

# ***Sequences within and upstream of the mouse *Ets1* gene drive high level expression in B cells, but are not sufficient for consistent expression in T cells***

**Alyssa Kearly<sup>1</sup>, Prontip Saelee<sup>1</sup>, Jonathan Bard<sup>1</sup>, Satrajit Sinha<sup>1</sup>, Anne Satterthwaite<sup>2</sup> and Lee Ann Garrett-Sinha<sup>1\*</sup>.**

<sup>1</sup>Department of Biochemistry, Jacobs School of Medicine and Biomedical Sciences, State University of New York at Buffalo, Buffalo, NY 14203; <sup>2</sup>Department of Internal Medicine, University of Texas Southwestern Medical Center, Dallas, TX 75390

## **\* Corresponding author**

Lee Ann Garrett-Sinha  
Department of Biochemistry  
State University of New York  
Jacobs School of Medicine and Biomedical Sciences  
955 Main St  
Buffalo, NY 14203  
Emails: [leesinha@buffalo.edu](mailto:leesinha@buffalo.edu)  
Telephone: 716-881-7995  
FAX: 716-829-2725

**Running Title:** Regulatory elements for the *Ets1* gene

**Funding Footnote:** This work was supported by grants from the Lupus Research Alliance and the National Institute of Allergy and Infectious Disease (NIAID R01 AI122720) as well as a National Cancer Institute Core Center grant to Roswell Park Cancer Institute (NCI P30CA016056).

# **Abstract:**

The levels of transcription factor Ets1 are high in resting B and T cells, but are downregulated by signaling through antigen receptors and Toll-like receptors (TLRs). Loss of Ets1 in mice leads to excessive immune cell activation and development of an autoimmune syndrome and reduced Ets1 expression has been observed in human PBMCs in the context of autoimmune diseases. In B cells, Ets1 serves to prevent premature activation and differentiation to antibody-secreting cells. Given these important roles for Ets1 in the immune response, stringent control of *Ets1* gene expression levels is required for homeostasis. However, the genetic regulatory elements that control expression of the *Ets1* gene remain relatively unknown. Here we identify a topologically-associating domain (TAD) in the chromatin of B cells that includes the mouse *Ets1* gene locus and describe an interaction hub that extends over 100 kb upstream and into the gene body. Additionally, we compile epigenetic datasets to find several putative regulatory elements within the interaction hub by identifying regions of high DNA accessibility and enrichment of active enhancer histone marks. Using reporter constructs, we determine that DNA sequences within this interaction hub are sufficient to direct reporter gene expression in lymphoid tissues of transgenic mice. Further analysis indicates that the reporter construct drives faithful expression of the reporter gene in mouse B cells, but variegated expression in T cells, suggesting the existence of T cell regulatory elements outside this region. To investigate how the downregulation of Ets1 transcription is associated with alterations in the epigenetic landscape of stimulated B cells, we performed ATAC-seq in resting and BCR-stimulated primary B cells and identified four regions within and upstream of the *Ets1* locus that undergo changes in chromatin accessibility that correlate to *Ets1* gene expression. Interestingly, functional analysis of several putative Ets1 regulatory elements using luciferase constructs suggested a high level of functional redundancy. Taken together our studies reveal a complex network of regulatory elements and transcription factors that coordinate the B cell-specific expression of *Ets1*.

# Introduction:

Autoimmune diseases are caused by the immune system attacking self-tissues. Genetic factors influence the development of autoimmune diseases and genome-wide association studies (GWAS) have been used to identify disease-associated single-nucleotide polymorphisms (SNPs) for many autoimmune diseases. GWAS has established that the human *ETS1* gene is a susceptibility locus for multiple autoimmune diseases, such as systemic lupus erythematosus (SLE), rheumatoid arthritis (RA), psoriasis and others (1). *Ets1* mRNA levels are reduced in PBMCs isolated from patients with a variety of different autoimmune diseases (2-7). Supporting a role for *Ets1* in autoimmunity, mice lacking *Ets1* develop a lupus-like autoimmune disease, with excessive B and T cell activation and secretion of IgM and IgG autoantibodies against self-antigens (8-11). Cell-type-specific deletion of *Ets1* has demonstrated cell-intrinsic roles for *Ets1* in blocking the aberrant activation of both B and T cells (10, 11).

*Ets1* expression is high in resting lymphocytes (12-14), but is downregulated by stimulation through antigen receptors or Toll-like receptors (15-19). Loss of *Ets1* leads to increased B cell differentiation into antibody-secreting plasma cells (20) and forced expression of *Ets1* blocks development of plasma cells (7, 15, 16, 21). Since the level of *Ets1* in B cells is crucial to regulating the differentiation program, expression of the *Ets1* gene must be tightly controlled. Maintenance of *Ets1* expression in resting B cells relies heavily on inhibitory signaling pathways that limit the transduction of activating signals. B cells that lack expression of inhibitory signaling components such as inhibitory receptors CD22 or SiglecG, proximal kinase Lyn, or downstream phosphatases SHP-1 or SHIP-1 have increased BCR signaling and express lower levels of *Ets1* mRNA and protein (16). Our studies have shown that the low level of *Ets1* in Lyn<sup>-/-</sup> B cells is not due to decreased stability of the *Ets1* protein or mRNA (16). Instead, BCR and TLR signaling result in a decrease in transcription of the *Ets1* gene and this does not require new protein synthesis, implicating pre-formed transcription factors (17).

Transcriptional regulation of gene expression is controlled by three-dimensional chromatin architecture in the nucleus as well as the binding of cell-type specific transcription factors. As part of this process, the genome is partitioned into an array of chromatin loops generated by the combined actions of

cohesin and CTCF proteins that define topologically-associated domains (TADs) (22-24). TADs counteract the formation of transcriptionally-silent heterochromatin domains (25, 26) and also function to insulate promoters of genes within them from the influence of response elements outside the TAD (27-30). TADs are further sub-divided into sub-TAD regions that contain one or a few genes along with regulatory elements such as gene enhancers or silencers that bind specific transcription factors (26-28). Recently, the organization TAD structures in the *Ets1* locus in mouse and human T cells have been published (31). This analysis also identified a super-enhancer region located approximately 250 kb downstream of the *Ets1* gene, which is required for CD4<sup>+</sup> T helper 1 (Th1) differentiation (31). In contrast to the situation in T cells, the genomic organization of the *Ets1* locus and the regulatory elements and transcription factors that bind to them in B cells remain unknown.

To identify B cell TADs and sub-TADs around *Ets1* and also regulatory elements that coordinate expression of *Ets1*, we examined chromatin interactions, epigenetic profiles and DNA accessibility in mouse B cells. To further validate the roles of the regulatory elements, we generated a bacterial artificial chromosome (BAC) reporter construct in which eGFP expression was inserted under the direction of sequences from -113 kb to +105 kb of the mouse *Ets1* gene and observed reporter expression that recapitulated endogenous *Ets1* in B cells, but did not fully recapitulate *Ets1* expression in T cells. Four putative regulatory elements experience changes in accessibility upon BCR stimulation and motif analysis identified transcription factors that may bind to these regulatory elements. Transient transfection assays with numerous putative regulatory element sequences failed to detect sequences that individually activate or repress reporter gene transcription at an appreciable level. These results suggest the possibility that *Ets1* gene transcription in B cells may be regulated by a network of multiple regulatory elements that coordinately stimulate gene transcription, rather than being reliant on one or a few dominant regulatory elements.

