**B cell–specific and stimulation-responsive enhancers derepress the *Aicda* gene by overcoming the effects of silencers**

Thinh Huy Tran[1,2], Mikiyo Nakata[1], Keiichiro Suzuki[2], Nasim A. Begum[1], Reiko Shinkura[1], Sidonia Fagarasan[2], Tasuku Honjo[1] & Hitoshi Nagaoka[1]

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

[2]Laboratory for Mucosal Immunity, RIKEN, Yokohama 1-7-22, Tsurumi, Yokohama, 230-0045, Japan.

Correspondence should be addressed to T.H., (honjo@mfour.med.kyoto-u.ac.jp).

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

Phone: 81-75-753-4371; Fax: 81-75-753-4388

Abbreviations: *Aicda*, activation-induced cytidine deaminase gene; CSR, class switch recombination; SHM, somatic hypermutation; TSS, transcription start site.

**Abstract**

Activation-induced cytidine deaminase (AID) is essential for the generation of antibody memory but also targets oncogenes among others. We investigated the transcriptional regulation of the *Aicda* gene, which encodes AID, in the class switch–inducible CH12F3-2 cells, and found that the *Aicda* regulation involves derepression by several layers of positive regulatory elements in addition to the 5′ promoter region. The 5′ upstream region contains functional motifs for the response to signaling by cytokines, CD40-ligand or stimuli that activate NF-κB. The first intron contains functional binding elements for the ubiquitous silencers c-Myb and E2f and for B cell–specific activator Pax5 and E-box-binding proteins. The results revealed that the Aicda is regulated by the balance between B-cell-specific and stimulation-responsive elements and ubiquitous silencers.

## Introduction

Activation-induced cytidine deaminase (AID) is essential for the physiological alterations in immunoglobulin genes (Igs) that generate antibody memory, namely, class switch recombination (CSR) and somatic hypermutation (SHM)[1,2]. SHM introduces non-templated point mutations at a high frequency in the variable (V) regions of antibody genes, which, in conjunction with cellular selection mechanisms, give rise to high-affinity antibodies. CSR is a region-specific DNA recombination that occurs between two switch (S) regions located 5′ to each heavy-chain constant region ($C_H$) gene. This recombination juxtaposes a downstream $C_H$ gene with the V region gene by excising the intervening $C_H$ genes. CSR results in immunoglobulin class switching without changing antigen specificity[3,4].

AID is responsible for DNA cleavage in both the V and S regions[5–8]. Since DNA cleavage by AID can be deleterious to the genome, its is tightly regulated; virtually no expression is observed in non-B cells or even in B cells, unless they have been activated by appropriate stimuli, such as lipopolysaccharide (LPS), CD40-ligand (CD40L; http://www.signaling-gateway.org/molecule/query?afcsid=A000536), or certain cytokines, including interleukin 4 (IL-4; http://www.signaling-gateway.org/molecule/query?afcsid=A001262), transforming growth factor-β (TGF-β; http://www.signaling-gateway.org/molecule/query?afcsid=A002271) and interferon-γ (IFN-γ)[9,10]. Consequently, the transcriptional regulatory system of *Aicda* should include both lineage-specific and stimulus-specific response elements.

Complicating this picture, AID appears to target not only Ig genes but also the genes for several oncogenes that are frequently mutated or translocated to Ig loci in B cell malignancies (reviewed in [11]). Studies on AID-deficient mice suggest that AID may be involved in the pathogenesis of B cell malignancy[12–16]. Aberrant AID expression in human

B cell lymphomas was also speculated to correlate with degree of malignancy[11,17]. Interestingly, infection by certain viruses or bacteria that are potentially tumorigenic can induce AID expression in B and non-B cells[18–20]. Such ectopic expression of AID may contribute to tumorigenesis when AID expression persists during chronic infection[20–22]. In support of this idea, systemic overexpression of AID in transgenic animals leads to tumor development in various organs[22–24].

Several studies have addressed the mechanism for the transcriptional regulation of *Aicda*. The putative promoter region has been identified immediately upstream of the transcription start site (TSS), but this promoter is not lymphocyte-specific[25]. Two tandem E-boxes, recognized by E proteins such as E2A, located in the first intron have been proposed to be important for the induction of AID in B cells[26]. The transcription factor Pax5 [http://www.signaling-gateway.org/molecule/query?afcsid=A000403] may play a role, in cooperation with E proteins, in the B lineage-specific control of AID expression[27]. However, the location of the functional Pax5 motif is controversial[25,27]. In addition, the functional contributions of proposed putative binding sites for the signal transducer and transactivator 6 (STAT6; http://www.signaling-gateway.org/molecule/query?afcsid=A002236) and nuclear factor kappa B (NF-κB; http://www.signaling-gateway.org/molecule/query?afcsid=A001645), located upstream of the *Aicda* promoter, are not yet clear[28]. Finally, most previous studies were done either in cell lines whose AID expression is constitutive, or in primary B cells that had been fully stimulated by mitogens. Additional information requires detailed and thorough studies on the regulatory elements of *Aicda* in stable cell lines in which AID expression is inducible.

In this study, we performed an extensive analysis using luciferase reporter assays to determine the functional regulatory elements of the *Aicda* locus in CH12F3-2 cells. These cells carry out CSR with an efficiency as high as 70% upon stimulation with CD40L, IL-4,

and/or TGF-β (CIT)[29]. Because each of these CIT factors is also important for CSR stimulation in primary B cells, the control of the induction of AID expression in the CH12F3-2 cells is likely to reflect the physiological regulation of *Aicda*. We identified and characterized two clusters of regulatory elements, one about 8 kb upstream of the TSS and the other within the first intron. The former contains the major enhancers that respond to CIT stimulation, and the latter confers B cell specificity through a combination of B cell–specific positive elements and ubiquitous negative regulators. Our results revealed a sophisticated regulatory network in which AID expression is controlled by the balance among more than a dozen enhancers and silencers.

## Results

### Conserved regions in and around the *Aicda* gene

To identify candidates for *cis*-regulatory elements of the AID gene, we compared the mouse and human genomic DNA sequences within 50 kb upstream and downstream of the *Aicda* TSS by using PipMaker and mVISTA pairwise alignments. Consistent with similar analyses by others[25], we found four relatively well conserved non-coding regions (**Fig. 1a** and **Supplementary Fig. 1**). Region-1 (–1500 to +101), located immediately upstream of the TSS, contains a ubiquitously active promoter[25] and the following putative transcription factor binding motifs: one for NF-κB, one for STAT6, two for Sp that can bind both Sp1 and Sp3, and one for HoxC4-Oct (**Supplementary Fig. 1b**).

Region-2 (+121 to +2221), located in the first intron, part of which shows more than 70% local nucleotide sequence identity between mouse and human. This region contains the following binding motifs: two for NF-κB, two for Mzf1, two for Cp-2, three for c-Myb, one for Nkx2.5, one for Pax5, two E-boxes, and one E2f-binding site (**Supplementary Fig. 1b**). Only the two high-affinity E-box have been well characterized[26,27].

Region-3 (+16278 to +18378) is located downstream of *Aicda* at approximately 6 kb and 25 kb from exon 5 in mouse and human, respectively. This region was reported to be required for normal AID expression by BAC (bacterial artificial chromosome) transgenic mouse system[30]. Region-4 (–9224 to –7424), located approximately 8 kb upstream of the TSS, has not been studied before, although it contains candidate binding motifs for positive regulatory transcription factors, including two for NF-κB, two for STAT6, three for enhancer binding protein (C/EBP), and one for Smad3/4 (**Supplementary Fig. 1b**).

**The *Aicda* promoter is not CIT-responsive**

To study the *cis* regulatory elements of *Aicda* by the luciferase reporter assay, we used CH12F3-2 cells, in which AID expression is barely detectable and strongly up-regulated by stimulation with CIT, resulting in efficient CSR to IgA[9]. First, we generated plasmids containing serial 5′ deletions of region-1, to examine whether region-1 contributes to AID induction by CIT. This region contains several elements that were proposed or demonstrated to be involved in *Aicda* regulation, including an Sp-binding motif[25] that was originally reported as a putative Pax5 site[27], STAT6 and NF-kB sites[28], and HoxC4-Oct motifs[31]. We found all the tested constructs showed an approximately 4- to 8-fold increase in the luciferase activity, compared with pGL3, the promoter-less luciferase vector. We also found weak positive and negative effects in Sp motif and GA-rich sequence of which binding factor is unkown, respectively (**Fig. 1b**). Unexpectedly, however, none of these constructs showed any response to the CIT stimulation. The HoxC4-Oct element was previously demonstrated to be involved in AID expression using *Hoxc4*$^{-/-}$ B cells[31], but this fragment alone did not significantly respond to CIT. The HoxC4-Oct element may be involved in the basal expression of AID.

Therefore, we extended our analysis to the other conserved regions in order to identify CIT-responsive enhancer elements. We also sought stronger B cell–specific enhancer elements, because Sp1 and Sp3 are general transcription factors and the activity of the HoxC4-Oct motif was minimal, although it is lymphoid specific[31,32]. Since the 100-bp fragment immediately upstream of TSS, containing the Sp and HoxC4-Oct motifs, still activated transcription 5-fold, to a level almost equivalent to that induced by the entire region-1, we used this fragment as the minimal promoter.

**Regulation of the *Aicda* promoter by regions-2 and -4**

To delineate the role of the conserved regions in the regulation of AID expression, we constructed reporter vectors harboring region-1 in combination with regions -2, -3, or -4. The addition of region-2 to region-1 decreased the luciferase activity by 50%, and this construct did not respond to CIT stimulation (**Fig. 1c**). On the other hand, a construct containing both region-1 and region-4 induced enhanced luciferase activity in response to CIT stimulation, and its basal activity was also slightly higher than that of the region-1-only construct, suggesting that region-4 contains DNA elements responsive to IL-4, TGF-β and/or CD40L.

By contrast, the addition of region-3 did not result in any significant change in either the basal or CIT-induced AID promoter activity, in contrast to a previous report[30] (**Fig. 1c**). Serially deleted region-3 fragments did not show any significant effects on the luciferase activity driven by region-1, either (data not shown). Taken together, our data indicate that the basal expression of AID in CH12F3-2 is regulated negatively by region-2 and positively by region-4 while no enhancer activity was detectable in region-3 by the luciferase assay. In addition, region-4 was primarily responsible for the up-regulation of the *Aicda* gene in response to CIT stimulation.

**Region-2 confers B cell specificity**

To dissect the regulatory elements in region-2, serially deleted region-2 fragments were ligated downstream of the luciferase reporter construct driven by the minimal *Aicda* promoter (–101 to +1) (**Fig. 2a,b**). CIT stimulation did not cause a significant increase in the luciferase activity of CH12F3-2 cells expressing any of these constructs, supporting the idea that the CIT responsive element is located outside region-2.

Further study of region-2, to look for B cell–specific regulators, showed that a deletion of the 3′ sequence (+1709 to +1783) increased the luciferase activity approximately six times over that induced by full-length region-2, suggesting negative regulatory elements in the deleted sequence (**Fig. 2a**, line 4). The further serial deletion of the two E-boxes reduced the activity to almost equal that of the minimal promoter (**Fig. 2a**, lines 5, 6), in agreement with a previous report on the enhancer function of these E-boxes[26]. When the Pax5 element was also deleted, the luciferase activity was reduced further (**Fig. 2a**, line 7).

To confirm the positive roles of the Pax5 motif and E-boxes, we made constructs with only these motifs because the suppressing effect from the adjacent sequences masked their activity. The fragment containing only the Pax5 site and the two E-boxes enhanced the luciferase activity of the minimal promoter about 3-fold (**Fig. 2c**, line 4). Mutating any one of the three motifs obviously impaired the enhancer activity, indicating that all three contributed to the positive regulation of *Aicda* (**Fig. 2c**, lines 5-7). Since Pax5 and E proteins are well-established B cell–specific co-activators, which actually bind to this region in activated B cells[26, 27], they could be responsible for the activation of *Aicda* in B cells. In contrast, the NF-κB-3 site (see **Fig. 2a** for the positions of the numbered motifs) did not seem to play a relevant role in either basal or CIT induced- AID expression (**Fig. 2a**, line 10 and **Fig. 2b**, line 4).

The negative regulatory region (+1709 to +1789) contains the binding motifs for NF-κB-4, E2f, Cp2-2, c-Myb-3, and Mzf1-2, but only the E2f-binding site appeared to be responsible for the negative regulation by mutagenesis experiments (**Fig. 2c**, lines 4, 8 and 9, and **Supplementary Fig. 2**, lines 8, 9). Deletion of the upstream sequence (+1203 to +1530) that contains the Mzf1-1, Cp2-1, c-Myb-1, c-Myb-2, and Nkx2.5 binding motifs enhanced the luciferase activity, suggesting the presence of other negative regulatory motifs within this region (**Fig. 2a**, line 9; **Fig. 2b**, line 7). We excluded the contribution from the Mzf1-1 and Cp2-1 motifs by mutagenesis experiments as no significant differences were observed relative the intact region (**Supplementary Fig. 2**, compare lines 6, 7 with line 5). Although the suppressor activity of a 350-bp CT-rich sequence from +874 to +1221 was obvious only in the 5′ deletion construct (**Fig. 2a,b**), a similar sequence has a reported suppressor function in another gene, *Cyp1a1*, in epidermal keratinocytes[33]. Notably, the activity of the B cell-specific enhancers (the Pax5 site and E-boxes) was not sufficient to counter the effect of the suppressive elements, including c-Myb-1, c-Myb-2, Nkx2.5, E2f, and the CT-rich sequence of region-2 (**Fig. 1c**, **Fig. 2a,b**).

**Suppressor elements in region-2 counteract enhancers**

The interaction between regions-2 and -4 was examined. Region-2 suppressed the enhancer activity of region-4 regardless of CIT stimulation (**Fig. 2d**, line 4). Since region-4 amplified the luciferase activity strongly, we could confirm suppressor activities observed in the deletion experiments. Deletion of the CT-rich sequence together with the inactive NF-κB-3 motif led to a slightly higher luciferase activity (**Fig. 2d**, line 5). Additional point mutations in one of the three motifs in this region, namely c-Myb-1, E2f, and c-Myb-2 (overlapping the Nkx2.5 site), reduced the suppression activity (**Fig. 2d**, lines 6–8). The combined disruption of all these motifs along with the CT-rich region

almost completely abolished the suppressor activity of region-2 (**Fig. 2d**, line 9). Although the involvement of Nkx2.5 cannot be formally excluded, the knockdown experiment with siRNA oligonucleotides for c-Myb but not for Nkx2.5 increased the luciferase activities (**Supplementary Figs. 3,4**). Taken together, these results indicate that the binding motifs in region-2 for c-Myb and E2f, and the CT-rich sequence, function independently as suppressor elements to counteract the positive elements in region-2 and region-4.

**Elements responsible for CIT stimulation in region-4**

IL-4, TGF-β and CD40L induce AID expression independently and additively in CH12F3-2 cells[9]. To confirm that region-4 contains independent response elements for IL4, TGF-β or CD40L, each stimulation was separately examined. IL-4, TGF-β, and CD40L individually up-regulated the luciferase activity of the construct carrying region-4 (**Fig. 3a**). Combining the three stimuli resulted in the highest activity, and the increment appeared to be additive.

To identify the functional elements in region-4, a series of deletion constructs was generated (**Fig. 3b,c**). The inductive response to CIT stimulation greatly decreased by the deletions of the DNA region containing STAT6-1 (see **Fig. 3b** line 5), STAT6-2 (**Fig. 3c**, line 4), NF-kB-2 (**Fig. 3b**, line 6), Smad3/4 (**Fig. 3c**, line 6), C/EBP-1 (**Fig. 3b**, line 7), and C/EBP-2 (**Fig. 3c**, line 7). Among these elements, NF-kB-2 and two tandem C/EBP motifs seem to have strong contribution to the basal activity as well. In contrast, the deletions of the DNA region containing NF-kB-1 (**Fig. 3b**, line 4) and C/EBP-3 (**Fig. 3c**, line 5) elements had no effect on the luciferase activity.

