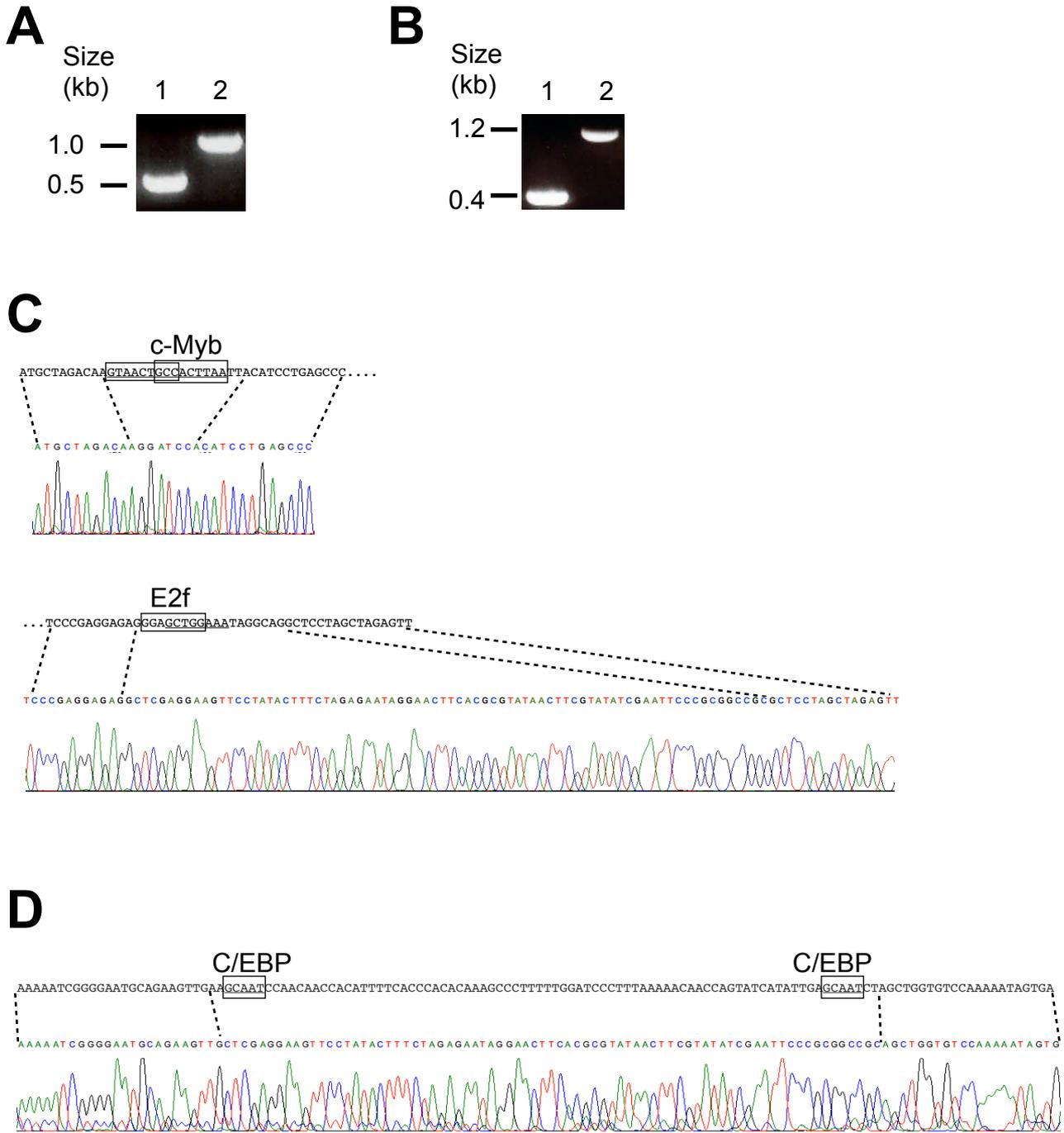


Figure 6



Supplementary Figure 1



Supplementary Figure 1

Confirmation of deletion/replacement of regulatory elements on BAC constructs.

PCR analysis confirmed the deleted of region 2 (0.5-kb) in dR2-ac-cre and that of region 4 (0.8-kb) in dR4-ac-cre (A) (B). Sequencing analyses to confirm the modifications introduced to dME-ac-cre (C) and dCEBP -ac-cre (D) were shown. The wild type sequence around the targeted elements (boxed) are presented. Sequencing chromatograms were shown below; corresponding region were indicated by dotted lines.