## **Materials and Methods:**

### **Analysis of publicly-available Hi-C, ChIP-seq and ATAC-seq datasets:**

Hi-C data (GSE82144) (32) from resting mouse splenic B cells was visualized using the Juicebox desktop application (33). ChIP-seq datasets analyzed include: CTCF ChIP-seq of mouse spleen cells (GSE29184) (34); H3K4me1, H3K4me3 and H3K27ac ChIP-seq of FACS-sorted (CD3<sup>-</sup>, B220<sup>+</sup>, CD19<sup>+</sup>) mouse splenic B cells (GSE60103) (35); p300 ChIP-seq of resting mouse splenic B cells (GSE82144) (32); and c-Jun and JunD ChIP-seq of mouse the CH12 B lymphoma cell line (GSE36030) (36). ATAC-seq and RNA-seq datasets were retrieved from the ImmGen data series (GSE100738, ATAC-Seq and GSE109125, RNA-Seq) with the sorted immune cell populations (37). The ImmGen cell populations analyzed were: bone marrow hematopoietic stem cells (HSC), bone marrow multi-potent progenitors (MPP), bone marrow common lymphoid progenitors (CLP), bone marrow pro-B cells (pro-B), bone marrow immature B cells (Imm B), total splenic B cells, splenic follicular B cells, splenic marginal zone B cells, splenic germinal center B cells, splenic plasma cells, naïve splenic CD4<sup>+</sup> T cells, naïve splenic CD8<sup>+</sup> T cells, splenic NK cells, splenic macrophages, splenic neutrophils and splenic dendritic cells (DC). Sorting parameters for these populations are listed in Yoshida et al (37). In addition, ATAC-seq data from FACS-sorted (CD3<sup>-</sup>, B220<sup>+</sup>, CD19<sup>+</sup>) mouse splenic B cells was also used (GSE60103) (35). The JASPAR 2020 database of transcription factor binding sites was used to identify transcription factor motifs (38).

For human cells, the following datasets were analyzed: CD20<sup>+</sup> B cell CTCF ChIP-seq (GSE29611) (39); CD19<sup>+</sup> B cell H3K27ac ChIP-seq (GSE18927) (40); GM12878 B lymphoma H3K27ac ChIP-seq (GSE51334) and naïve B cell ATAC-seq (GSE118253) (41). All of the datasets were visualized using the UCSC Genome Browser (42).

### **BAC recombineering and transgenic mouse generation:**

The RP23-350A20 BAC (BACPAC Resources) was recombineered according to the protocol previous described (43) using a targeting construct that contained eGFP, an SV40 polyA sequence and a

neomycin/kanamycin resistance cassette between homology arms to replace the first exon of the *Ets1* gene. Clones were selected and PCR-tested for eGFP insertion. FLP-FRT was induced to remove the neomycin/kanamycin resistance cassette. The final recombined BAC contains the eGFP cDNA and SV40 polyA sequence inserted in place of the first exon of the major isoform of Ets1 (mouse Ets1 isoform 1) and under the control of Ets1 gene regulatory elements.

The recombined BAC was microinjected into fertilized C57BL/6 mouse eggs by the Roswell Park Comprehensive Cancer Center (RPCCC) Gene Targeting and Transgenic Facility. BAC Transgenic (BACtg) mice were identified by PCR genotyping. All mice were housed at RPCCC Lab Animal Shared Resource and experiments were performed in accordance with Institutional Animal Care and Use Committee protocol number UB1104M. Mice were euthanized by CO<sub>2</sub> inhalation followed by cervical dislocation.

### **BACtg flow cytometry:**

Single-cell suspensions were made from the spleens of wild type and BACtg animals and stained with Ghost Dye Violet 510 (Tonbo Biosciences) and fluorescent dye-conjugated antibodies against B lineage surface markers, including B220, CD19, CD21, CD23, CD80, PDL2, Fas, PNA, CD138, and CD98, or T lineage surface markers, including CD3, CD4, and CD8. Cells were fixed and permeabilized for intracellular staining and stained with unconjugated rabbit monoclonal anti-Ets1 (D8O8A, Cell Signaling Technologies) or rabbit monoclonal IgG isotype control, followed by PE-conjugated anti-rabbit IgG. Data was collected using a LSR II flow cytometer and analyzed using FlowJo software. Samples were gated on singlets and live cells, and specific populations were identified based on marker expression.

### **BACtg B cell isolation, stimulation, and qPCR:**

Splenic B cells from BACtg and wild type littermate control mice were isolated using the EasySep Mouse B Cell Isolation Kit (StemCell Technologies) and rested for 30 minutes and then either left unstimulated or stimulated with the addition of 10 µg/ml goat anti-IgM F(ab')<sub>2</sub> crosslinking antibody (Jackson

ImmunoResearch) for two hours before RNA isolation. Primers for qPCR include Ets1 mRNA (F-AGTCTTGTCAGTCCTTTATCAGC, R-TTTCCTCTTTCCCCATCTCC); Ets1-pre-mRNA Ets1 pre-mRNA set A (F-TCGATCTCAAGCCGACTCTC, R-GTCTTGGGCCACCAACAGTC); eGFP (F-CCATCTTCTTCAAGGACGAC, R-GCCATGATATAGACGTTGTGG); and beta-actin (F-GCAGCTCCTTCGTTGCCGGTC, R-TTGCACATGCCGGAGCCGTTG).

### **Mouse B cell ATAC-seq and identification of DARs:**

B cells were isolated from wild type mouse spleens and stimulated via the BCR for two hours. 100,000 cells from each sample were pelleted and stored frozen prior to ATAC-seq (Active Motif). Illumina NextSeq500 sequencing was used to generate paired-end 42 bp sequencing reads that were mapped to the mm10 genome build using the BWA algorithm with default settings (44). Duplicate reads were removed and analysis was performed on reads that passed Illumina's purity filter, aligned with a maximum of 2 mismatches, and mapped uniquely. Data was normalized across samples by reducing read count to that of the smallest sample. Peaks were called using the MACS2 algorithm (45). DARs were identified by comparing the number of sequencing tags in each region and identifying regions with greater than 2-fold differences in tag count. Gene ontology analysis of genes within 10 kb of DARs was performed using DAVID (46), limited to Gene Ontology (GO)'s biological process annotation set. *De novo* motif analysis was performed using HOMER's findMotifsGenome.pl program (47). The FIMO program within the MEME Suite of tools was used to scan the four DARs for transcription factor motifs (48). ATAC-seq data are available in the GEO database under the accession number GSE273513.

### **Luciferase construct design and luciferase assays:**

A region of the mouse Ets1 proximal promoter of 469 bp (-451 to +17 bp) was cloned into the pGL3-Basic firefly luciferase vector. Sequences covering regions of Sites 1-6 (defined in the Results section below) were cloned upstream of this minimal promoter fragment. In addition, a shorter fragment of the minimal promoter of 193 bp (-114 to +59 bp) was cloned into pGL3-Basic along with shorter segments of putative

regulatory regions. A20 B lymphoma cells were co-transfected with 2 µg of the firefly luciferase plasmids and 0.25 µg of pRL-Renilla luciferase plasmid using Nucleofection (Lonza) and analyzed at 24 hours after transfection. Firefly luciferase activity was normalized to Renilla luciferase activity.