The mutagenesis experiments confirmed the involvement of the region-4 elements. Mutations of either or both of the STAT6-binding sequences strongly reduced the luciferase activity induced by IL-4 alone as well as by CIT, indicating that both sites are

response elements for the IL-4 signal (**Fig. 4a**, lines 4–7). STAT6 knockdown almost completely blocked the luciferase induction by IL-4 stimulation (**Fig. 4b** and **Supplementary Fig. 5**). We conclude that the IL-4 signal activates the *Aicda* promoter through the STAT6-1 and –2 binding sites on region-4.

Destruction of the NF-κB-2 motif severely impaired the response to both CD40L alone and to CIT (**Fig. 4c**, line 4), and the NF-κB-specific inhibitor, sulfasalazine, almost completely suppressed the luciferase induction by CD40L through region-4 (**Fig. 4d** and **Supplementary Fig. 6**). The inactivation of NF-κB also reduced the basal activity.

When the Smad3/4 site was mutated, the TGF-β response and the basal activity were clearly suppressed (**Fig. 4e**, line 4). Furthermore, in agreement with the deletion experiment results, the two C/EBP binding sites were required for the response to both TGF-β alone and to CIT (**Fig. 4f**, lines 4, 5). Thus, the pair of C/EBP sites was required for the region-4 response to all three CIT components: IL-4, CD40L and TGF-β. However, the C/EBP sites alone did not respond to CIT stimulation (**Fig. 4g**), indicating that the pair of C/EBP sites is required but not sufficient.

Finally, we examined whether the CIT response was dependent on the B cell–specific enhancers (Pax5-binding site and E-boxes) of region-2 when the repressor elements co-exist. Deletion of Pax5-binding site and E-boxes from the reporter construct containing regions-2 and -4 still showed clear induction by CIT although the luciferase activity was reduced to about 35% (**Fig. 5a**, lines 4, 5). These results indicate that the B cell–specific enhancers are not absolutely essential for the CIT response through region-4, at least in our luciferase system.

The same constructs used above were tested in two non-B cells, 2B4.11 T cells and NIH3T3 fibroblastic cells. The basal luciferase activity in CH12F3-2 cells by the region-4 reporter alone was approximately twice that observed in the non-B cells, indicating that

CH12F3-2 cells express more intrinsic factors for AID expression than non-B cells (**Fig. 5b**, line 3). Addition of region-2 repressed the basal activity not only in the B cells but also in the other lineage of cells, indicating that the region-2 silencers are active in non-B as well as B cells (**Fig. 5b**, line 4). Importantly, the deletion of the Pax5-binding site and E-boxes reduced the luciferase activity only in B cells (**Fig. 5b**, lines 5).

**Involvement of candidate transcription factors**

To examine *in vivo* binding of the transcription factors to the candidate regulatory elements, we performed chromatin immunoprecipitation (ChIP) assay of region-2 or -4 (**Fig. 6a**). Twelve-hour cultivation with CIT significantly induced the specific binding of transcription factors, NF-κB, C/EBPβ, STAT6 and Smad4 to region-4 in CH12F3-2 cells (**Fig. 6b**). We tried c-Myb and some of E2f proteins (E2f-1, -4 and -5) without success partly because of the absence of high-quality antibodies (data not shown). C/EBPβ and STAT6 also bound to region-4 in spleen B cells activated by LPS and IL-4 (**Fig. 6c**). Binding of Pax5 and E-proteins to region-2 was already demonstrated[26,27].

To confirm involvement of these transcription factors in *Acida* regulation, we examined their expression profile in stimulated spleen B cells (**Fig. 7**). mRNAs of STAT6, NF-kB, Smads, C/EBPs, Pax5, E2a, cMyb, E2f-2 and -4 were rapidly induced by stimuli of LPS, IL-4 and TGF-β, declined by 12 h, and maintained thereafter. Expression of E2f-7 and -8, which can repress by binding to the E2f motif[34], increased 20 h after stimulation in parallel with AID expression. We also compared the expression of these factors in naïve and germinal center (GC) B cells. Most of the activators were already well-expressed in naïve cells, suggesting the activation of their function by the protein modification after stimulation (**Supplementary Fig. 7**). Expression of repressive E2fs, namely E2f-5, -7 and -8, and cMyb were increased in GC B cells. Since the suppressive elements were

functional in CIT stimulated CH12F3-2 cells, the increased expression of candidate repressor proteins (E2f and c-Myb) in activated B cells may be a feedback response to avoid the over-expression of AID (**Supplementary Fig. 8**).

**Discussion**

In the present study, we identified three functionally important regions in the 10-kb region upstream of the *Aicda* gene and within it. Region-1 serves as the basic promoter, and contains positive elements, Sp-binding sites, and a HoxC4-Oct motif, and a weak negative element, the GA-rich sequences. Region-2, located in the first intron of *Aicda*, also contains both positive (Pax5-binding site and E-boxes) and negative (c-Myb- and E2f-binding sites) elements. On balance, region-2 functions as a negative regulator when combined with region-1. Here we showed that the Pax5-binding site and E-boxes in region-2, contribute to the restriction of AID expression to the B cell lineage. Region-4 contains the functional binding sites for NF-κB, STAT6, and Smad3/4, which are, respectively, response elements for CD40L, IL-4 and TGF-β. A tandem pair of C/EBP binding sites is required for each of these elements to respond to CIT stimulation. A similar requirement for a tandem pair of C/EBP elements has been reported for the gene encoding the common β-chain of the IL-3, IL-5, and GM-CSF receptors[35].

Because AID promotes genome-wide damage and leads tumorigenesis, the promoter activity should be finely regulated. In fact, AID expression virtually restricted to activated B cells. How could such a restricted expression be achieved by the elements we found in this study? Perhaps, the key is the negative elements in region-2. These counteract against positive signals mediated from cytokine- or Toll-like-receptors (TLR) which are widely expressed. They also work against B cell-specific enhancers, of which binding proteins are functional even in non-activated B cells, thereby would confer the active transcription

only in activated B cells.

C-Myb often functions as a transcription activator, but it also represses many genes, including *Cd4*, *Nras*, *c-erbB-2* (refs. [36-38]). Interestingly the repressive c-Myb-binding site of the *Cd4*, whose expression is tightly regulated during T cell development, is also located in the first intron[38]. Some E2f proteins such as E2f-4, E2f-5, E2f-6 can reversibly repress a promoter in conjunction with other co-repressors[39]. The E2f protein can also induce stable repression by recruiting the polycomb complex[40]. Such a mechanism might be involved in the complete silencing of *Aicda* in non-lymphoid cells. Interestingly, we found increase of E2f-7 and -8 in *in vitro* activated and GC B-cell. E2f-7 and -8 counteract activating E2fs [34]. Given activating E2fs were also up-regulated in GC cells, they might bind to region-2 in certain condition, and E2f-7 and -8 would counteract against them. We could not detect any enhancer activity to the E2f site in region-2, but it could be due to the expression pattern of E2fs in CH12F3-2 which is repressor type dominant (not shown). It still remains unsolved which E2f proteins are actually responsible to the repressive function in region-2.

We found discrepancies between our results and previous studies on AID regulation. An active Pax5 binding site in region-1 was proposed, but later another group concluded the site is in fact an Sp-binding site[25,27]. Our mutagenesis experiments strongly indicated the Pax5 involvement, but the responsive site is the one in region-2. Another discrepancy was our inability to detect enhancer activity in region-3, which was reported to be required for AID expression in the BAC transgenic mouse[30]. Region-3 might contain a positive regulator which works only in the chromatin context.

Our study suggests that the region-4 is essential for the AID response in B cells to the environmental stimulation delivered by T cells, dendritic cells, and other cytokine-producing cells. Immature B cells were reported to express low amounts of AID by

signaling through TLR[41,42]. Furthermore it was shown that even gastric epithelial cells or hepatocytes can express AID upon infection[20,21]. The mechanism for this AID induction is not clear from our studies in B lineage cells. The NF-κB site in region-4 could be involved in the AID-induction signal via TLR or other pathways in response to bacterial or viral infection[43]. Negative regulators binding to the elements in region-2 should counteract NF-κB under normal conditions, to suppress the *Aicda* gene and retain genomic stability. We were unable to identify a candidate response element for IFN-γ (STAT1-binding site). Since IFN-γ can activate the NF-κB pathway[44], and also because STAT1 can form a heterodimer with other STAT proteins including STAT6 (ref. [45]), the *Aicda* regulation by IFN-γ would not need an obvious STAT1-binding motif. We could not test this hypothesis because CH12F3-2 cells do not respond to IFN-γ. We were unable to test the effect of B cell receptor signaling in CH12F3-2 cells because it affects CSR only marginally (data not shown).

Our current assay system has obvious limitations for recapitulating the complete repression of *Aicda* by silencers. The luciferase assay depends on transient transfection with high dose of plasmid DNA whose regulation might be different from that of the endogenous *Aicda* gene. However, ChIP analyses and expression profiles of candidate transcription factors in CH12F3-2 and spleen B cells are in general agreement with the conclusion derived from the luciferase assay. The next step to confirm our findings will be *in vivo* mutagenesis using either gene targeting or BAC transgenic strategies.

In summary, our findings emphasize that two separate regions, region-2 and region-4, carry out physiologically distinct regulatory functions in *Aicda* expression in which derepression from the effects of silencers by B lineage–specific and stimulation-responsive enhancers is critical. The evolution of such a sophisticated regulatory system for the *Aicda* locus seems reasonable, even necessary, since its product is essential to

antibody memory but has the potential to induce genomic instability.

**Acknowledgement**

The authors are grateful to I. Taniuchi (RCAI, RIKEN) for his critical comments and to Y. Shiraki and T. Kanda for their help for the preparation of the manuscript. This research was supported by a Grant-in Aid for Specially Promoted Research 17002015 from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

**Author Contributions**

T.H.T., T.H. and H.N. designed the study. T.H.T., M.N., K.S. and H.N. performed experiments. N.B., R.S. and S.F. provided new reagents and instructive suggestions. T.H.T., T.H. and H.N. wrote the manuscript.

**Figure legends:**

**Figure 1. The general characterization of regions-1 ~ -4 by luciferase reporter plasmids.** (**a**) A diagram of mouse *Aicda* locus. The positions of region-1 to -4 relative to five exons are indicated. (**b**) Luciferase constructs with full-length or deletion mutants of region-1, which carries the *Aicda* promoter, were analyzed. The structure of each construct is shown schematically. The plasmids were introduced into CH12F3-2 cells and cultured with or without CIT stimulation for 24 h. The positions of the candidate regulatory elements are indicated by the specific symbols at the top of the diagram. Open and filled bars indicate cultures with and without CIT stimulation, respectively. (**c**) Effects of regions-2, -3, and -4 on the *Aicda* promoter in region-1 were analyzed. Each region was ligated to the luciferase coding sequence as shown. Luciferase activities of each construct with or without CIT stimulation are shown as the ratio relative to non-stimulated pGL3

(top of each panel). Luc, luciferase coding sequence; R1, region-1; R2, region-2; R3, region-3, and R4, region-4. Data are representative of three independent experiment (mean +/– s.d.).

**Figure 2. Region-2 contains both enhancer and suppressor elements.** Firefly-luciferase reporter plasmids for region-2 were made in conjunction with the minimal *Aicda* promoter (-101 ~ -1; black bar). Full-length and region-2 fragments with serial deletions from either the 5′-end (**a**) or 3′-end (**b**) were used. Numbers on the diagram give positions relative to the TSS. Candidate elements are indicated at the top of panel **a**. Note that the c-Myb-2 site includes Nkx2.5, but is illustrated as two symbols that contact each other. Binding motifs that appear in more than one place are distinguished by a number after the motif name. (**c**) Constructs carrying the Pax5 motif and E boxes of region-2 were made as shown. (**d**) The luciferase construct with full-length region-4 and -2, designated as pR4-R2paidLuc, and its derivative plasmids with the deletions and mutations shown in the diagram, were generated and analyzed. The luciferase activities of each construct with or without CIT stimulation are shown as the ratio relative to non-stimulated pGL3 (top of each panel). Filled and open bars indicate activities with and without CIT stimulation, respectively. Elements destroyed by mutagenesis are indicated by arrows. Data are representative of three independent experiment (mean +/– s.d.).

**Figure 3. Analysis of CIT-responsive elements in region-4.** (**a**) CH12F3-2 cells were transfected with a luciferase construct containing the minimal *Aicda* promoter and region-4, pR4paidLuc, to analyze the response to IL-4 (line 4), TGF-β (line 5), and CD40L (line 6). Candidate response elements in region-4 are indicated by specific symbols on the diagram with their names indicated above. The stimuli used for each experiment are

indicated at the left of each construct. Luciferase activities of each construct with or without CIT stimulation are shown as the ratio relative to the non-stimulated pGL3 (top). The enhancer activity was identified by a making serial deletions of region-4 from the 5′-end (**c**) and the 3′-end (**c**) as indicated. Filled and open bars indicate with and without CIT stimulation, respectively. Data are representative of three independent experiment (mean +/– s.d.).

**Figure 4. Elements in region-4 that are responsive to the CIT stimulation.** Transfected CH12F3-2 cells were cultured for 24 h with no stimulation, IL-4, CD40L, TGF-β, or CIT as indicated in each panel. (**a**) Mutations or deletions were introduced to STAT6 elements on the pR4paidLuc plasmid as illustrated by arrows or the deletion of symbols, respectively. (**b**) CH12F3-2 cells were transfected with pR4paidLuc constructs and three kinds of RNAi oligonucleotide (#1-#3) for STAT6 (**Supplementary Table 1**), as well as with the control low GC oligonucleotide. (**c**) Mutation of the NF-kB-2 element in region-4 in pR4paidLuc reduced the luciferase activity in response to CD40L and CIT. (**d**) An NF-κB specific inhibitor, sulfasalazine, inhibited the enhancer activity of region-4 in pR4paidLuc in response to CD40L and CIT (**e**) Mutations of the Smad3/4 element reduced the luciferase activity of pR4paidLuc in response to TGF-β and CIT. (**f**) Luciferase assay showing the requirement for C/EBPs for the TGF-β response of pR4paidLuc and its derivative plasmids that have deletions or mutations as indicated in the diagram. (**g**) Modified pR4paidLuc plasmids with deleted C/EBP motifs or only the two C/EBP motifs of region-4 were constructed and analyzed. Data are representative of three independent experiment (mean +/– s.d.).

**Figure 5. Roles of Pax5 and E proteins in non-B cells.** Deletion of binding sites for

Pax5 and E proteins in region-4 of the pR4-R2paidLuc plasmid were generated as shown. (**a**) Pax5 and E protein contribute to the induction of AID expression in B cells. CH12F3-2 cells were transiently transfected with reporter constructs and cultured for 24 h with (filled bars) or without (open bars) CIT stimulation. Luciferase activities of each construct with or without CIT stimulation are shown as the ratio relative to non-stimulated pGL3 (top). (**b**) The reporter plasmids were introduced by transient transfection into CH12F3-2, 2B4.11 (T cell line), and NIH 3T3 (fibroblast cell line) cells and cultured for 24 h without stimulation. Luciferase activities of each construct in each cell line are shown as the ratio relative to that in CH12F3-2 cells transfected with pGL3 (top). Data are representative of three independent experiment (mean +/– s.d.).

**Figure 6. *In vivo* binding of the transcription factors examined by ChIP assay**. (**a**) The diagram shows the positions detected by real-time PCR. (**b**) CH12F3-2 cells cultured with or without CIT for 12 h were fixed for indicated time periods, and used for ChIP with antibodies against transcription factors shown in each panel. Difference between values from CIT stimulated (filled) and non-stimulated (open) samples was evaluated by one tail paired t-test. Data are representative of three (Stat6 and Smad4), five (p65), or six(C/EBPβ and control IgG) independent experiment (mean + s.d.). (**c**) Spleen B cells isolated from wild-type mice were cultured with LPS and IL-4 for 3 days and analyzed by ChIP. Difference between each region and silent gene, *Icos*, was analyzed. Data are representative of three independent experiment (mean + s.d.). *, $P < 0.05$; **, $P < 0.01$. Actual *P* values are also shown.

**Figure 7. Expression time-course of candidate transcription factors for *Aicda* regulatory regions.** Purified spleen B cells from 3 mice were pooled and stimulated with

indicated stimulants. Cells were harvested at indicated time points after the stimulation and mRNA expression of indicated proteins were measured by RT-qPCR. Signals were mormalized by GAPDH signals. Data representative of two independent experiments. a.u., arbitrary unit.

## Methods

### Cells, reagents and mice

CH12F3-2 cells, a subline of CH12F3 cells, were cultured and stimulated by a combination of 2.5 μg/ml IL-4, 1 μg/ml TGF-β, and CD40L as described[29]. Culture supernatant from J558 cells transfected with pHβApr-1neomCD40L-mCD8α plasmid was used as the source of CD40L at 40% volume of culture medium[29]. The NF-κB-specific inhibitor, sulfasalazine, was purchased from Sigma-Aldrich. The sequences of the siRNA oligonucleotides for *Stat6*, *Myb*, and *Nkx2-5* knockdown (Invitrogen) are indicated in **Supplementary Table 1**. Animal studies were approved by Animal Research Committee of Graduate School of Medicine, Kyoto University. Mice were kept in specific pathogen free condition. Spleen B cells were purified by magnetic sorting with CD43 beads (Miltenyi Biotec). B cells were cultured with combination of IL-4 (10 ng/ml), LPS (20 μg/ml) and TGF-β (10 ng/ml).

### Sequence alignment and identification of regulatory elements

The DNA sequence alignment and dotplot matrix sequence comparisons were performed using PipMaker[46] and the BLAST2 algorithms[47]. Identification of regulatory elements in the conserved region of the AID gene was conducted with the programs MOTIF search (http://motif.genome.jp/) and rVISTA[48].

**Construction of plasmids and PCR-mediated mutagenesis**

Region-1 to region-4, which contained putative *Aicda* regulatory elements, were amplified from genomic clones containing *Aicda* by PCR. The amplified region-1 fragment was digested with *Nco* I and *Xho* I and cloned into the pGL3-basic vector (Promega) upstream of the firefly luciferase coding region. The region-4 DNA was digested with *Xho* I and *Kpn* I and ligated upstream of region-1. Regions-2 and -3 were digested by *Bam* HI and *Sal* I and cloned downstream of the luciferase sequence. The relative positions and orientations of the regions were chosen according to their arrangement in the *Aicda* locus. Serial deletion mutants were generated by PCR with primers corresponding to each deletion site (**Supplementary Table 2**). Mutations at the putative binding sites of regulatory elements were generated by the QuickChange Site-Directed Mutagenesis kit (Stratagene). Point mutations introduced into each element are shown in **Supplementary Table 3.** All constructs were verified by DNA sequencing and restriction enzyme mapping.

**Transient reporter assay**

CH12F3-2 cells were transiently transfected with equimolar amounts of reporter plasmids, using the Amaxa nucleofection technology™ (Amaxa), according to the manufacturer's instructions. The general transfection efficiency of this procedure was more than 75%, as monitored by a GFP reporter plasmid. All transfections included 50 ng of phRL Renilla luciferase expression construct (Promega), driven by the CMV immediate early promoter, as an internal control. The cells were cultured for 24 h with or without stimulation before being harvested. Luciferase activity was measured using the Dual Luciferase Assay kit (Promega) by EnVision 2103 Multilabel Reader (PerkinElmer). The Firefly luciferase activities were normalized to the Renilla luciferase activity, and calculated as the fold

change from pGL3-basic (promoter-less control plasmid; Promega), without CIT stimulation.

**RNA extraction, semiquantitative RT-PCR.**

Total RNA was extracted by using TRIzol (Invitrogen) or RNeasy mini-kit (QIAGEN), according to the manufacturer's instructions. Complementary DNA was synthesized with the TaqMan RT-PCR kit (Applied Biosystems) and analyzed by ABI 7900HT real-time PCR machine (Applied Biosystems). In time course experiment, qPCR was done with LightCycler 480 thermal cycler using SYBER Green Supermix (Roche). The primers used are indicated in **Supplementary Table 4 and 5**.

**Chromatine immunoprecipitation assay and statistical analysis.**

Cells were fixed by 1% formaldehyde for 2 or 30 min at 22-26 º C . The cross-linking was terminated by 150 mM of glycine. After washing by ice cold PBS twice, cells were lysed by sonication in the SDS lysis buffer (1% SDS, 10 mM EDTA, 50 mM Tris, pH8.0). Debris were removed by centrifugation and cleared lysate was used for immuno-precipitation. Rabbit polyclonal antibodies for Stat6 (sc-981), NF-kB p65 (sc-109), Smad4 (sc-7145), and C/EBPβ (sc-150) were purchased from Santa Cruz biotechnology. DNA was purified from the precipitates and contents of each specific DNA locus were measured by real-time PCR (iQ sybr green supermix, Bio-Rad). Multiple results were evaluated by paired one-tail *t*-test. Primers used were: region-2a (R2a), 5′-CCACTTAATTACATCCTGAGCCC-3′ & 5′- CTATAAACCCAGAAGCAGCTCA-3′; region-2b (R2b), 5′- AGGGACATCCCGAGGAGAG-3′ & 5′-GAGCTGCACAGTCCTGACCT-3′; region-4#a (R4a), 5′-

AAATCGGGGAATGCAGAAGT-3′ & 5′-TCCTCGGGTCACTATTTTTGG-3’; region-4b (R4b), 5′-GGGACTGGAAAATGGTTTTTG-3′ & 5′-GATGCCGTACACATGCCATA-3′.

**Gene array analysis**

Naïve and GC B cells were sorted from B57BL/6 mice immunized with keyhole limpet hemocyanin by FACSAria (Becton Dickinson) according to the staining with B220, PNA and Fas. Total RNA was purified by Trizol (Invitrogen) and used for gene chip analysis (mouse genome 430 2.0 array; Affymetrix). Signals were normalized by Genespring GX 7.3 software with GC-RMA method and selected transcription factors were plotted.

**Accession number of gene array data.**

The micro-array data in this publication have been deposited in NCBI database with the accession number GSE18746.

**References**


1. **Muramatsu, M. *et al.* Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. *Cell* 102, 553-563 (2000).**
2. **Revy, P. *et al.* Activation-induced cytidine deaminase (AID) deficiency causes the autosomal recessive form of the Hyper-IgM syndrome (HIGM2). *Cell* 102, 565-575 (2000).**
3. **Honjo, T., Kinoshita, K. & Muramatsu, M. Molecular mechanism of class switch recombination: linkage with somatic hypermutation. *Annu Rev Immunol* 20, 165-196 (2002).**
4. **Stavnezer, J., Guikema, J.E. & Schrader, C.E. Mechanism and regulation of class switch recombination. *Annu Rev Immunol* 26, 261-292 (2008).**
5. **Petersen, S. *et al.* AID is required to initiate Nbs1/gamma-H2AX focus formation and mutations at sites of class switching. *Nature* 414, 660-665 (2001).**
6. **Doi, T. *et al.* The C-terminal region of activation-induced cytidine deaminase is responsible for a recombination function other than DNA cleavage in class switch recombination. *Proceedings of the National Academy of Sciences of the United States of America* 106, 2758-2763 (2009).**
7. **Nagaoka, H., Muramatsu, M., Yamamura, N., Kinoshita, K. & Honjo, T.**

**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. *The Journal of experimental medicine* 195, 529-534 (2002).**

8. **Begum, N.A. *et al.* De novo protein synthesis is required for activation-induced cytidine deaminase-dependent DNA cleavage in immunoglobulin class switch recombination. *Proceedings of the National Academy of Sciences of the United States of America* 101, 13003-13007 (2004).**
9. **Muramatsu, M. *et al.* Specific expression of activation-induced cytidine deaminase (AID), a novel member of the RNA-editing deaminase family in germinal center B cells. *The Journal of biological chemistry* 274, 18470-18476 (1999).**
10. **Gao, N., Dang, T. & Yuan, D. IFN-gamma-dependent and -independent initiation of switch recombination by NK cells. *J Immunol* 167, 2011-2018 (2001).**
11. **Okazaki, I.M., Kotani, A. & Honjo, T. Role of AID in tumorigenesis. *Advances in immunology* 94, 245-273 (2007).**
12. **Ramiro, A.R. *et al.* AID is required for c-myc/IgH chromosome translocations in vivo. *Cell* 118, 431-438 (2004).**
13. **Kotani, A. *et al.* Activation-induced cytidine deaminase (AID) promotes B cell lymphomagenesis in Emu-cmyc transgenic mice. *Proceedings of the National Academy of Sciences of the United States of America* 104, 1616-1620 (2007).**
14. **Kovalchuk, A.L. *et al.* AID-deficient Bcl-xL transgenic mice develop delayed atypical plasma cell tumors with unusual Ig/Myc chromosomal rearrangements. *The Journal of experimental medicine* 204, 2989-3001 (2007).**
15. **Pasqualucci, L. *et al.* AID is required for germinal center-derived lymphomagenesis. *Nature genetics* 40, 108-112 (2008).**
16. **Takizawa, M. *et al.* AID expression levels determine the extent of cMyc oncogenic translocations and the incidence of B cell tumor development. *The Journal of experimental medicine* 205, 1949-1957 (2008).**
17. **Feldhahn, N. *et al.* Activation-induced cytidine deaminase acts as a mutator in BCR-ABL1-transformed acute lymphoblastic leukemia cells. *The Journal of experimental medicine* 204, 1157-1166 (2007).**
18. **He, B., Raab-Traub, N., Casali, P. & Cerutti, A. EBV-encoded latent membrane protein 1 cooperates with BAFF/BLyS and APRIL to induce T cell-independent Ig heavy chain class switching. *J Immunol* 171, 5215-5224 (2003).**
19. **Machida, K. *et al.* Hepatitis C virus induces a mutator phenotype: enhanced mutations of immunoglobulin and protooncogenes. *Proceedings of the National Academy of Sciences of the United States of America* 101, 4262-4267 (2004).**
20. **Matsumoto, Y. *et al.* Helicobacter pylori infection triggers aberrant expression of activation-induced cytidine deaminase in gastric epithelium. *Nature medicine* 13, 470-476 (2007).**
21. **Endo, Y. *et al.* Expression of activation-induced cytidine deaminase in human hepatocytes via NF-kappaB signaling. *Oncogene* 26, 5587-5595 (2007).**
22. **Morisawa, T. *et al.* Organ-specific profiles of genetic changes in cancers caused by activation-induced cytidine deaminase expression. *International journal of cancer* 123, 2735-2740 (2008).**
23. **Okazaki, I.M. *et al.* Constitutive expression of AID leads to tumorigenesis. *The Journal of experimental medicine* 197, 1173-1181 (2003).**

**24. Rucci, F. *et al.* Tissue-specific sensitivity to AID expression in transgenic mouse models. *Gene* 377, 150-158 (2006).**
**25. Yadav, A. *et al.* Identification of a ubiquitously active promoter of the murine activation-induced cytidine deaminase (AICDA) gene. *Molecular immunology* 43, 529-541 (2006).**
**26. Sayegh, C.E., Quong, M.W., Agata, Y. & Murre, C. E-proteins directly regulate expression of activation-induced deaminase in mature B cells. *Nature immunology* 4, 586-593 (2003).**
**27. Gonda, H. *et al.* The balance between Pax5 and Id2 activities is the key to AID gene expression. *The Journal of experimental medicine* 198, 1427-1437 (2003).**
**28. Dedeoglu, F., Horwitz, B., Chaudhuri, J., Alt, F.W. & Geha, R.S. Induction of activation-induced cytidine deaminase gene expression by IL-4 and CD40 ligation is dependent on STAT6 and NFkappaB. *International immunology* 16, 395-404 (2004).**
**29. Nakamura, M. *et al.* High frequency class switching of an IgM+ B lymphoma clone CH12F3 to IgA+ cells. *International immunology* 8, 193-201 (1996).**
**30. Crouch, E.E. *et al.* Regulation of AID expression in the immune response. *The Journal of experimental medicine* 204, 1145-1156 (2007).**
**31. Park, S.R. *et al.* HoxC4 binds to the promoter of the cytidine deaminase AID gene to induce AID expression, class-switch DNA recombination and somatic hypermutation. *Nature immunology* (2009).**
**32. Suske, G. The Sp-family of transcription factors. *Gene* 238, 291-300 (1999).**
**33. Nagy, S.R. & Denison, M.S. Specificity of nuclear protein binding to a CYP1A1 negative regulatory element. *Biochemical and biophysical research communications* 296, 799-805 (2002).**
**34. DeGregori, J. & Johnson, D.G. Distinct and Overlapping Roles for E2F Family Members in Transcription, Proliferation and Apoptosis. *Curr Mol Med* 6, 739-748 (2006).**
**35. van Dijk, T.B. *et al.* A composite C/EBP binding site is essential for the activity of the promoter of the IL-3/IL-5/granulocyte-macrophage colony-stimulating factor receptor beta c gene. *J Immunol* 163, 2674-2680 (1999).**
**36. Mizuguchi, G. *et al.* c-Myb repression of c-erbB-2 transcription by direct binding to the c-erbB-2 promoter. *The Journal of biological chemistry* 270, 9384-9389 (1995).**
**37. Ganter, B. & Lipsick, J.S. Myb binding sites within the N-ras promoter repress transcription. *Oncogene* 15, 193-202 (1997).**
**38. Allen, R.D., 3rd, Kim, H.K., Sarafova, S.D. & Siu, G. Negative regulation of CD4 gene expression by a HES-1-c-Myb complex. *Molecular and cellular biology* 21, 3071-3082 (2001).**
**39. Frolov, M.V. & Dyson, N.J. Molecular mechanisms of E2F-dependent activation and pRB-mediated repression. *Journal of cell science* 117, 2173-2181 (2004).**
**40. Trimarchi, J.M. & Lees, J.A. Sibling rivalry in the E2F family. *Nature reviews* 3, 11-20 (2002).**
**41. Han, J.H. *et al.* Class switch recombination and somatic hypermutation in early mouse B cells are mediated by B cell and Toll-like receptors. *Immunity* 27, 64-75 (2007).**
**42. Ueda, Y., Liao, D., Yang, K., Patel, A. & Kelsoe, G. T-independent activation-induced cytidine deaminase expression, class-switch recombination, and antibody production by immature/transitional 1 B cells. *J Immunol* 178, 3593-**

**3601 (2007).**

**43. Gourzi, P., Leonova, T. & Papavasiliou, F.N. Viral induction of AID is independent of the interferon and the Toll-like receptor signaling pathways but requires NF-kappaB. *The Journal of experimental medicine* 204, 259-265 (2007).**