# Results:

## The 3-D chromatin organization of the mouse *Ets1* gene in B cells

To define topologically-associated domains (TADs) around the mouse *Ets1* gene, we examined Hi-C data from a previously-published study of resting mouse B cells (49). This analysis showed that the mouse *Ets1* gene is contained within a CTCF-flanked TAD of ~900 kb that stretches from -445 kb to +452 kb from the major transcriptional start site (TSS) of *Ets1* (boundaries of the TAD are shaded in aqua in Figure 1A) (34). In addition to the *Ets1* gene, this TAD contains fellow ETS family member *Fli1*, as well as two potassium channel-encoding genes, *Kcnj1* and *Kcnj5*. A closer look at the contact map reveals a CTCF-flanked sub-TAD stretching from -107 kb to +244 kb (indicated by the large aqua triangle in Figure 1B). This region contains only the *Ets1* gene itself, with no other known genes. Within the *Ets1* sub-TAD, there is a local hub of more intense interactions that stretches from -107 kb to the end of the *Ets1* gene (approximately +62 kb from the TSS) (denoted by the smaller aqua triangle in Figure 1B). It is likely that regulatory elements crucial for B cell-specific expression of mouse *Ets1* are present within the strong interaction hub. We suspect that the CTCF binding site located at -107 kb likely functions to insulate the *Ets1* promoter from regulatory elements of the *Fli1* gene, which is oriented in a head-to-head manner with *Ets1* and separated by 154 kb (Figure 1B). Indeed, Hi-C detects little interaction between the *Fli1* promoter and sequences in the *Ets1* sub-TAD.

## A region extending from -113 kb to +105 kb from the mouse *Ets1* TSS is directs expression in B cells

Prior transgenic mouse studies have shown that sequences located from -5.3 kb to +9 kb from the TSS of *Ets1* were not sufficient to drive expression of a transgene in B cells or other lymphoid lineages (50). To identify sequences that drive B cell-specific expression of *Ets1*, we generated BAC transgenic mice harboring a 219 kb piece of mouse DNA including sequences from -113 kb upstream of the *Ets1* gene (stretching just beyond the CTCF-bound boundary) to +105 kb downstream of the transcriptional start site. This BAC construct contains the entire local hub of robust interactions around the *Ets1* gene (Figure 2A).

The first exon of the *Ets1* gene was replaced with an eGFP cDNA, followed by a stop codon and an SV40 polyA signal, allowing eGFP expression from the *Ets1* promoter under the direction of putative regulatory elements located within the BAC (Figure 2A). Analysis of both mRNA and protein showed high eGFP expression in lymphoid tissues, including the thymus, spleen, and lymph nodes, where B and T lymphocytes are abundant and where Ets1 is normally expressed at high levels (Figure 2B-C). Non-lymphoid tissues such as the brain and kidneys showed much lower expression as expected.

To compare the potential overlap in the Ets1 and eGFP expression in these BACtg mice at a cellular level, we performed intracellular flow cytometry using an antibody specific for Ets1. B220<sup>+</sup>CD19<sup>+</sup> B cells from both wild type mice and BACtg mice showed high levels of intracellular Ets1 expression compared to isotype control staining (Figure 3A, bottom). Mature B cell subsets within this B220<sup>+</sup>CD19<sup>+</sup> population, including CD23<sup>hi</sup>CD21<sup>lo</sup> follicular B cells, CD23<sup>lo</sup>CD21<sup>hi</sup> marginal zone B cells, and Fas<sup>+</sup>PNA<sup>+</sup> germinal center B cells, all showed strong positive staining for both Ets1 and eGFP (Figure 3B). Ets1 is downregulated upon B cell differentiation to plasma cells (12, 15, 51). In keeping with this, eGFP levels were lower in B220<sup>lo</sup>CD138<sup>+</sup>CD98<sup>+</sup> plasma cells than in B220<sup>+</sup>CD19<sup>+</sup> B cells of BACtg mice (Figure 3B). However, in contrast to the B cells, eGFP expression did not fully reflect Ets1 expression in the T cell populations of BACtg mice. While both CD3<sup>+</sup>CD4<sup>+</sup> and CD3<sup>+</sup>CD8<sup>+</sup> T cells had uniformly high staining for intracellular Ets1, GFP expression showed a broad range in these populations (Figure 3B), indicating variegation of BACtg expression.

BCR stimulation induces the downregulation of Ets1 in B cells, mainly through a decrease in *Ets1* gene transcription (16, 17). To determine if the response elements required for this phenomenon are within the 219 kb of the BAC, B cells from the BACtg animals were isolated and stimulated *ex vivo*, and expression of Ets1 and eGFP were assessed. As shown in Supplemental Figure S1, stimulation by the BCR resulted in a decrease in Ets1 pre-mRNA and mRNA in B cells. Because the pre-mRNA is a short-lived product in the cells, its measurement serves as a surrogate for the transcriptional level of the *Ets1* gene. In the BACtg B cells, eGFP mRNA was similarly downregulated in response to BCR stimulation, although not quite as dramatically. We surmise that the attenuated downregulation of eGFP compared to Ets1 is due to a

difference in mRNA stability of the two transcripts (16, 52). Overall, eGFP expression in the mature B cell populations of these BACtg mice recapitulates that of endogenous *Ets1* under both basal conditions and upon BCR stimulation.

### **Multiple putative regulatory elements have the potential to regulate *Ets1***

To identify candidate regulatory elements that might control *Ets1* expression in mouse B cells, we examined epigenetic characteristics of regions included in the BAC transgene in more detail. We identified regions that had open chromatin, as defined by peaks of ATAC-seq accessibility in mouse B cells (highlighted in aqua in Figure 4A). H3K4me3 histone methylation, a mark of active promoters, was enriched at the TSS. H3K4me1 histone methylation and H3K27ac histone acetylation, both marks of active regulatory regions (53), were enriched at numerous DNA segments upstream and within the *Ets1* gene. On the other hand, regions downstream of the *Ets1* gene did not harbor any regions that showed enrichment for H3K27Ac or H3K4me1 in B cells. These observations, combined with the Hi-C data shown in Figure 1 and the BAC transgenic data shown in Figures 2-3, suggest that in mouse B cells *Ets1* is likely regulated by DNA sequences upstream and within the gene itself, rather than sequences located downstream of the gene.

The BAC transgenic mice have variegated expression of eGFP in T cell populations, suggesting that the BAC lacks all the elements required for reproducible expression of *Ets1* in T cells. In contrast to B cells which lack putative regulatory elements downstream of the *Ets1* gene, ATAC-seq data from mouse T cell populations shows that T cells have accessible sites downstream of the *Ets1* gene (highlighted in aqua in Supplemental Figure S2). These sites are located beyond the downstream regions included in the BAC (highlighted in yellow in Supplemental Figure S2) and could represent T cell-specific enhancers required for optimal and sustained expression. Since some of the T cells in BACtg mice express high levels of GFP, the 219 kb region of the BAC must encompass regulatory elements that can induce expression in T cells. However, these elements appear to be insufficient to prevent spreading of heterochromatin into the *Ets1*

locus in T cells thus resulting in variegated expression of the transgene. Recently a region located ~250 kb downstream of *Ets1* was shown to function as a multi-enhancer for mouse Th1 cells (31).

While peaks of H3K27Ac and H3K4me1 histone marks were discrete in regions 5' to the *Ets1* gene, there was broad enrichment of these marks across the body of the *Ets1* gene (Figure 4A). To better define potentially important regulatory regions within the *Ets1* gene itself, we examined histone marks in a human B cell line GM12878 that expresses Ets1. As shown in Figure 4B, GM12878 cells have a more discrete pattern of enrichment of H3K27Ac at seven specific regions within and around the human *ETS1* gene. These regions of H3K27Ac enrichment are characterized by the presence of DNase I hypersensitivity and were labeled Site 1-7 (Figure 4B). Site 4 corresponds to the *Ets1* proximal promoter and beginning of intron 1, sequences previously tested in transgenic mice and shown to not mediate lymphoid-specific expression (50). Thus, this region was not further analyzed. Site 1 encompasses the CTCF binding peaks located at -104 kb upstream of the mouse *Ets1* promoter. Sites 1-7 contained regions of DNA sequence conservation between mouse and human, supporting the idea that they may be functionally important. In summary, there are multiple potential regulatory elements found within and upstream of the mouse *Ets1* gene that may be crucial for B cell-specific expression.