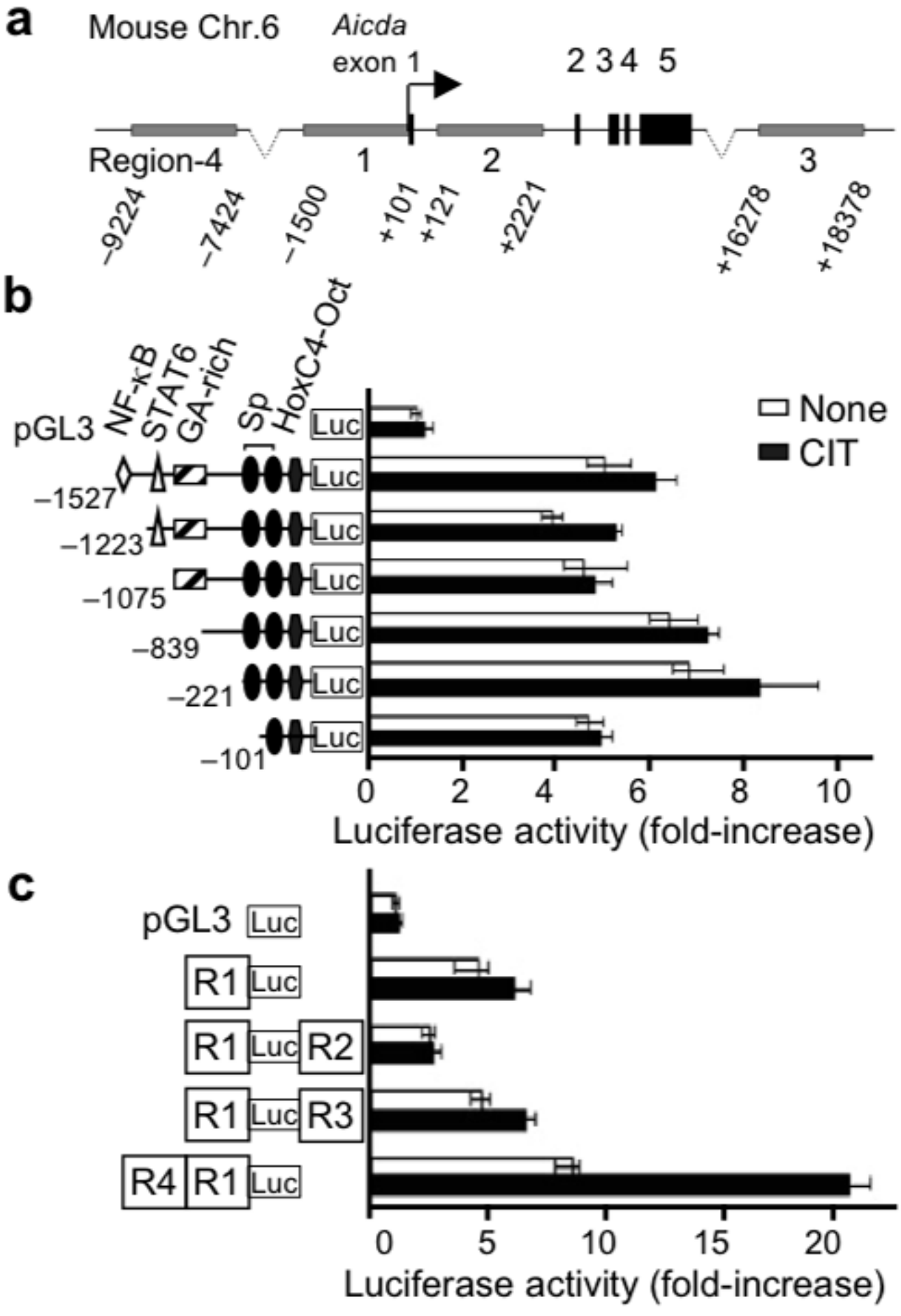
**44. Deb, A., Haque, S.J., Mogensen, T., Silverman, R.H. & Williams, B.R. RNA-dependent protein kinase PKR is required for activation of NF-kappa B by IFN-gamma in a STAT1-independent pathway. *J Immunol* 166, 6170-6180 (2001).**

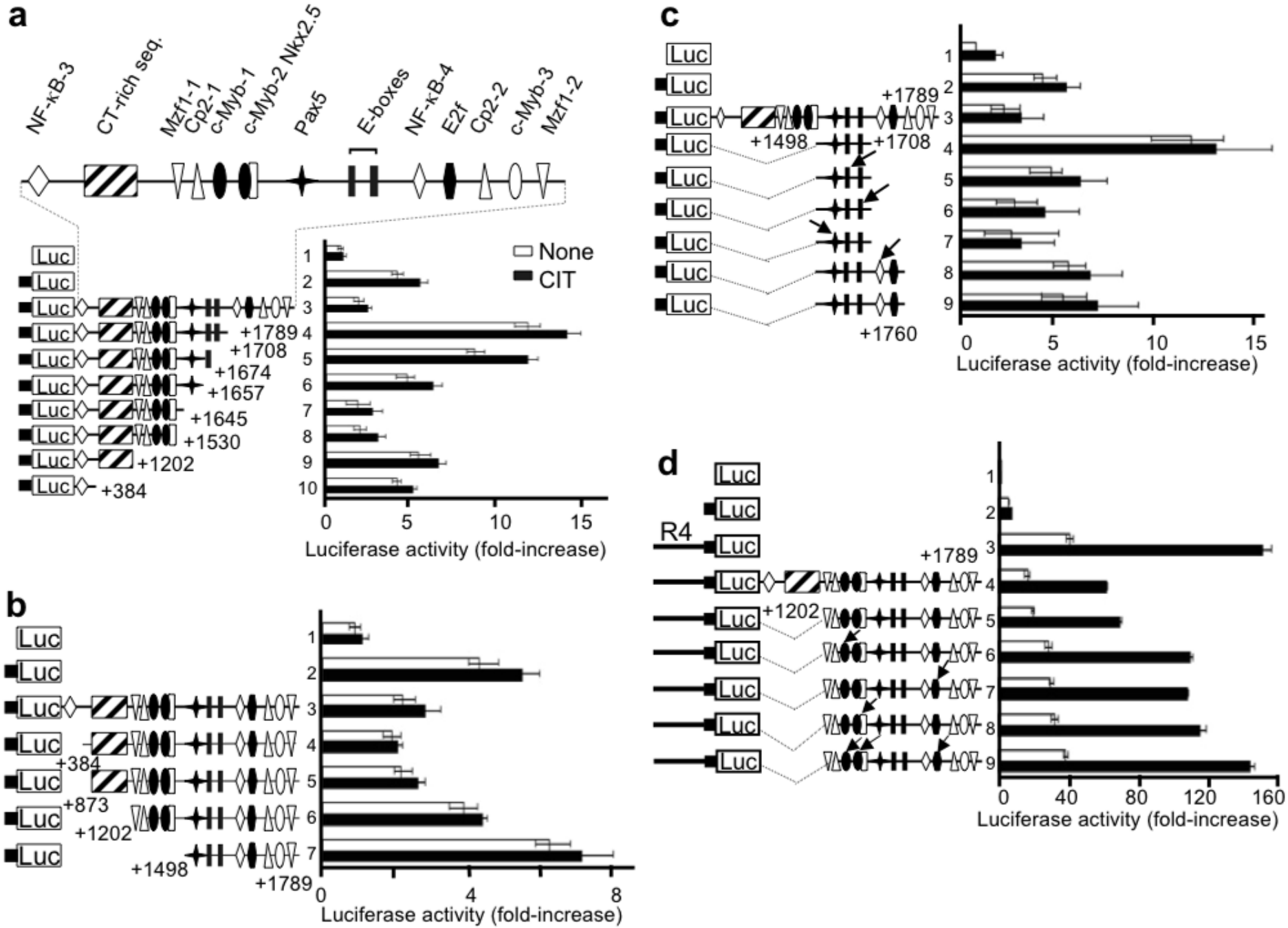
**45. Park, H.J., So, E.Y. & Lee, C.E. Interferon-gamma-induced factor binding to the interleukin-4-responsive element of CD23b promoter in human tonsillar mononuclear cells: role in transient up-regulation of the interleukin-4-induced CD23b mRNA. *Molecular immunology* 35, 239-247 (1998).**

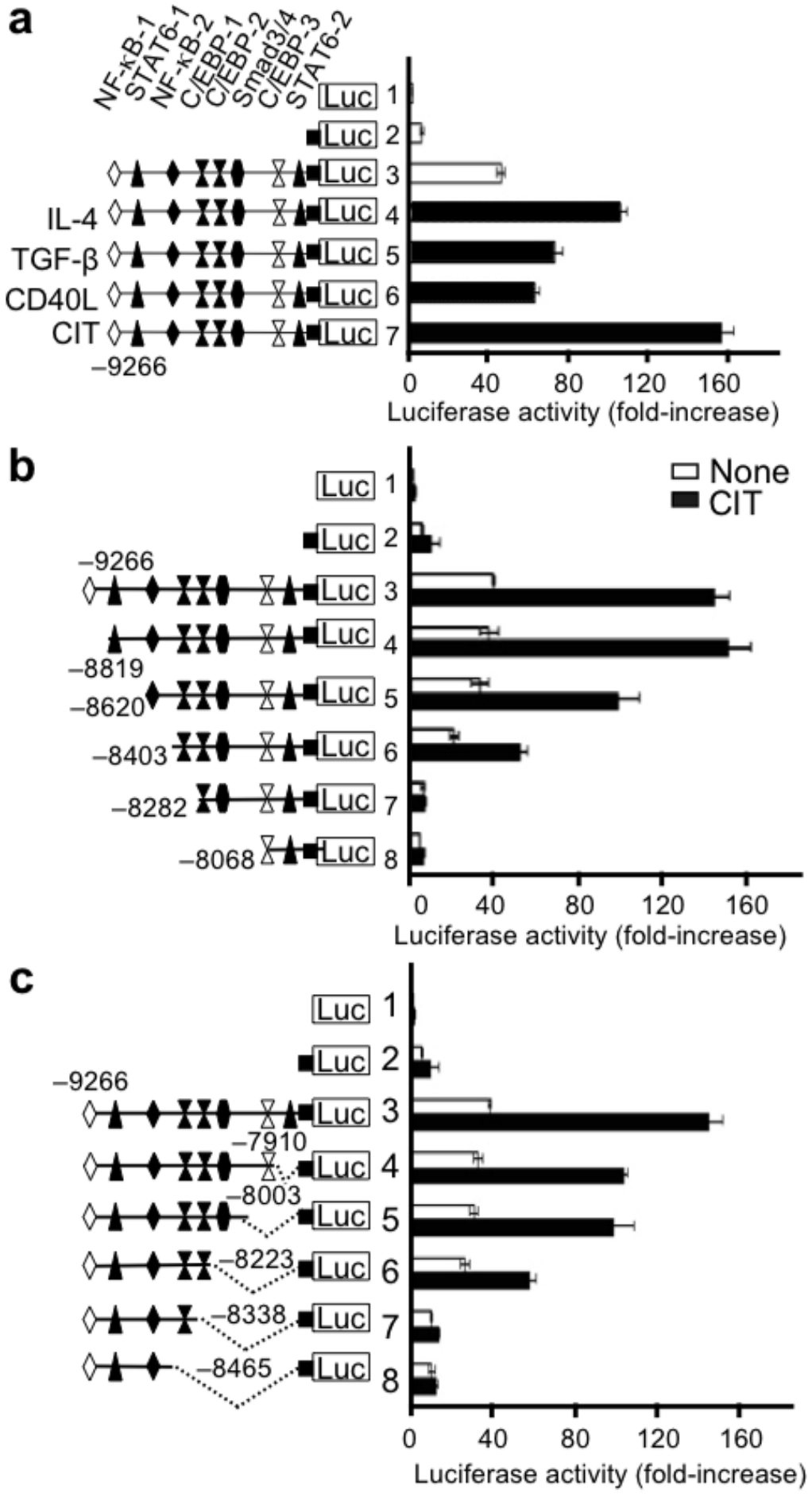
**46. Schwartz, S. *et al.* PipMaker--a web server for aligning two genomic DNA sequences. *Genome research* 10, 577-586 (2000).**

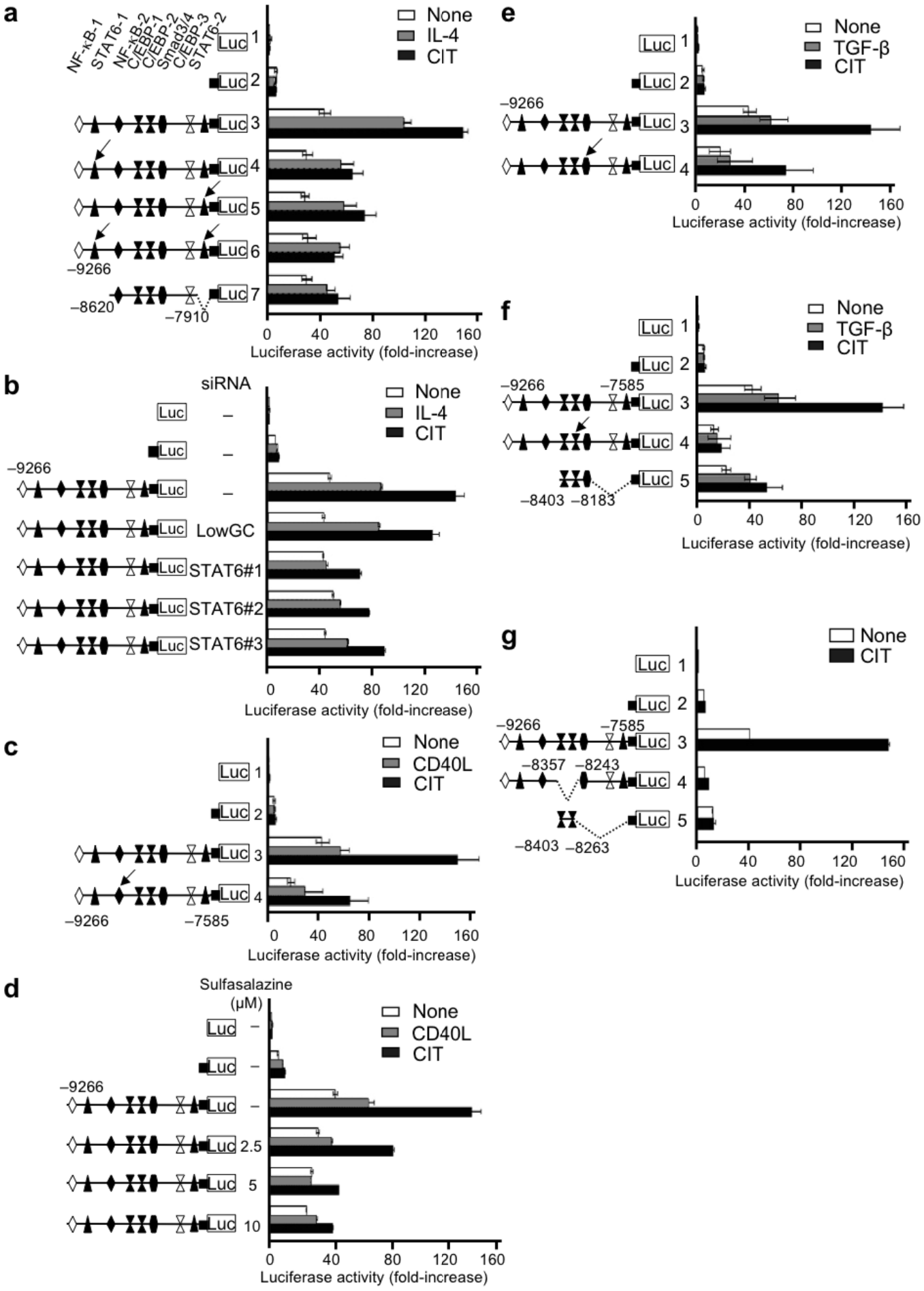
**47. Tatusova, T.A. & Madden, T.L. BLAST 2 Sequences, a new tool for comparing protein and nucleotide sequences. *FEMS microbiology letters* 174, 247-250 (1999).**

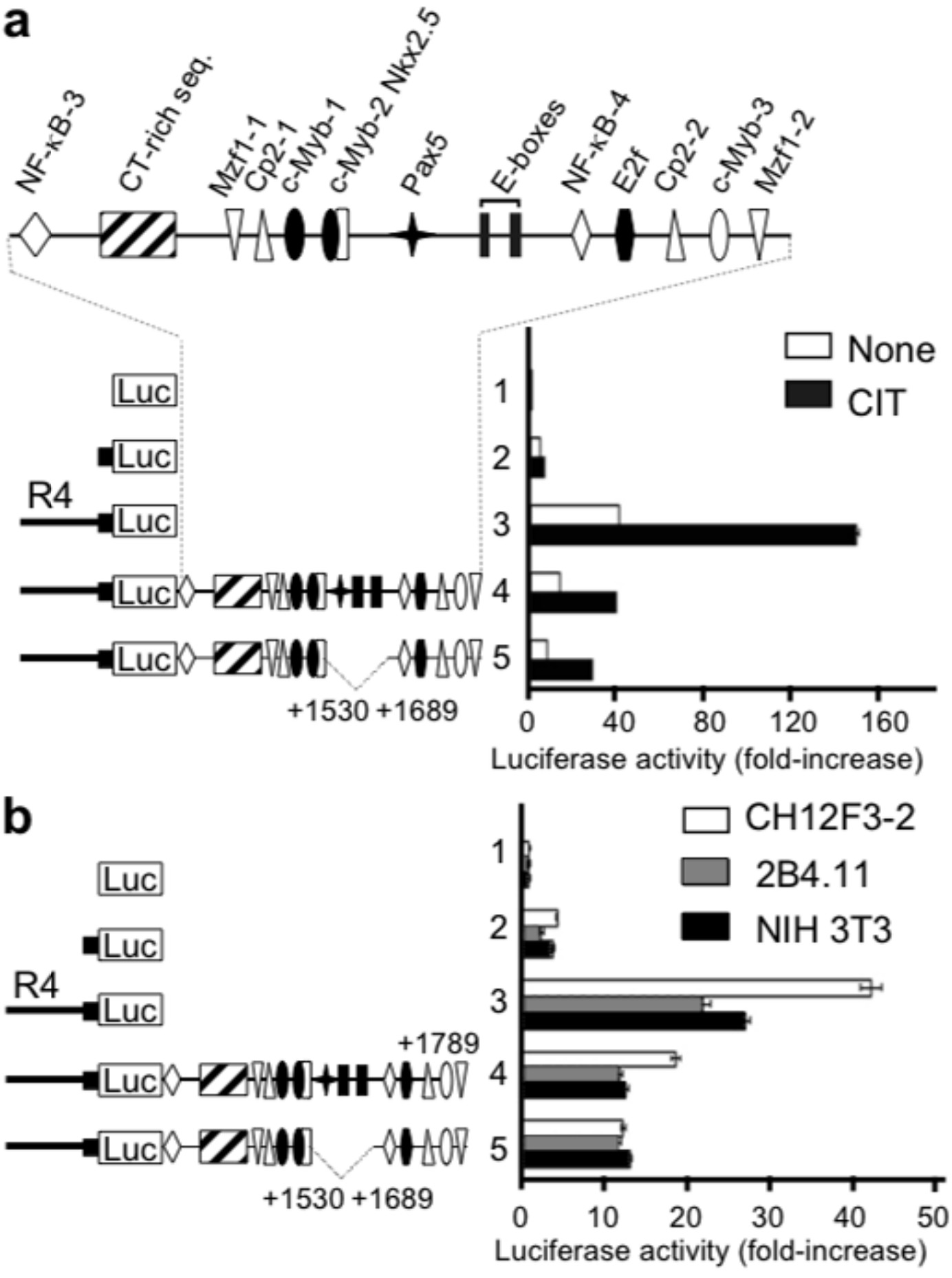
**48. Loots, G.G., Ovcharenko, I., Pachter, L., Dubchak, I. & Rubin, E.M. rVista for comparative sequence-based discovery of functional transcription factor binding sites. *Genome research* 12, 832-839 (2002).**

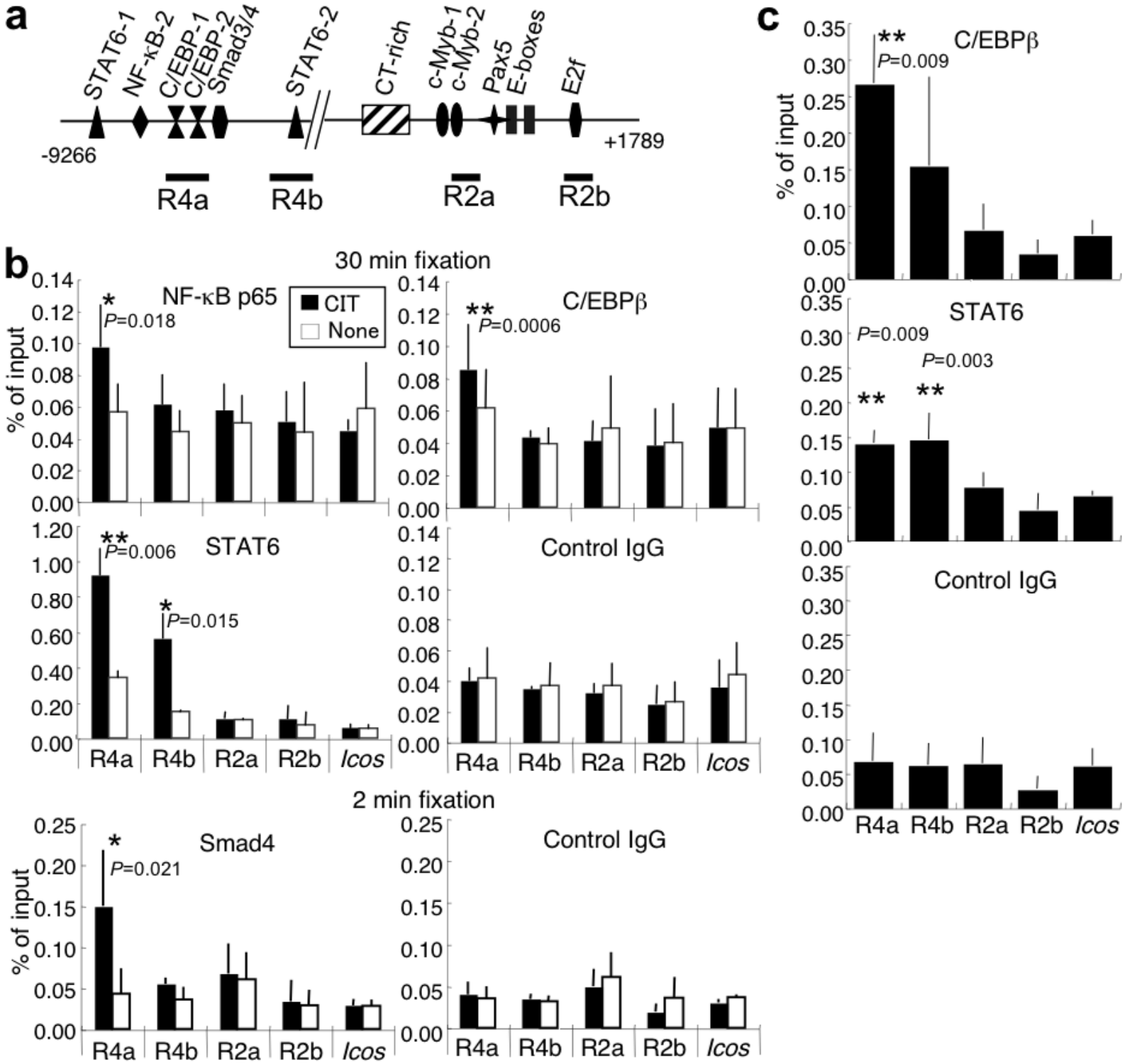

a
Mouse Chr.6
Aicda
exon 1
2 3 4 5
Region-4
1
2
3
−9224
−7424
−1500
+101
+121
+2221
+16278
+18378
b
NF-κB
STAT6
GA-rich
Sp
HoxC4-Oct
pGL3
Luc
−1527
−1223
−1075
−839
−221
−101
None
CIT
0
2
4
6
8
10
Luciferase activity (fold-increase)
c
pGL3
R1
R2
R3
R4
0
5
10
15
20
Luciferase activity (fold-increase)

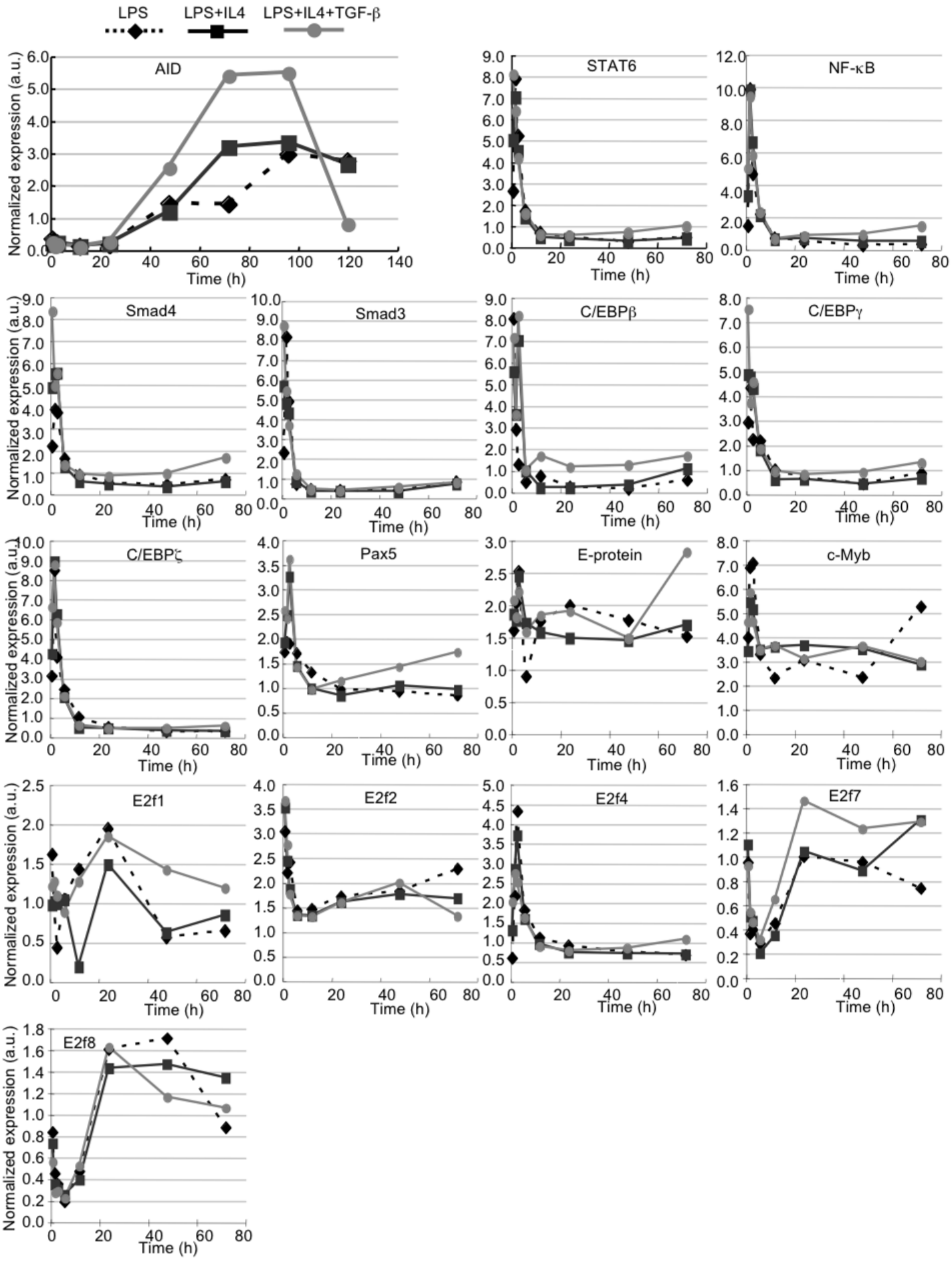

a
NF-κB-3
CT-rich seq.
Mzf1-1
Cp2-1
c-Myb-1
c-Myb-2 Nkx2.5
Pax5
E-boxes
NF-κB-4
E2f
Cp2-2
c-Myb-3
Mzf1-2
Luc
+1789
+1708
+1674
+1657
+1645
+1530
+1202
+384
None
CIT
Luciferase activity (fold-increase)
b
+384
+873
+1202
+1498
+1789
Luciferase activity (fold-increase)
c
+1789
+1498
+1708
+1760
Luciferase activity (fold-increase)
d
R4
+1789
+1202
Luciferase activity (fold-increase)


a
NF-κB-1
STAT6-1
NF-κB-2
C/EBP-1
C/EBP-2
Smad3/4
C/EBP-3
STAT6-2
Luc
1
2
3
IL-4
4
TGF-β
5
CD40L
6
CIT
7
–9266
0
40
80
120
160
Luciferase activity (fold-increase)
b
None
CIT
–9266
–8819
–8620
–8403
–8282
–8068
c
–7910
–8003
–8223
–8338
–8465


a
NF-κB-1
STAT6-1
NF-κB-2
C/EBP-1
C/EBP-2
Smad3/4
C/EBP-3
STAT6-2
Luc 1
Luc 2
Luc 3
Luc 4
Luc 5
Luc 6
Luc 7
−9266
−8620
−7910
None
IL-4
CIT
0 40 80 120 160
Luciferase activity (fold-increase)
b
siRNA
Luc –
Luc –
−9266
Luc –
Luc LowGC
Luc STAT6#1
Luc STAT6#2
Luc STAT6#3
None
IL-4
CIT
0 40 80 120 160
Luciferase activity (fold-increase)
c
Luc 1
Luc 2
Luc 3
Luc 4
−9266
−7585
None
CD40L
CIT
0 40 80 120 160
Luciferase activity (fold-increase)
d
Sulfasalazine (μM)
Luc –
Luc –
−9266
Luc –
Luc 2.5
Luc 5
Luc 10
None
CD40L
CIT
0 40 80 120 160
Luciferase activity (fold-increase)
e
Luc 1
Luc 2
−9266
Luc 3
Luc 4
None
TGF-β
CIT
0 40 80 120 160
Luciferase activity (fold-increase)
f
Luc 1
Luc 2
−9266
−7585
Luc 3
Luc 4
Luc 5
−8403 −8183
None
TGF-β
CIT
0 40 80 120 160
Luciferase activity (fold-increase)
g
Luc 1
Luc 2
−9266
−7585
Luc 3
−8357 −8243
Luc 4
Luc 5
−8403 −8263
None
CIT
0 40 80 120 160
Luciferase activity (fold-increase)