### **Chromatin patterns in resting B cells versus plasma cells**

Analysis of previously-published RNA-seq data from sorted purified mouse bone marrow progenitors (37) shows that Ets1 is expressed at low levels in hematopoietic stem cells (HSC), multi-potent progenitors (MPP) and common lymphoid progenitors (CLP), but is elevated once cells commit to the B cell lineage and again reduced in fully-differentiated plasma cells (PC) (Figure 5A). Open chromatin patterns assessed by ATAC-seq in the same sorted populations showed that there were differences in the regions of open chromatin between resting follicular (FO) B cells, where Ets1 levels are high, and PC, where Ets1 levels are low (Figure 5B). The ATAC-seq peaks located near the sub-TAD boundary where CTCF binds (-100 to -106 kb region) and the ATAC-seq peaks located in the proximal promoter of *Ets1* were equally present in both FO B cells and PC. When examining the other ATAC-seq peaks, we found

that peaks located at approximately -87 kb, -44 kb, -16 kb, -10 kb, +39 kb and +44 kb were strong in B cells, but reduced in PC (Figure 5B, blue highlighting). The differences in chromatin accessibility at these regions may indicate the presence of regulatory elements bound by activating transcription factors promoting *Ets1* expression in B cells.

### **Putative regulatory elements within and upstream of the *Ets1* gene change in accessibility upon BCR stimulation**

As mouse *Ets1* gene transcription is decreased in response to BCR stimulation and this effect is mediated by response elements within the BAC, we explored changes in chromatin accessibility in B cells in response to BCR stimulation. The downregulation of *Ets1* gene transcription occurs early after activation, within 2 hours of BCR engagement (16). We therefore generated ATAC-seq datasets for resting mouse B cells and B cells stimulated for 2 hours with a BCR crosslinking antibody. Regions that experienced a two-fold or greater change in accessibility between resting and stimulated B cells were deemed as differentially accessible regions (DARs) (Figure 6A). Upon stimulation, changes in chromatin accessibility were induced genome-wide, with 2259 DARs increasing and 598 DARs decreasing in accessibility. The disproportionate amount of DARs increasing in accessibility versus those decreasing is in line with previous studies of immune cell *ex vivo* stimulation (54, 55). Genes within 10 kb of DARs that gained or lost accessibility were found to be functionally associated with immune responses and signal transduction (Figure 6B). Motif enrichment analysis revealed that DARs that gained accessibility frequently contained binding sites for AP-1 and NF- $\kappa$ B proteins, which are downstream effectors of pathways activated by BCR signaling, while those that lost accessibility were most highly enriched for the *Ets1* motif (Figure 6C). The loss of *Ets1* binding in BCR-stimulated B cells is consistent with the decreased transcription of the *Ets1* gene upon BCR stimulation. However, it is also likely that BCR-induced phosphorylation of the *Ets1* protein, which has been shown to decrease its DNA binding activity and in T cells leads to its nuclear export (56, 57), contributes to loss of *Ets1* binding in stimulated B cells.

Four ATAC-Seq peaks in the region contained within the BACtg were identified as DARs, with 2 gaining and 2 losing accessibility after 2 hours of BCR stimulation (Figure 7A, DARs highlighted in blue and red, respectively). These sites were located at -85 kb (included within Site 2 defined in Figure 4B), at +36 and +39 kb (included within Site 6) and at +49 kb (included within Site 7) from the TSS. The changes in DNA accessibility could reflect stimulation-induced alterations in transcription factor binding. A decrease in accessibility upon stimulation could reflect the loss of binding of a transcriptional activator, while an increase in accessibility may be caused by increased binding of a transcriptional repressor. Both scenarios could lead to downregulation of *Ets1* gene transcription.

DNA sequences in the +39 kb and +49 kb DARs showed conservation with corresponding sequences in human *ETS1* gene (Supplemental Figure S3). On the other hand, the other two mouse DAR sequences at -85 kb and +36 kb didn't show strong DNA sequence conservation, although these DARs were located within Sites 2 and 6 that did have regions of conserved sequence. To identify transcription factors that may bind to the four DAR sites, we used the FIMO program to scan the sequences for DNA-binding motifs. The five strongest motif matches for each site are mapped in Figure 7B. Transcription factors sites found within the DARs that lose accessibility after BCR stimulation include those for NF- $\kappa$ B proteins, SpiC, Smad proteins and Irf5, among others. Both sites that gained accessibility contained motifs for AP-1 family members and for Srebf2, a sterol RE-binding protein (Figure 7B). ChIP-seq data from the CH12 mouse B cell lymphoma line indicates that both c-Jun and JunD can bind at the +49 kb DAR site that gains accessibility (Figure 7C) (36). With the exception of Hoxd3 and Hnf4g, all of the transcription factors shown in Figure 7B are expressed in B cells (37) and could potentially be involved in regulating *Ets1* transcription. These changes in chromatin accessibility that occur in response to BCR stimulation could highlight the response elements important for regulation of *Ets1* expression and implicate AP-1, in particular, in its downregulation.

## **A selection of regulatory elements tested in transient transfections are insufficient to activate the *Ets1* promoter**

To determine if the potential regulatory elements identified above govern *Ets1* expression, we next generated a series of luciferase reporter constructs with putative regulatory elements cloned upstream of the *Ets1* minimal promoter (Figure 8). We generated two sets of constructs. In the first set, a 469 bp fragment of the *Ets1* minimal promoter (451 bp to +18 bp from the TSS) was cloned along with putative regulatory element fragments containing regions of Sites 1-3 and 5-6 (shown as blue boxes in Figure 8). The second set of constructs was more focused and included a smaller fragment of the *Ets1* promoter (-114 bp to +60 bp from the TSS) along with shorter segments of the putative regulatory elements (shown in green boxes in Figure 8). For the second set of constructs, we chose elements that were located in regions of H3K27ac enrichment and that were more accessible in either resting B cells or plasma cells (as highlighted in Figure 5B), assuming that differences in chromatin accessibility would reflect the high and low expression of *Ets1* in these cell types.

The reporter constructs were transfected into the A20 mouse B lymphoma cell line, which expresses *Ets1* and downregulates it transcriptionally in response to BCR stimulation similar to primary B cells (16, 17). Out of the putative regulatory elements tested, ranging from -110 kb upstream to +49 kb downstream of the *Ets1* TSS, none showed a strong ability to activate or repress expression (Figure 8C-D). The strongest activity was found with sequences in Site 2 that were able to stimulate transcription modestly at approximately 1.5-fold over the promoter alone (Figure 8C). Given that none of these sequences alone was able to direct high level expression of a reporter gene in B cells, these data suggest that *Ets1* may be regulated by a combinatorial network of response elements that cooperate to stimulate gene expression in B cells.

# Discussion:

In mice, loss of *Ets1* leads to autoimmune disease with inappropriate activation of B cells, including autoreactive B cells, and their differentiation into plasma cells that secrete antibodies (8, 15, 16). Numerous autoimmune disease-associated SNPs map near the human *ETSI* gene (1) and decreased *Ets1* expression has been found in immune cells from autoimmune disease patients (2-7, 58, 59). Given that aberrant expression of the *Ets1* gene is found in autoimmune diseases, it is important to understand the regulatory mechanisms controlling its expression pattern.