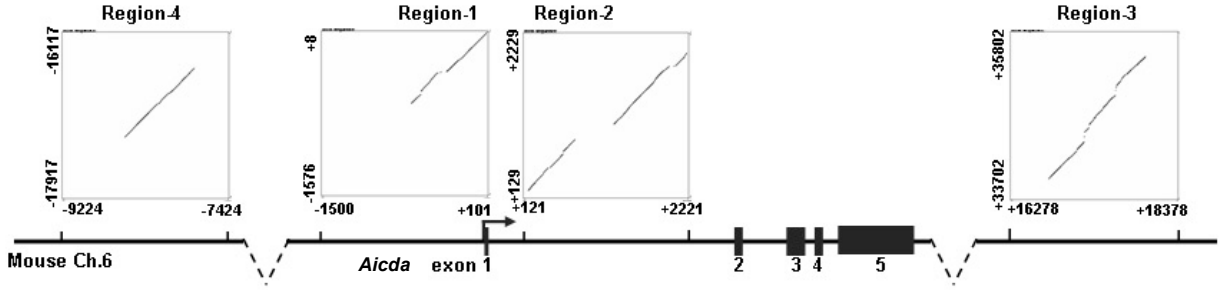

a
NF-κB-3
CT-rich seq.
Mzf1-1
Cp2-1
c-Myb-1
c-Myb-2 Nkx2.5
Pax5
E-boxes
NF-κB-4
E2f
Cp2-2
c-Myb-3
Mzf1-2
Luc
R4
+1530 +1689
None
CIT
1
2
3
4
5
0
40
80
120
160
Luciferase activity (fold-increase)
b
+1789
CH12F3-2
2B4.11
NIH 3T3
0
10
20
30
40
50
Luciferase activity (fold-increase)


a
STAT6-1
NF-κB-2
C/EBP-1
C/EBP-2
Smad3/4
STAT6-2
CT-rich
c-Myb-1
c-Myb-2
Pax5
E-boxes
E2f
-9266
+1789
R4a
R4b
R2a
R2b
b
30 min fixation
NF-κB p65
CIT
None
% of input
*
P=0.018
C/EBPβ
**
P=0.0006
STAT6
**
P=0.006
*
P=0.015
Control IgG
R4a
R4b
R2a
R2b
Icos
2 min fixation
Smad4
*
P=0.021
Control IgG
c
C/EBPβ
**
P=0.009
STAT6
P=0.009
P=0.003
**
**
Control IgG


LPS
LPS+IL4
LPS+IL4+TGF-β
AID
STAT6
NF-κB
Smad4
Smad3
C/EBPβ
C/EBPγ
C/EBPζ
Pax5
E-protein
c-Myb
E2f1
E2f2
E2f4
E2f7
E2f8
Normalized expression (a.u.)
Time (h)

Supplementary Figure 1

**a**




**b**

**Region-1**

```
 -1389      NF-κB-3      -1171      STAT6        -129        Sp
...GCTGGTTACTCCCGG...ATTTCACACAACA...ACAGCCCCGCCCATCT...
   |  |||    ||||       | |||  |||||       ||||||||  |||||||
...GTGGGTCTTTCCCAT...AGTTCCCACAATT...ACAGCCCCACCCATCT...
      GGGNNNTCCC         TTCNNNCAA          GCCCCGCCCC

   -70        Sp              -34    HoxC4-Oct          +91
...TGGGGGAGGAGCCCACA...CAGATTTGAATGCC...CGATATGGAC...
     |||||||||||||  |    ||  |||||||||       |   |||||||
...-AGGGGAGGAGCCCAAA...CAAATTTGAATGTG...CACTATGGAC...
     GGGGAGGAGCCC           ATTTGCAT
```

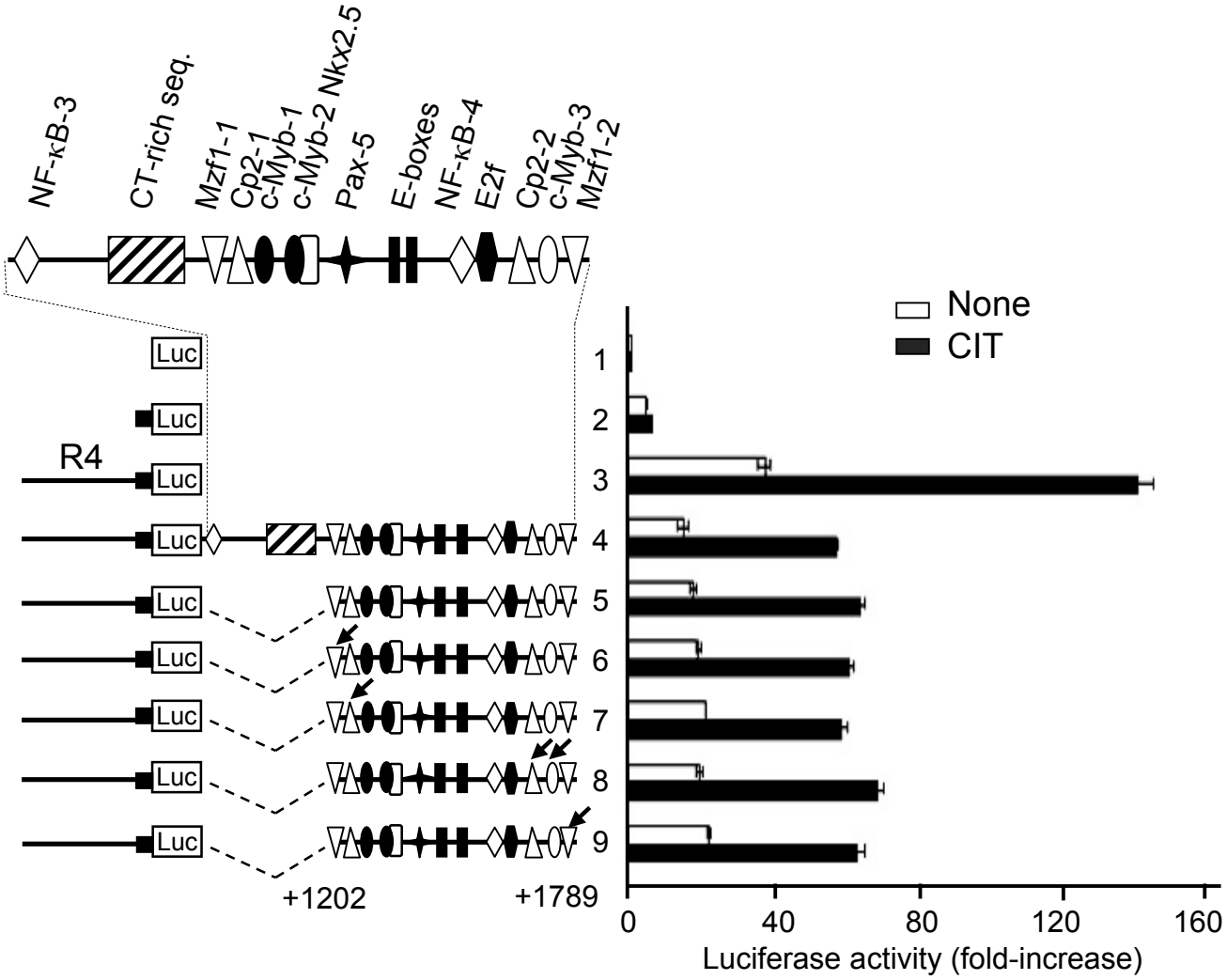
**Region-2**

```
 +178      NF-κB          +1208         Mzf1-1                    Cp2-1
...TGCGCGAATTTCCTCGT...CCCTCACCCCTATC-AGGGACCACATTGAAGACCTCACACATGC...
          ||||   |   |      |   ||||||  | ||  ||     |   ||     |||||  ||
...CATATATATTTTTTGAT...AACAGGCCCCTAACCAGAGAAGGGAAGTAATGGATCACAAATTA...
      GGGNNNTCCC              TCCCCNCN              NTWCAAAACCYNWKNN
                               T AA                  G   T   T G   T

 +1252   c-Myb-1  c-Myb-2/Nkx2.5  +1642      Pax5        +1659     E-box          E-box
...ACAAGTAATCTGCCACTTAATTA..TAAGGTTCTGCCTCA...AGCACAGCTGTCTC-AGCAGCTGGCGATC..
   |      || | ||    | ||||    || |||||||||||    ||||||||||||   ||||||||  |    |
...AGG-TAAAACAGCAGTTAAATT...TAGGGTTCTGCCTCA...AGCACAGCTGTCCAGAGCAGCTGTCAGCC..
       NNNAACKGNC               GGTCACGCC           CAGCTG        CAGCTG
              GNCKGTTNNN
                CACTTYA
         TCGGC     AGC              AG A TA           GT  AT        GT  AT

+1727      NF-κB-4                    E2f                 Cp2-2/c-Myb-3    Mzf1-2
...GGAGGGACATCCCGAGGAGAGGGAGCTGGAAATAGGCAGGCTCCTAGCTAGAGTTGAGGGGAGTG...
   |||   || ||   ||||  | |||||||||||||| ||||||| |||| | ||||    |||| |||
...GGACAGAAATGACGAGAACAGGGAGCTGGAAACAGGCAGGCCCCTAACCAGAG--AAGGGAAGTA...
      GGGNNNTCCC            GCTGGAAA             CTMGMTNMNGC
       TA     AT            CG    TGC             GNCKGTTNN
                                                        NGNGGGGA
                                                   GA C A     C CAA
```

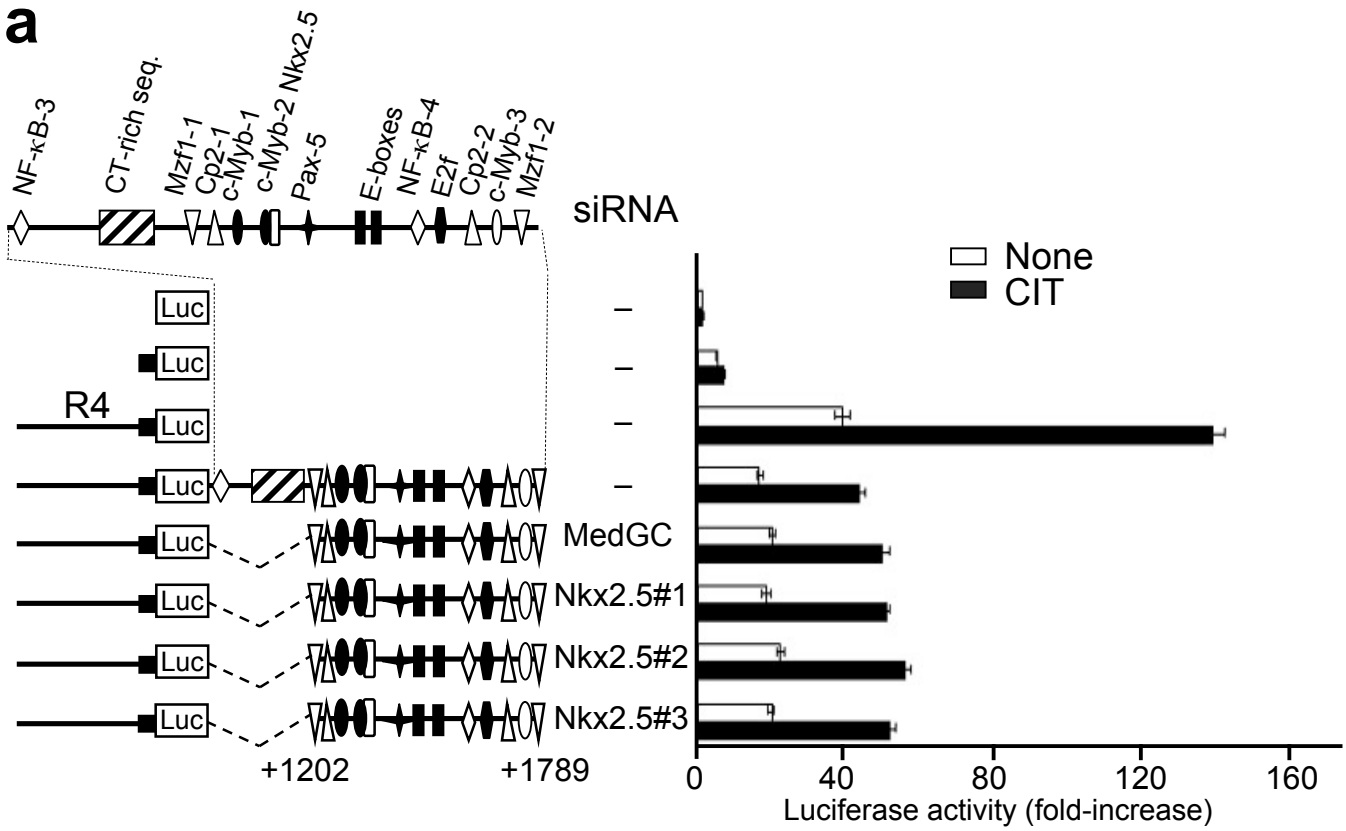
**Region-4**

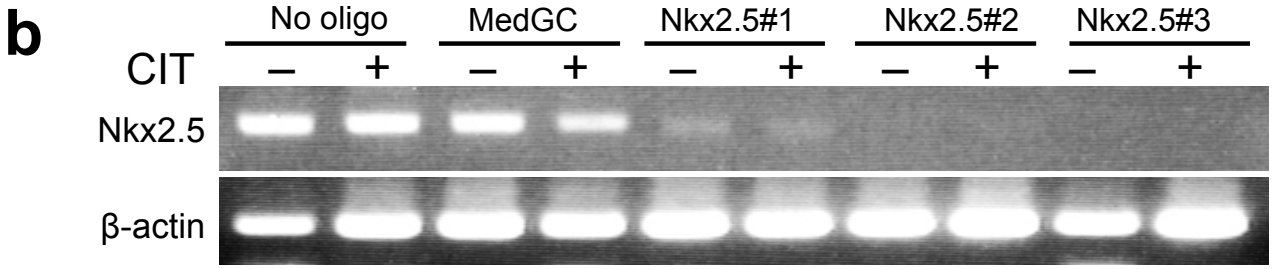
```
 -8973    NF-κB-1             -8636   STAT6-1       -8447    NF-κB-2       -8355   C/EBP-1
...GGTGTGGGGATGAGGGTGTA...TTTTTCTTTCAAACCT..TGAGGGCTGTCCCATG...TGAAGCAATCCAA...
   ||  ||||| ||  ||   |    ||||||||||   |   | | |||||||| ||      ||||||  |||||
...GGC-TGGGCGTGCTGGCTCA...CTTTTCTTTCATCAT...TAAAGGCTGTCCTATC...AGAAGCACTCCAA...
       TGGGGA                TTCnnnCAA          GGGnnnTCCC           GCAAT
                              CG      GC          CA     AT

 -8280   C/EBP-2      -8219    Smad3/4       -8002   C/EBP-3      -7823       STAT6-2
...TTGAGCAATCTAG...TCTGAGACACAGGAAT...GTTTTTCAATAGAA...TGGCTTTCTTACAAATCA-A...
   ||||  ||||||||     || ||||||||||||   |||||||||||||   ||  ||||  |||| ||  |
...TTGAACAATCTAG...-CTCAGACACAGGAAT...GTTTTTCAATAGAA...TGTATTTCCCACAATTCCCA...
       GCAAT            AGACACA             GCAAT            TTCnnnCAA
       CTCGC            CT  TA                                GC      GC
```

**Supplementary Figure 1. Conserved regions in the *Aicda* locus.** (**a**) Highly conserved regions between human (vertical) and mouse (horizontal) are designated as regions-1 to –4, as shown by dot-matrix plots in which the numbers indicate base positions relative to the mouse TSS: Region-1 (-1500 ~ +101), region-2 (+121 ~ +2221), region-3 (+16278 ~ +18378), and region-4 (-9224 ~ -7424). A schematic diagram of mouse *Aicda* is shown below. Boxes and an arrow indicate the exons and TSS, respectively. (**b**) DNA sequence alignment of transcription regulatory elements conserved between mouse (upper) and human (lower). Vertical lines indicate conserved nucleotides in the two species. Candidate elements are indicated by italic letters with the consensus binding motifs below. Mutations generated for the functional assays are shown below the consensus sequences. The numbers above the alignments indicate the nucleotide positions relative to the TSS. The underlined sequence is the ATG initiation codon.

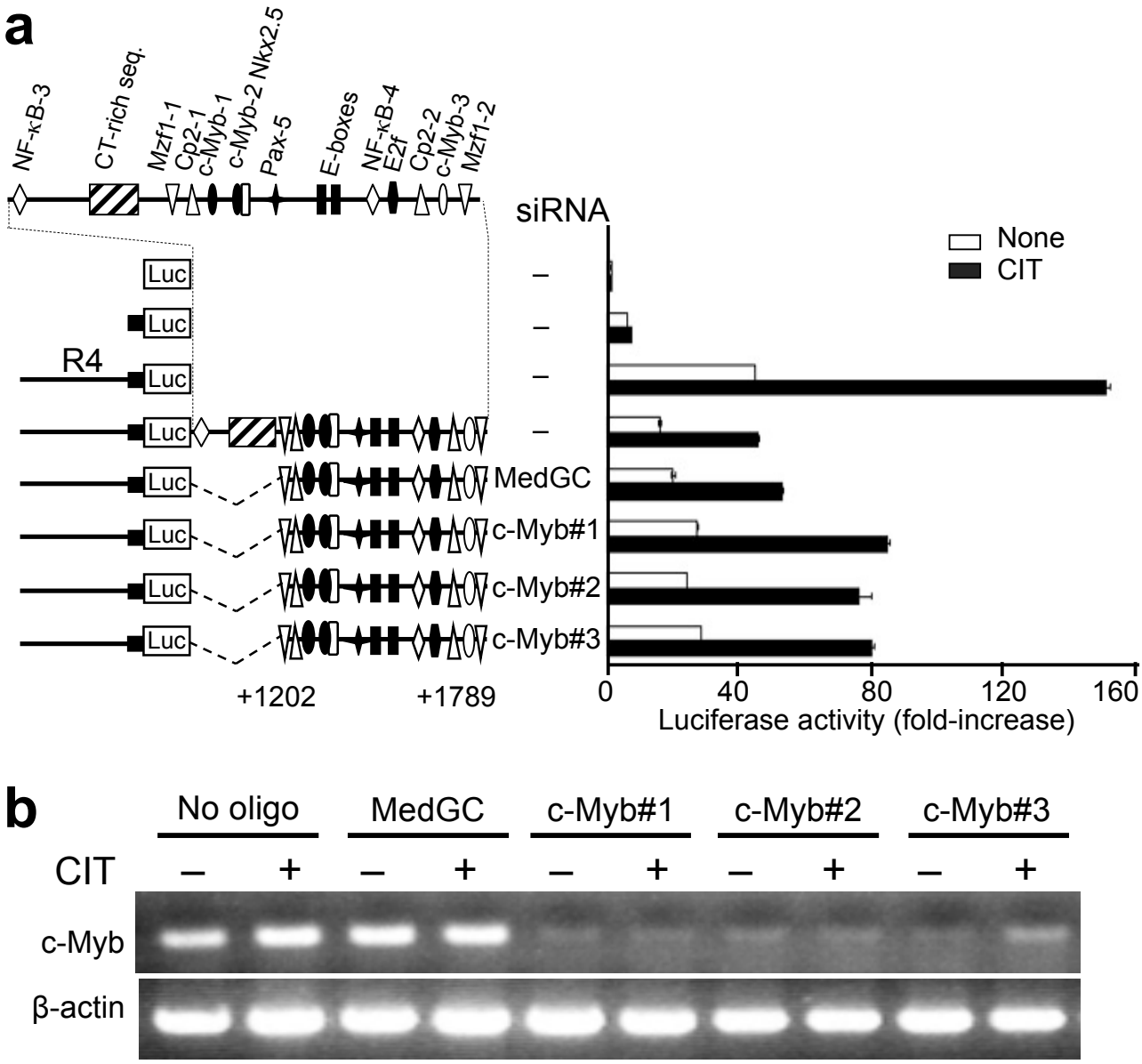



**Supplementary Figure 2. Mutagenesis of Mzf1, Cp-2 and c-Myb elements in region-2.** The luciferase construct with full length of region-4 and -2 (pR4-R2paidLuc) and its derivative plasmids with deletion and mutations as illustrated in the diagram were generated and analyzed. Mutated elements are indicated by arrows. Constructs were transiently transfected into CH12F3-2 cells and cultured for 24 h with no stimulant (open bars) or with CIT (filled bars). Data are representative of 3 independent experiment (mean +/– s.d.).

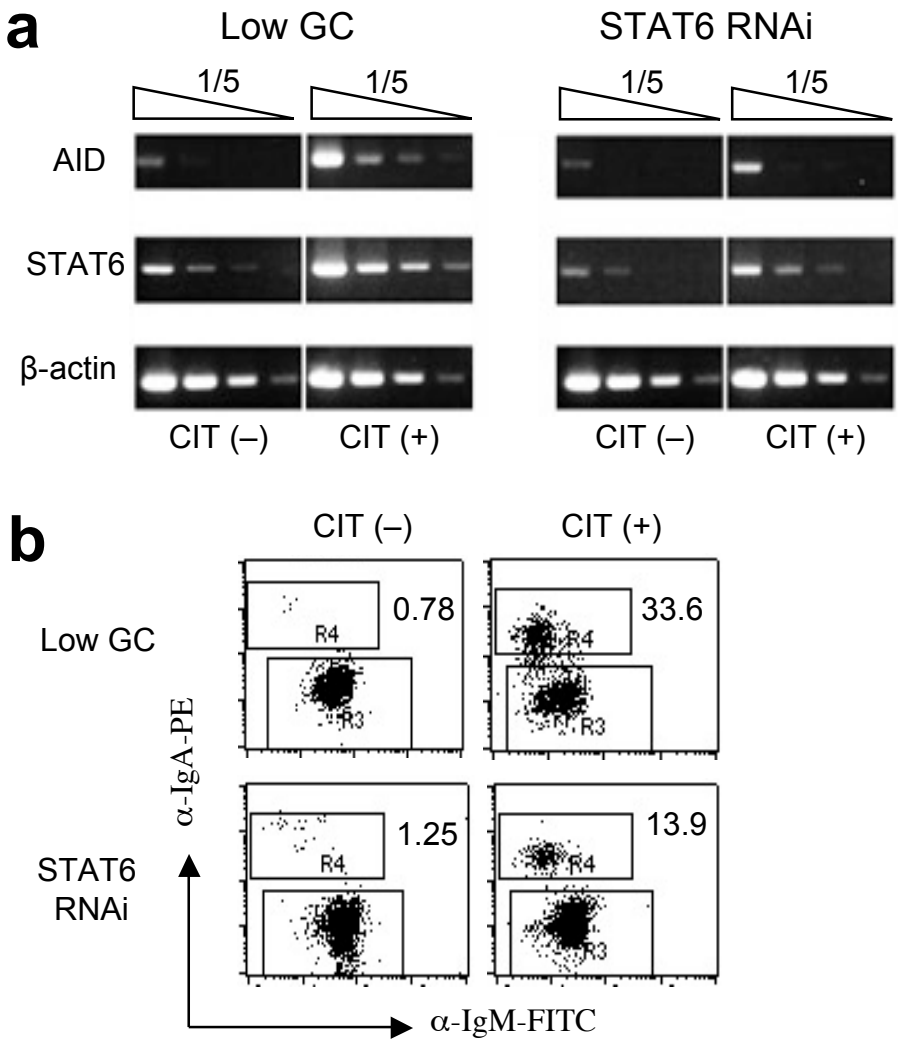


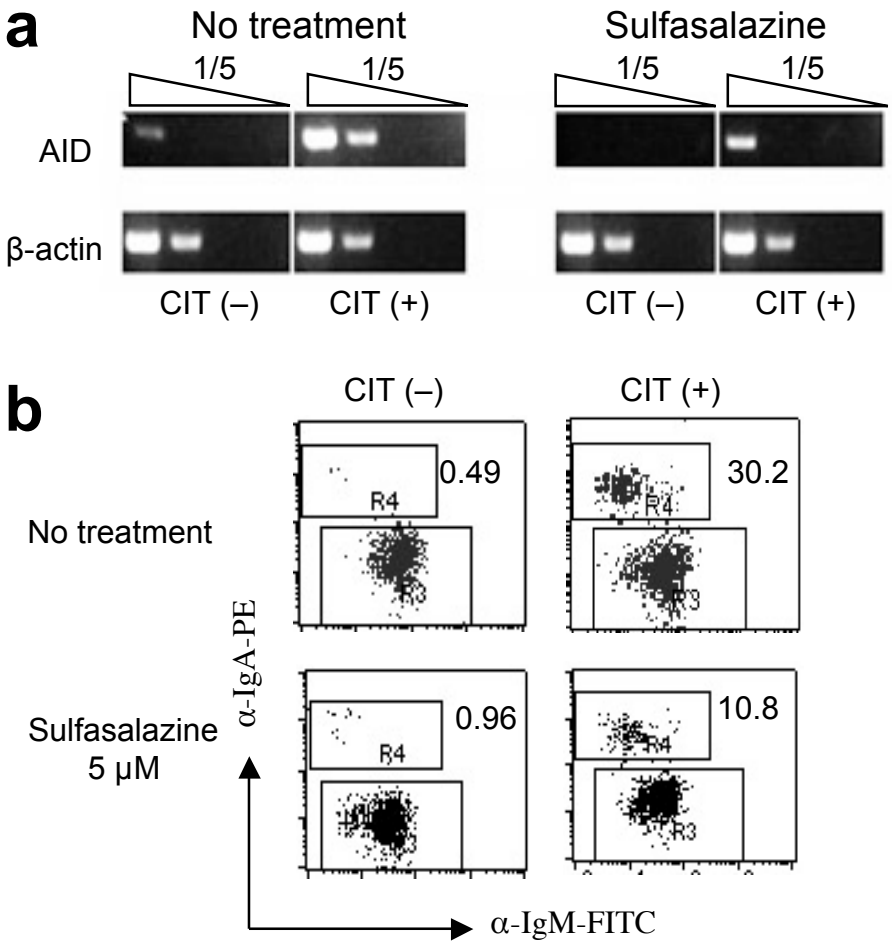

**Supplementary Figure 3. Minor role of Nkx2.5 in *Aicda* repression.** (**a**) RNAi oligonucleotides for Nkx2.5 (#1-#3) and the control med GC oligo were co-transfected into CH12F3-2 cells with the plasmid carrying region-4 and repressor sequence of region-2 as indicated. The activity was analyzed after 24 h culturing with no stimulant (open bars) or CIT (filled bars). Data are representative of three independent experiments (mean +/– s.d.). RNA oligonucleotide sequences are shown in **Supplementary Table 3**. (**b**) RT-PCR with Nkx2.5 specific primers indicated the suppression by RNAi oligos. Data are representative of two independent experiments.



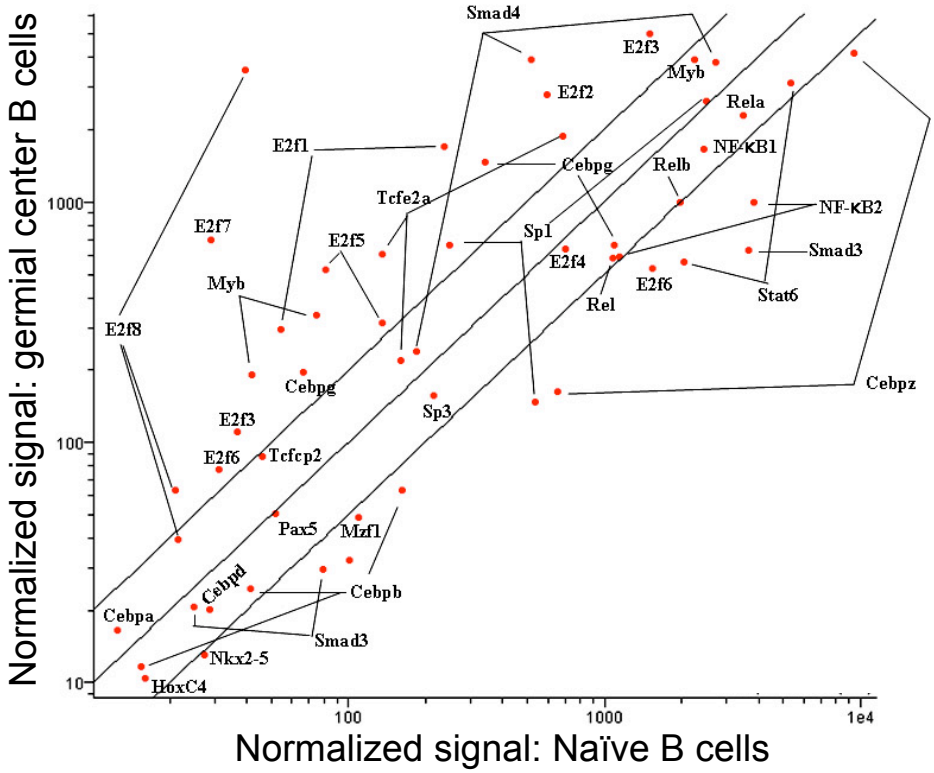

**Supplementary Figure 4. Knockdown of c-Myb reveals its suppressive function.** (**a**) A plasmid carrying region-4 and region-2 repressive elements was transfected into CH12F3-2 cells with three kinds of c-Myb RNAi oligonucleotide (#1-#3) as well as with the control med GC oligonucleotide as indicated. Transfected cells were cultured for 24 h with no stimulant (open bars) or CIT (filled bars). Data are representative of three independent experiment (mean +/– s.d.). RNA oligonucleotide sequences are shown in **Supplementary Table 3**. (**b**) c-Myb expression were evaluated by RT-PCR. Data are representative of two independent experiments.




**Supplementary Figure 5. STAT6 RNAi suppressed endogenous AID expression and IgA class switching in CH12F3-2 cells.** CH12F3-2 cells were transfected with STAT6 RNAi oligonucleotide (#1) or low GC oligonucleotide control and cultured for 24 h with or without CIT stimulation. (**a**) Total RNA was extracted and semi-quantitative RT-PCR was performed to evaluate endogenous AID and STAT6 expression. (**b**) CSR activity was assessed by percentage of IgA surface expression in STAT6 RNAi and low GC control transfected cells. The result shown in this figure is oligo #1; the other oligonucleotides (STAT6 #2 and #3) gave essentially the same result. The data representative of two independent experiments.



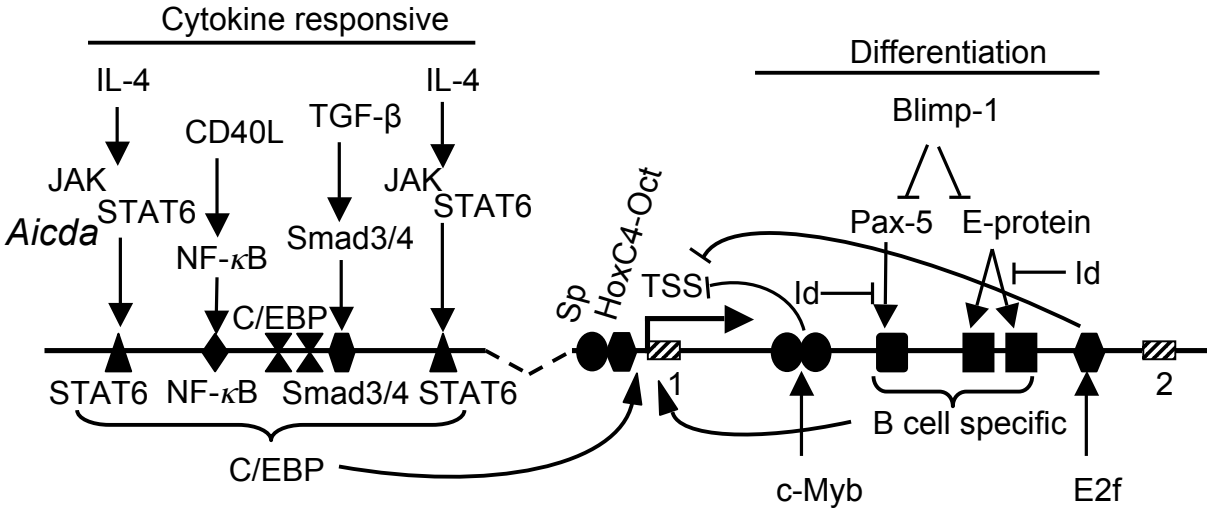

**Supplementary figure 6. Effect of sulfasalazine to endogenous AID expression and IgA class switching in CH12F3-2 cells.** CH12F3-2 cells were cultured with NF-κB specific inhibitor, sulfasalazine, at 5 μM for 24 h with or without CIT stimulation. (**a**) Total RNA was extracted and the expression of AID was evaluated by semi-quantitative RT-PCR. (**b**) FACS profile shows the CSR activity by percentage of IgA surface expression. Data are representative for two independent experiments.