Previous studies to examine cell-type specific regulation of the *Ets1* gene focused on sequences ranging from -5.3 kb to +9 kb with respect to the major transcriptional start site of the gene. However, these sequences fail to drive lymphocyte expression of a reporter gene in transgenic mice (50). To determine regulatory elements for *Ets1* in B cells, we first identified an 835 kb topologically-associated domain (TAD) flanked by CTCF binding sites that contains the mouse *Ets1* and other genes. This large TAD region includes a smaller 289 kb CTCF-flanked sub-TAD containing only the *Ets1* gene. Within the sub-TAD, Hi-C analysis shows that B cells have an interaction hotspot extending from approximately -110 kb to the end of the final exon of *Ets1*. The TAD structure of the mouse *Ets1* gene is conserved in CD4<sup>+</sup> T cells, as shown by Chandra et al (31).

We generated a BAC-transgenic reporter system carrying sequences from -113 kb to +105 kb from the *Ets1* TSS, which contains the Hi-C interaction hotspot of *Ets1*. Sequences in the BAC transgene recapitulated well the expression of *Ets1* in B cells, indicating that all regulatory elements needed for B cell expression are present within the interaction hotspot. B cell activation by BCR crosslinking or incubation with TLR ligands leads to downregulation of *Ets1* and this is mediated mainly at the level of gene transcription and does not require new protein synthesis (16, 17). The sequences contained in the BACtg are sufficient to mediate downregulation of the GFP reporter in response to BCR stimulation. In contrast to the pattern in B cells, the expression of eGFP in T cells from BACtg mice was variegated, suggesting the existence of further T cell-specific enhancers outside the BAC segment. Variegation of gene expression is typically due to stochastic spread of heterochromatin in a subset of cells. Thus, there may be sequences



located downstream of the 3' end of the BAC that are required in T cells, but not B cells, to prevent heterochromatin spread. A previously-described super-enhancer segment located ~250 kb downstream of *Ets1* was shown to be essential for maximal *Ets1* expression in double positive thymocytes and in Th1 cells and for proper Th1 cell differentiation (31). This super-enhancer region is not present in the BACtg and its absence may lead to variegated expression of the GFP marker in T cells. However, it is important to note that there is also a region of accessible chromatin in T cells located at approximately +120 kb from the *Ets1* TSS that is missing from the BACtg (see Figure S2). This sequence could also play a role in regulation of *Ets1* expression in T cells.

To identify potential regulatory elements of *Ets1*, we examined epigenetic datasets from B cells and defined seven broad regions that have properties of regulatory sequences (Sites 1-7). All seven of these sites have regions with high DNA sequence homology between the mouse and human genomes. However, these large regions of sequence showed only weak activity in a luciferase reporter assay in A20 B lymphoma cells. To refine our analysis, we compared open chromatin regions in follicular B cells, which express high levels of *Ets1*, to open chromatin regions in plasma cells, which have very low levels of *Ets1*. This analysis identified 6 regions with differential accessibility, two of which overlapped with sequences in Sites 1-7. We tested these six shorter segments in luciferase assays, but they also failed to mediate strong activation of the promoter. Therefore, we did not identify any regions that were able to strongly activate or repress expression of a reporter gene in transient transfection assays. This result suggests the possibility that transcription of the *Ets1* gene is more complicated and relies on the combinatorial action of multiple regulatory regions that must function cooperatively.

Since sequences in the BACtg were sufficient to mediate down-regulation of *Ets1* gene transcription upon BCR stimulation, we identified differentially-accessible regions (DARs) in the chromatin of stimulated and unstimulated B cells. In the region surrounding the *Ets1* locus, four DAR regions were identified and three of these overlapped with sequences in Sites 1-7. The two sites that gained accessibility after stimulation contained motifs for AP-1 family members. AP-1 motifs were also enriched globally in sites that gained accessibility with stimulation, in line with other studies examining stimulation-

responsive chromatin and gene expression that have implicated AP-1 in mediating early genome-wide changes (60, 61). AP-1 is activated downstream of MAPK pathways, particularly JNK and ERK, and we have previously demonstrated that JNK activity is required for *Ets1* downregulation (16). Binding of AP-1 subunits c-Jun and JunD was detected at the +49 DAR in CH12 B lymphoma cells. While AP-1 proteins often serve to activate gene expression, considerable evidence shows that they can also serve as repressors of transcription in appropriate contexts (62-66). Thus, recruitment of AP-1 proteins to the -85 kb and +49 kb DAR regions might serve to inhibit *Ets1* gene transcription after BCR stimulation. Further studies will be needed to fully determine the role of AP-1 family proteins in the regulation of *Ets1* expression.

The two regions that lost accessibility upon stimulation could represent regulatory elements bound by transcription factors responsible for maintaining *Ets1* expression in resting B cells. These two regions contained potential binding sites for a variety of transcription factors including Smad proteins, NFκB family members, and ETS family member SpiC. SpiC is a member of the ETS family that has been shown to be a negative regulator of B cell function (67). However, examining ChIP-seq data for SpiC in a mouse pro-B cell line (68) did not reveal any binding of SpiC to its putative site in the DAR at +39 kb (data not shown). Therefore, SpiC seems to be a less likely candidate for promoting basal expression of *Ets1* in resting B cells. Canonical NFκB proteins are sequestered in the cytoplasm in the resting state and hence also seem unlikely to mediate basal expression of *Ets1* in resting B cells. On the other hand, it is possible that non-canonical NFκB signaling, such as that triggered by the BAFF receptor or CD40 receptor, contributes to the maintaining *Ets1* expression in resting B cells. The Smad2/3/4 complex is activated downstream of TGFβ signaling, which has been shown to suppress B cell responses through upregulation of inhibitory signaling components (69), and mice lacking the receptor TGFβ RII have hyper-responsive B cells (70). Therefore, the TGFβ signaling pathway is an attractive possibility for regulating basal levels of *Ets1* in resting B cells. Future studies on potential roles for Smad proteins and non-canonical NFκB proteins will be important to determine whether they are involved in maintaining *Ets1* in resting B cells.

Overall, in this report we have analyzed the regions surrounding the mouse *Ets1* gene and determined that a BAC transgene carrying sequences from -113 kb to +105 kb from the TSS is sufficient

for high level and reproducible expression of a reporter gene in B cells. Unexpectedly, this region is not sufficient for maintaining consistent expression of the reporter gene in T cells and instead resulted in variegated expression in both CD4 and CD8 T cells. Therefore, there are differences in the regulatory sequences required for optimal expression in B and T cells. Further analysis of epigenetic marks and the comparison of open chromatin in resting B cells versus activated B cells as well as in follicular B cells versus plasma cells resulted in the identification of numerous regions with features of regulatory elements that could participate in controlling *Ets1* expression in B cells. However, none of the elements was sufficient by itself to strongly activate or repress transcription in transient transfection assays, suggesting that *Ets1* may be regulated by a network of several regulatory elements located upstream and within its introns. One limitation worth pointing out is the episomal nature and the lack of intact chromatin conformations in plasmid-based reporters that were utilized in this report. Further studies are needed to fully dissect the *Ets1* cis-regulatory network and the role it may play in B cell activation.

### **Acknowledgements:**

We thank Alex Glather, Akinsola Oyelakin and Dr. Yungki Park (all at State University of New York at Buffalo) for helpful discussions relevant to this study. We thank Kirsten Smalley for help with animal husbandry. This work was funded by NIH grant R01 AI122720.