Summary of the differential expression plot:
expression in GC cells relative to naïve cells.

| | |
|---|---|
| Up | E2f-1, E2f-2, E2f-3, E2f-5, E2f-7, E2f-8, E2a, Myb, C/EBPγ |
| No change | Rel, RelA, Relb, NF-κB-1, NF-κB-2, E2f-4, E2f-6, Sp1, Sp3, Smad4,STAT6, Pax5, C/EBPβ |
| Down | Smad3, C/EBPζ |

**Supplementary Figure 7. The mRNA expression of transcription factors in naive and GC B cells.** Expression of transcription factors in naive and GC B cells were analyzed by Affymetrix mouse genome 430 2.0 gene chip array. Data from two independent gene chip hybridization results for each condition were compiled. Two dimensional plot of normalized signal from each probe spot were shown. For the summary table, the expression pattern of the candidate factors were categorized. A plot within two diagonal lines of double-fold to both direction was judged as No change. Because generally more than one values were obtained to each genes from multi-probes, genes that showed the shift in more than half of plots were judged as Up or Down.




**Supplementary Figure 8. Proposed model for the transcriptional regulation of *Aicda* expression in B cells.** The cytokine-responsive and differentiation-dependent regions in the *Aicda* locus are illustrated. Hatched boxes indicate *Aicda* exons. Functional elements in each region are indicated with symbols. Binding of c-Myb and E2f independently repress AID expression. When B cells are activated by appropriate stimuli for inducing the binding of STAT6, Smad3/4, C/EBP, or NF-κB, AID expression is derepressed. B cell–specific factors, namely Pax5 and E protein, contribute to restricting the induction of AID to B lineage cells.

**Supplementary Table 1. RNAi oligos used in the knockdown assay.**

| | | |
|---|---|---|
| c-Myb #1 | Sense | UUAACAUGCAAUGCGACAGGAUAGG |
| RNAi_stealth_275871 | Anti-sense | CCUAUCCUGUCGCAUUGCAUGUUAA |
| c-Myb #2 | Sense | UUGGCAAUAACAGACCAACGCUUCG |
| RNAi_stealth_206853 | Anti-sense | CGAAGCGUUGGUCUGUUAUUGCCAA |
| c-Myb #3 | Sense | UUCUAAGAGGUUCUUAACAUUGUCC |
| RNAi_stealth_206854 | Anti-sense | GGACAAUGUUAAGAACCUCUUAGAA |
| Nkx2.5 #1 | Sense | UUAUCCGCCCGAGGGUCUUUGGCUG |
| RNAi_stealth_276028 | Anti-sense | CAGCCAAAGACCCUCGGGCGGAUAA |
| Nkx2.5 #2 | Sense | UACCGCUGUUGCUUGAAGCGCCGCU |
| RNAi_stealth_276029 | Anti-sense | AGCGGCGCUUCAAGCAACAGCGGUA |
| Nkx2.5 #3 | Sense | UAGGCAUUGAGACCCACGCCGUAGG |
| RNAi_stealth_276030 | Anti-sense | CCUACGGCGUGGGUCUCAAUGCCUA |
| STAT6 #1 | Sense | UGUUACUGAUCUCUCCUGUGCUCUC |
| RNAi_stealth_1288 | Anti-sense | GAGAGCACAGGAGAGAUCAUGAACA |
| STAT6 #2 | Sense | UUACCAUGCACGAUGACCACCAAGG |
| RNAi_stealth_1511 | Anti-sense | CCUUGGUGGUCAUCGUGCAUGGUAA |
| STAT6 #3 | Sense | UAAUAAAGCCAAUGAUCAGCCGAUC |
| RNAi_stealth_1888 | Anti-sense | GAUCGGCUGAUCAUUGGCUUUAUUA |

**Supplementary Table 2. Primers used for generating deletions in conserved regions.** Serial deletions in the conserved regions were generated both from 3'- and 5'-end. A restriction endonuclease site was inserted to each terminus of primer for cloning. Underline sequence were indicated for restriction sites.

| | | |
|---|---|---|
| R1-del5'-end | Sense | CCGCTCGAGACCTAACTCTGACCACGAAG |
| | | CCGCTCGAGCCCCTCAACTCTAGCTAGGA |
| | | CCGCTCGAGTAGGTGGTGTCCCAAACCTG |
| | | CCGCTCGAGAGAACAGCTGTGCTCCTTG |
| | | CCGCTCGAGCTTGAGGCAGAACCTTAGCTG |
| | | CCGCTCGAGCTTAGCTGCTTCCCCAGCAA |
| | | CCGCTCGAGCTCCTGTGTCTCTCCACTGTC |
| | | CCGCTCGAGGGAATCTGAAGAACACCCTGC |
| | Anti-sense | ATGCCATGGAGGAGTTGCTACGACCTCTT |
| R2-del5'-end | Sense | CGCGGATCCACCTAACTCTGACCACGAAG |
| | | CGCGGATCCCCCCTCAACTCTAGCTAGGA |
| | | CGCGGATCCTAGGTGGTGTCCCAAACCTG |
| | | CGCGGATCCAGAACAGCTGTGCTCCTTG |
| | | CGGGATCCCTTGAGGCAGAACCTTAGCTG |
| | | CGCGGATCCTTAGCTGCTTCCCCAGCAA |
| | | CGCGGATCCTCCTGTGTCTCTCCACTGTC |
| | | CGGGATCCGGAATCTGAAGAACACCCTGC |
| | Anti-sense | CGCGGATCCAGGAGTTGCTACGACCTCTT |
| R2-del3'-end | Sense | CGCGGATCCACCTAACTCTGACCACGAAG |
| | Anti-sense | CGCGGATCCCCCCTCAACTCTAGCTAGGA |
| | | CGCGGATCCTAGGTGGTGTCCCAAACCTG |
| | | CGCGGATCCAGAACAGCTGTGCTCCTTG |
| | | CGGGATCCCTTGAGGCAGAACCTTAGCTG |
| | | CGCGGATCCTTAGCTGCTTCCCCAGCAA |
| | | CGCGGATCCTCCTGTGTCTCTCCACTGTC |
| | | CGGGATCCGGAATCTGAAGAACACCCTGC |
| | | CGCGGATCCAGGAGTTGCTACGACCTCTT |
| R4-del5'-end | Sense | CGGGGTACCCACCACTCTCGGCTGTTCTG |
| | | CGGGGTACCTGTATATGGGTGTGGGGATGA |
| | | CGGGGTACCGCTGGACTTGAATTCAACAC |
| | | CGGGGTACCTGTCCCATGGCTTTTCATACC |
| | | CGGGGTACCATTGAGCAATCTAGCTGGTGT |
| | | CGGGGTACCCAGGCGACTCGATTTCCAAG |
| | Anti-sense | CCGCTCGAGACTCTTCGGGCCAATGAGATG |
| R4-del3'-end | Sense | CGGGGTACCCACCACTCTCGGCTGTTCTG |
| | Anti-sense | CCGCTCGAGACTCTTCGGGCCAATGAGATG |
| | | CCGCTCGAGAACTGGTACAGCGGCAAAG |
| | | CCGCTCGAGACCACAAGCCAACTTGGTC |
| | | CCGCTCGAGCAGCACACTCAGACATGACTC |
| | | CCGCTCGAGAAGTTCTGCATTCCCCGATT |
| | | CCGCTCGAGATTACTCAGCTTCTGCTGTTG |

**Supplementary Table 3. Primers used to introduce point mutations in the putative binding site of regulatory elements.** Mutation was introduced in the consensus sequence of each motif. Underline italic letters indicated for substitution nucleotides.

| | | |
|---|---|---|
| R4-mut.Stat6-1 | Sense | TGAAAATTCTTTTT*CG*CTTTC*GC*CCTGCCTGCTGGCC |
| | Anti-sense | GGCCAGCAGGCAGG*GC*GAAAG*CG*AAAAAGAATTTTCA |
| R4-mut.Stat6-2 | Sense | TGAGAGAATTTGGCT*GC*CTTAC*GC*ATCAAGGGTTTTGT |
| | Anti-sense | ACAAAACCCTTGAT*GC*GTAAG*GC*AGCCAAATTCTCTCA |
| R4-mut.NF-kB-2 | Sense | TCTGTGTCTGAG*CA*CTGT*AT*CATGGCTTTTCATACC |
| | Anti-sense | TATGAAAAGCCATG*AT*ACAG*TG*CTCAGACACAGACC |
| R4-mut.C/EBP-2 | Sense | AACCAGTATCATATTGA*CTCGC*CTAGCTGGTGTCCAAA |
| | Anti-sense | TTTGGACACCAGCTAG*GCGAG*TCAATATGATACTGGTT |
| R4-mut.Smad3/4 | Sense | TCTGAGTGTGCTGGCTCTG*CT*AC*TA*AGGAATGCAGCC |
| | Anti-sense | GGCTGCATTCCT*TA*GT*AG*CAGAGCCAGCACACTCAGA |
| R2-mut.Ebox-1 | Sense | TCTGCCTCAAGGAGCA*GT*GC*AT*TCTCAGCAGCTGGCGAT |
| | Anti-sense | ATCGCCAGCTGCTGAGA*AT*GC*AC*TGCTCCTTGAGGCAGA |
| R2-mut.Ebox-2 | Sense | AGCACAGCTGTCTCAG*GT*GC*AT*GCGATCTACAGGTTTG |
| | Anti-sense | CAAACCTGTAGATCGC*AT*GC*AC*CTGAGACAGCTGTGCT |
| R2-mut.Pax-5 | Sense | TGGGGAAGCAGCTAAG*AG*T*A*T*TA*CTCAAGGAGCACAGCT |
| | Anti-sense | AGCTGTGCTCCTTGAG*TA*A*T*A*CT*CTTAGCTGCTTCCCCA |
| R2-mut.c-Myb-1 | Sense | CACATGCTAGACAAGT*TCGGC*GCCACTTAATTACATCCT |
| | Anti-sense | AGGATGTAATTAAGTGGC*GCCGA*ACTTGTCTAGCATGTG |

| | | |
|---|---|---|
| R2-mut.c-Myb-1&<br>Nkx2.5 | Sense<br>Anti-sense | CTAGACAAGTAACTGCCA*AGC*TTAATTACATCCTGAGCC<br>GGCTCAGGATGTAATTAA*GCT*TGGCAGTTACTTGTCTAG |
| R2-mut.E2f | Sense<br>Anti-sense | ATCCCGAGGAGAGGGA*CG*TGG*TGC*TAGGCTCCTAGCTAGA<br>TCTAGCTAGGAGCCTA*GCA*CCA*CG*TCCCTCTCCTCGGGAT |
| R2-mut.Cp2-1 | Sense<br>Anti-sense | ATCAGGGACCACA*G*TG*T*AG*T*C*G*TC*T*CACATGCTAGACAAG<br>CTTGTCTAGCATGTG*A*GA*C*G*A*CT*A*CA*C*TGTGGTCCCTGAT |
| R2-mut.Cp2-2&<br>C-Myb-3 | Sense<br>Anti-sense | TGCAAATAGGCAGGCTCCT*GA*C*C*A*A*AGTTGAGGGGAGT<br>ACTCCCCTCAACT*T*T*G*G*TC*AGGAGCCTGCCTATTTGCA |
| R2-mut.Mzf1-1 | Sense<br>Anti-sense | CGTTCAACTCTACCCTCA*T*C*AA*TATCAGGGACCACATT<br>AATGTGGTCCCTGATA*TT*G*A*TGAGGGTAGAGTTGAACG |
| R2-mut.Mzf1-2 | Sense<br>Anti-sense | GCAGGCTCCTAGCTGAGTT*C*A*CAA*GAGTGCTGGATGGA<br>TCCATCCAGCACTC*TTG*T*G*AACTCAGCTAGGAGCCTGC |
| R2-mut.NF-kB-4 | Sense<br>Anti-sense | TCCTCCTTACCGGGAG*TA*ACAT*TA*CGAGGAGAGGGAGCT<br>AGCTCCCTCTCCTCG*TA*ATGT*TA*CTCCCGGTAAGGAGGA |

**Supplementary Table 4. Primers used for RT-PCR analysis.**

| | | |
|---|---|---|
| AID | Sense<br>Anti-sense | AGATAGTGCCACCTCCTGCTCACTGG<br>GGCTGAGGTTAGGGTTCCATCTCAG |
| Stat6 | Sense<br>Anti-sense | GGCCTAATTTCCAAGATGTCC<br>ACACTTGTCCAGTCTTAGCC |
| c-Myb | Sense<br>Anti-sense | CTTTCGACACATGGCTCCTC<br>AAAAATGCACTTGGTGCTGC |
| β-actin | Sense<br>Anti-sense | CCTTCCTGGGCATGGAGTCCTG<br>GGAGCAATGATCTTGATCTTC |

**Supplementary Table 5. Primers used for RT-qPCR**

| | | |
|---|---|---|
| AID | Sense | CGTGGTGAAGAGGAGAGATAGTG |
| | Anti-sense | CAGTCTGAGATGTAGCGTAGGAA |
| STAT6 | Sense | TGGCTGACTGGCTGGAGAG |
| | Anti-sense | CTGAAGACGCTGGACTGTGG |
| NF-Kb1 | Sense | GCAACAGATGGGCTACACAGAGG |
| | Anti-sense | AGGAAGACGAGAGAGGCAGACAG |
| Smad3 | Sense | CGAGAAGGCGGTCAAGAG |
| | Anti-sense | CAGGCGGCAGTAGATAACG |
| Smad4 | Sense | GTCTGTCTGCTGCTGCTG |
| | Anti-sense | GCTCGGTGAAGGTGAATCTC |
| Pax5 | Sense | AGTCTCCAGTGCCGAATG |
| | Anti-sense | TCCGTGGTGGTGAAGATG |
| c-Myb | Sense | CGGACGGACTGATAATGC |
| | Anti-sense | GCTGAGATGGAGGTGAGG |
| CEBPb | Sense | CGACGAGTACAAGATGCG |
| | Anti-sense | TGCTCCACCTTCTTCTGC |
| CEBPg | Sense | TTACAGCAGGTTCCTCAGC |
| | Anti-sense | GCGGTATTCGTCACTATTCC |
| CEBPz | Sense | ACCTGTGGCGTTCCTTGC |
| | Anti-sense | TCCTCCTCGTCTTCCTCATCC |
| E2f1 | Sense | GGAAAGGGAGAGGGAGAC |
| | Anti-sense | CATAGGAAGGACGCATACC |
| E2f2 | Sense | GGCAACTTCAAGGAGCAGAC |
| | Anti-sense | GGCACAGGTAGACTTCAATGG |
| E2f4 | Sense | TTCTACACCTCCACCTCTG |
| | Anti-sense | CACCACTGTCCTTGTTCTC |
| E2f7 | Sense | CCTCAAGTTCCACAGACC |
| | Anti-sense | CACTCAGAAGCAACATAGC |
| E2f8 | Sense | GAGTCGGAAGGAGAAGAGC |
| | Anti-sense | CCAGGCAGATGTCGTTATTC |
| Tcfe2a | Sense | TTCCCTCCCTGACCTCTC |
| | Anti-sense | GGCTACTGATGCGATTTCC |
| GAPDH | Sense | TGTGTCCGTCGTGGATCTGA |
| | Anti-sense | CCTGCTTCACCACCTTCTTGAT |