## Citations

1. Garrett-Sinha, L. A., A. Kearly, and A. B. Satterthwaite. 2016. The Role of the Transcription Factor Ets1 in Lupus and Other Autoimmune Diseases. *Crit Rev Immunol* 36: 485-510.
2. Shan, S., J. Dang, J. Li, Z. Yang, H. Zhao, Q. Xin, X. Ma, Y. Liu, X. Bian, Y. Gong, and Q. Liu. 2014. ETS1 variants confer susceptibility to ankylosing spondylitis in Han Chinese. *Arthritis Res Ther* 16: R87.
3. Li, Y., L. D. Sun, W. S. Lu, W. L. Hu, J. P. Gao, Y. L. Cheng, Z. Y. Yu, S. Yao, C. F. He, J. L. Liu, Y. Cui, and S. Yang. 2010. Expression analysis of ETS1 gene in peripheral blood mononuclear cells with systemic lupus erythematosus by real-time reverse transcription PCR. *Chin Med J (Engl)* 123: 2287-2288.
4. Yang, W., N. Shen, D. Q. Ye, Q. Liu, Y. Zhang, X. X. Qian, N. Hirankarn, D. Ying, H. F. Pan, C. C. Mok, T. M. Chan, R. W. Wong, K. W. Lee, M. Y. Mok, S. N. Wong, A. M. Leung, X. P. Li, Y. Avihingsanon, C. M. Wong, T. L. Lee, M. H. Ho, P. P. Lee, Y. K. Chang, P. H. Li, R. J. Li, L. Zhang, W. H. Wong, I. O. Ng, C. S. Lau, P. C. Sham, Y. L. Lau, and C. Asian Lupus Genetics. 2010. Genome-wide association study in Asian populations identifies variants in ETS1 and WDFY4 associated with systemic lupus erythematosus. *PLoS Genet* 6: e1000841.
5. Wei, L., Q. Zhou, S. Hou, L. Bai, Y. Liu, J. Qi, Q. Xiang, Y. Zhou, A. Kijlstra, and P. Yang. 2014. MicroRNA-146a and Ets-1 gene polymorphisms are associated with pediatric uveitis. *PLoS One* 9: e91199.
6. Xiang, N., X. P. Li, X. M. Li, G. S. Wang, J. H. Tao, H. F. Pan, X. Fang, Q. Ma, and N. Yu. 2014. Expression of Ets-1 and FOXP3 mRNA in CD4(+)CD25(+) T regulatory cells from patients with systemic lupus erythematosus. *Clin Exp Med* 14: 375-381.
7. Zhang, R., B. Pan, Y. Li, and X. Li. 2019. SNP rs4937333 in the miRNA-5003-Binding Site of the ETS1 3'-UTR Decreases ETS1 Expression. *Front Genet* 10: 581.
8. Wang, D., S. A. John, J. L. Clements, D. H. Percy, K. P. Barton, and L. A. Garrett-Sinha. 2005. Ets-1 deficiency leads to altered B cell differentiation, hyperresponsiveness to TLR9 and autoimmune disease. *Int Immunol* 17: 1179-1191.
9. Mouly, E., K. Chemin, H. V. Nguyen, M. Chopin, L. Mesnard, M. Leite-de-Moraes, O. Burlen-defranoux, A. Bandeira, and J. C. Bories. 2010. The Ets-1 transcription factor controls the development and function of natural regulatory T cells. *J Exp Med* 207: 2113-2125.
10. Kim, C. J., C. G. Lee, J. Y. Jung, A. Ghosh, S. N. Hasan, S. M. Hwang, H. Kang, C. Lee, G. C. Kim, D. Rudra, C. H. Suh, and S. H. Im. 2018. The Transcription Factor Ets1 Suppresses T Follicular Helper Type 2 Cell Differentiation to Halt the Onset of Systemic Lupus Erythematosus. *Immunity* 49: 1034-1048 e1038.
11. Sunshine, A., D. Goich, A. Stith, K. Sortino, J. Dalton, S. Metcalfe, E. C. Svensson, and L. A. Garrett-Sinha. 2019. Ets1 Controls the Development of B Cell Autoimmune Responses in a Cell-Intrinsic Manner. *Immunohorizons* 3: 331-340.
12. Bhattacharya, D., M. T. Cheah, C. B. Franco, N. Hosen, C. L. Pin, W. C. Sha, and I. L. Weissman. 2007. Transcriptional profiling of antigen-dependent murine B cell differentiation and memory formation. *J Immunol* 179: 6808-6819.
13. Hollenhorst, P. C., D. A. Jones, and B. J. Graves. 2004. Expression profiles frame the promoter specificity dilemma of the ETS family of transcription factors. *Nucleic Acids Res* 32: 5693-5702.
14. Kassambara, A., T. Reme, M. Jourdan, T. Fest, D. Hose, K. Tarte, and B. Klein. 2015. GenomicScape: an easy-to-use web tool for gene expression data analysis. Application to investigate the molecular events in the differentiation of B cells into plasma cells. *PLoS Comput Biol* 11: e1004077.

15. John, S. A., J. L. Clements, L. M. Russell, and L. A. Garrett-Sinha. 2008. Ets-1 regulates plasma cell differentiation by interfering with the activity of the transcription factor Blimp-1. *J Biol Chem* 283: 951-962.
16. Luo, W., J. Mayeux, T. Gutierrez, L. Russell, A. Getahun, J. Muller, T. Tedder, J. Parnes, R. Rickert, L. Nitschke, J. Cambier, A. B. Satterthwaite, and L. A. Garrett-Sinha. 2014. A balance between B cell receptor and inhibitory receptor signaling controls plasma cell differentiation by maintaining optimal Ets1 levels. *J Immunol* 193: 909-920.
17. Kearly, A., K. Ottens, M. C. Battaglia, A. B. Satterthwaite, and L. A. Garrett-Sinha. 2022. B Cell Activation Results in IKK-Dependent, but Not c-Rel- or RelA-Dependent, Decreases in Transcription of the B Cell Tolerance-Inducing Gene Ets1. *Immunohorizons* 6: 779-789.
18. Pognonec, P., K. E. Boulukos, J. C. Gesquiere, D. Stehelin, and J. Ghysdael. 1988. Mitogenic stimulation of thymocytes results in the calcium-dependent phosphorylation of c-ets-1 proteins. *EMBO J* 7: 977-983.
19. Bhat, N. K., C. B. Thompson, T. Lindsten, C. H. June, S. Fujiwara, S. Koizumi, R. J. Fisher, and T. S. Papas. 1990. Reciprocal expression of human ETS1 and ETS2 genes during T-cell activation: regulatory role for the protooncogene ETS1. *Proc Natl Acad Sci U S A* 87: 3723-3727.
20. Trezise, S., I. Y. Kong, E. D. Hawkins, M. J. Herold, S. N. Willis, and S. L. Nutt. 2023. An arrayed CRISPR screen of primary B cells reveals the essential elements of the antibody secretion pathway. *Front Immunol* 14: 1089243.
21. John, S., L. Russell, S. S. Chin, W. Luo, R. Oshima, and L. A. Garrett-Sinha. 2014. Transcription factor Ets1, but not the closely related factor Ets2, inhibits antibody-secreting cell differentiation. *Mol Cell Biol* 34: 522-532.
22. Lieberman-Aiden, E., N. L. van Berkum, L. Williams, M. Imakaev, T. Ragoczy, A. Telling, I. Amit, B. R. Lajoie, P. J. Sabo, M. O. Dorschner, R. Sandstrom, B. Bernstein, M. A. Bender, M. Groudine, A. Gnirke, J. Stamatoyannopoulos, L. A. Mirny, E. S. Lander, and J. Dekker. 2009. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* 326: 289-293.
23. Rao, S. S., M. H. Huntley, N. C. Durand, E. K. Stamenova, I. D. Bochkov, J. T. Robinson, A. L. Sanborn, I. Machol, A. D. Omer, E. S. Lander, and E. L. Aiden. 2014. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* 159: 1665-1680.
24. Dixon, J. R., S. Selvaraj, F. Yue, A. Kim, Y. Li, Y. Shen, M. Hu, J. S. Liu, and B. Ren. 2012. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* 485: 376-380.
25. Haarhuis, J. H. I., R. H. van der Weide, V. A. Blomen, K. D. Flach, H. Teunissen, L. Willems, T. R. Brummelkamp, B. D. Rowland, and E. de Wit. 2022. A Mediator-cohesin axis controls heterochromatin domain formation. *Nat Commun* 13: 754.
26. Kim, Y. J., K. R. Cecchini, and T. H. Kim. 2011. Conserved, developmentally regulated mechanism couples chromosomal looping and heterochromatin barrier activity at the homeobox gene A locus. *Proc Natl Acad Sci U S A* 108: 7391-7396.
27. Downen, J. M., Z. P. Fan, D. Hnisz, G. Ren, B. J. Abraham, L. N. Zhang, A. S. Weintraub, J. Schuijers, T. I. Lee, K. Zhao, and R. A. Young. 2014. Control of cell identity genes occurs in insulated neighborhoods in mammalian chromosomes. *Cell* 159: 374-387.
28. Hanssen, L. L. P., M. T. Kassouf, A. M. Oudelaar, D. Biggs, C. Preece, D. J. Downes, M. Gosden, J. A. Sharpe, J. A. Sloane-Stanley, J. R. Hughes, B. Davies, and D. R. Higgs. 2017. Tissue-specific CTCF-cohesin-mediated chromatin architecture delimits enhancer interactions and function in vivo. *Nat Cell Biol* 19: 952-961.
29. Lupianez, D. G., K. Kraft, V. Heinrich, P. Krawitz, F. Brancati, E. Klopocki, D. Horn, H. Kayserili, J. M. Opitz, R. Laxova, F. Santos-Simarro, B. Gilbert-Dussardier, L. Wittler, M. Borschiwer, S. A. Haas, M. Osterwalder, M. Franke, B. Timmermann, J. Hecht, M. Spielmann, A. Visel, and S. Mundlos. 2015. Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell* 161: 1012-1025.

30. Nora, E. P., A. Goloborodko, A. L. Valton, J. H. Gibcus, A. Uebersohn, N. Abdennur, J. Dekker, L. A. Mirny, and B. G. Bruneau. 2017. Targeted Degradation of CTCF Decouples Local Insulation of Chromosome Domains from Genomic Compartmentalization. *Cell* 169: 930-944 e922.
31. Chandra, A., S. Yoon, M. F. Michieletto, N. Goldman, E. K. Ferrari, M. Abedi, I. Johnson, M. Fasolino, K. Pham, L. Joannas, B. L. Kee, J. Henao-Mejia, and G. Vahedi. 2023. Quantitative control of Ets1 dosage by a multi-enhancer hub promotes Th1 cell differentiation and protects from allergic inflammation. *Immunity*.
32. Kieffer-Kwon, K. R., Z. Tang, E. Mathe, J. Qian, M. H. Sung, G. Li, W. Resch, S. Baek, N. Pruett, L. Grontved, L. Vian, S. Nelson, H. Zare, O. Hakim, D. Reyon, A. Yamane, H. Nakahashi, A. L. Kovalchuk, J. Zou, J. K. Joung, V. Sartorelli, C. L. Wei, X. Ruan, G. L. Hager, Y. Ruan, and R. Casellas. 2013. Interactome maps of mouse gene regulatory domains reveal basic principles of transcriptional regulation. *Cell* 155: 1507-1520.
33. Robinson, J. T., D. Turner, N. C. Durand, H. Thorvaldsdottir, J. P. Mesirov, and E. L. Aiden. 2018. Juicebox.js Provides a Cloud-Based Visualization System for Hi-C Data. *Cell Syst* 6: 256-258 e251.
34. Shen, Y., F. Yue, D. F. McCleary, Z. Ye, L. Edsall, S. Kuan, U. Wagner, J. Dixon, L. Lee, V. V. Lobanenkov, and B. Ren. 2012. A map of the cis-regulatory sequences in the mouse genome. *Nature* 488: 116-120.
35. Lara-Astiaso, D., A. Weiner, E. Lorenzo-Vivas, I. Zaretsky, D. A. Jaitin, E. David, H. Keren-Shaul, A. Mildner, D. Winter, S. Jung, N. Friedman, and I. Amit. 2014. Immunogenetics. Chromatin state dynamics during blood formation. *Science* 345: 943-949.
36. Mouse, E. C., J. A. Stamatoyannopoulos, M. Snyder, R. Hardison, B. Ren, T. Gingeras, D. M. Gilbert, M. Groudine, M. Bender, R. Kaul, T. Canfield, E. Giste, A. Johnson, M. Zhang, G. Balasundaram, R. Byron, V. Roach, P. J. Sabo, R. Sandstrom, A. S. Stehling, R. E. Thurman, S. M. Weissman, P. Cayting, M. Hariharan, J. Lian, Y. Cheng, S. G. Landt, Z. Ma, B. J. Wold, J. Dekker, G. E. Crawford, C. A. Keller, W. Wu, C. Morrissey, S. A. Kumar, T. Mishra, D. Jain, M. Byrska-Bishop, D. Blankenberg, B. R. Lajoie, G. Jain, A. Sanyal, K. B. Chen, O. Denas, J. Taylor, G. A. Blobel, M. J. Weiss, M. Pimkin, W. Deng, G. K. Marinov, B. A. Williams, K. I. Fisher-Aylor, G. Desalvo, A. Kiralusha, D. Trout, H. Amrhein, A. Mortazavi, L. Edsall, D. McCleary, S. Kuan, Y. Shen, F. Yue, Z. Ye, C. A. Davis, C. Zaleski, S. Jha, C. Xue, A. Dobin, W. Lin, M. Fastuca, H. Wang, R. Guigo, S. Djebali, J. Lagarde, T. Ryba, T. Sasaki, V. S. Malladi, M. S. Cline, V. M. Kirkup, K. Learned, K. R. Rosenbloom, W. J. Kent, E. A. Feingold, P. J. Good, M. Pazin, R. F. Lowdon, and L. B. Adams. 2012. An encyclopedia of mouse DNA elements (Mouse ENCODE). *Genome Biol* 13: 418.
37. Yoshida, H., C. A. Lareau, R. N. Ramirez, S. A. Rose, B. Maier, A. Wroblewska, F. Desland, A. Chudnovskiy, A. Mortha, C. Dominguez, J. Tellier, E. Kim, D. Dwyer, S. Shinton, T. Nabekura, Y. Qi, B. Yu, M. Robinette, K. W. Kim, A. Wagers, A. Rhoads, S. L. Nutt, B. D. Brown, S. Mostafavi, J. D. Buenrostro, C. Benoist, and P. Immunological Genome. 2019. The cis-Regulatory Atlas of the Mouse Immune System. *Cell* 176: 897-912 e820.
38. Fornes, O., J. A. Castro-Mondragon, A. Khan, R. van der Lee, X. Zhang, P. A. Richmond, B. P. Modi, S. Correard, M. Gheorghe, D. Baranasic, W. Santana-Garcia, G. Tan, J. Cheneby, B. Ballester, F. Parcy, A. Sandelin, B. Lenhard, W. W. Wasserman, and A. Mathelier. 2020. JASPAR 2020: update of the open-access database of transcription factor binding profiles. *Nucleic Acids Res* 48: D87-D92.
39. Consortium, E. P. 2012. An integrated encyclopedia of DNA elements in the human genome. *Nature* 489: 57-74.
40. Roadmap Epigenomics, C., A. Kundaje, W. Meuleman, J. Ernst, M. Bilenky, A. Yen, A. Heravi-Moussavi, P. Kheradpour, Z. Zhang, J. Wang, M. J. Ziller, V. Amin, J. W. Whitaker, M. D. Schultz, L. D. Ward, A. Sarkar, G. Quon, R. S. Sandstrom, M. L. Eaton, Y. C. Wu, A. R. Pfennig, X. Wang, M. Claussnitzer, Y. Liu, C. Coarfa, R. A. Harris, N. Shores, C. B. Epstein,



- E. Gjoneska, D. Leung, W. Xie, R. D. Hawkins, R. Lister, C. Hong, P. Gascard, A. J. Mungall, R. Moore, E. Chuah, A. Tam, T. K. Canfield, R. S. Hansen, R. Kaul, P. J. Sabo, M. S. Bansal, A. Carles, J. R. Dixon, K. H. Farh, S. Feizi, R. Karlic, A. R. Kim, A. Kulkarni, D. Li, R. Lowdon, G. Elliott, T. R. Mercer, S. J. Neph, V. Onuchic, P. Polak, N. Rajagopal, P. Ray, R. C. Sallari, K. T. Siebenthall, N. A. Sinnott-Armstrong, M. Stevens, R. E. Thurman, J. Wu, B. Zhang, X. Zhou, A. E. Beaudet, L. A. Boyer, P. L. De Jager, P. J. Farnham, S. J. Fisher, D. Haussler, S. J. Jones, W. Li, M. A. Marra, M. T. McManus, S. Sunyaev, J. A. Thomson, T. D. Tlsty, L. H. Tsai, W. Wang, R. A. Waterland, M. Q. Zhang, L. H. Chadwick, B. E. Bernstein, J. F. Costello, J. R. Ecker, M. Hirst, A. Meissner, A. Milosavljevic, B. Ren, J. A. Stamatoyannopoulos, T. Wang, and M. Kellis. 2015. Integrative analysis of 111 reference human epigenomes. *Nature* 518: 317-330.
41. Scharer, C. D., E. L. Blalock, T. Mi, B. G. Barwick, S. A. Jenks, T. Deguchi, K. S. Cashman, B. E. Neary, D. G. Patterson, S. L. Hicks, A. Khosroshahi, F. Eun-Hyung Lee, C. Wei, I. Sanz, and J. M. Boss. 2019. Epigenetic programming underpins B cell dysfunction in human SLE. *Nat Immunol* 20: 1071-1082.
42. Kent, W. J., C. W. Sugnet, T. S. Furey, K. M. Roskin, T. H. Pringle, A. M. Zahler, and D. Haussler. 2002. The human genome browser at UCSC. *Genome Res* 12: 996-1006.
43. Warming, S., N. Costantino, D. L. Court, N. A. Jenkins, and N. G. Copeland. 2005. Simple and highly efficient BAC recombineering using galK selection. *Nucleic Acids Res* 33: e36.
44. Li, H., and R. Durbin. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25: 1754-1760.
45. Zhang, Y., T. Liu, C. A. Meyer, J. Eeckhoutte, D. S. Johnson, B. E. Bernstein, C. Nusbaum, R. M. Myers, M. Brown, W. Li, and X. S. Liu. 2008. Model-based analysis of ChIP-Seq (MACS). *Genome Biol* 9: R137.
46. Huang da, W., B. T. Sherman, and R. A. Lempicki. 2009. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 4: 44-57.
47. Heinz, S., C. Benner, N. Spann, E. Bertolino, Y. C. Lin, P. Laslo, J. X. Cheng, C. Murre, H. Singh, and C. K. Glass. 2010. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell* 38: 576-589.
48. Grant, C. E., T. L. Bailey, and W. S. Noble. 2011. FIMO: scanning for occurrences of a given motif. *Bioinformatics* 27: 1017-1018.
49. Kieffer-Kwon, K. R., K. Nimura, S. S. P. Rao, J. Xu, S. Jung, A. Pekowska, M. Dose, E. Stevens, E. Mathe, P. Dong, S. C. Huang, M. A. Ricci, L. Baranello, Y. Zheng, F. Tomassoni Ardori, W. Resch, D. Stavreva, S. Nelson, M. McAndrew, A. Casellas, E. Finn, C. Gregory, B. G. St Hilaire, S. M. Johnson, W. Dubois, M. P. Cosma, E. Batchelor, D. Levens, R. D. Phair, T. Misteli, L. Tessarollo, G. Hager, M. Lakadamyali, Z. Liu, M. Floer, H. Shroff, E. L. Aiden, and R. Casellas. 2017. Myc Regulates Chromatin Decompaction and Nuclear Architecture during B Cell Activation. *Mol Cell* 67: 566-578 e510.
50. Jorcyk, C. L., L. J. Garrett, I. G. Maroulakou, D. K. Watson, and J. E. Green. 1997. Multiple regulatory regions control the expression of Ets-1 in the developing mouse: vascular expression conferred by intron I. *Cell Mol Biol (Noisy-le-grand)* 43: 211-225.
51. King, H. W., K. L. Wells, Z. Shipony, A. S. Kathiria, L. E. Wagar, C. Lareau, N. Orban, R. Capasso, M. M. Davis, L. M. Steinmetz, L. K. James, and W. J. Greenleaf. 2021. Integrated single-cell transcriptomics and epigenomics reveals strong germinal center-associated etiology of autoimmune risk loci. *Sci Immunol* 6: eabh3768.
52. Sacchetti, A., T. El Sewedy, A. F. Nasr, and S. Alberti. 2001. Efficient GFP mutations profoundly affect mRNA transcription and translation rates. *FEBS Lett* 492: 151-155.
53. Creighton, M. P., A. W. Cheng, G. G. Welstead, T. Kooistra, B. W. Carey, E. J. Steine, J. Hanna, M. A. Lodato, G. M. Frampton, P. A. Sharp, L. A. Boyer, R. A. Young, and R. Jaenisch. 2010. Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc Natl Acad Sci U S A* 107: 21931-21936.



54. Calderon, D., M. L. T. Nguyen, A. Mezger, A. Kathiria, F. Muller, V. Nguyen, N. Lescano, B. Wu, J. Trombetta, J. V. Ribado, D. A. Knowles, Z. Gao, F. Blaesche, A. V. Parent, T. D. Burt, M. S. Anderson, L. A. Criswell, W. J. Greenleaf, A. Marson, and J. K. Pritchard. 2019. Landscape of stimulation-responsive chromatin across diverse human immune cells. *Nat Genet* 51: 1494-1505.
55. Comoglio, F., M. Simonatto, S. Polletti, X. Liu, S. T. Smale, I. Barozzi, and G. Natoli. 2019. Dissection of acute stimulus-inducible nucleosome remodeling in mammalian cells. *Genes Dev* 33: 1159-1174.
56. Fisher, C. L., J. Ghysdael, and J. C. Cambier. 1991. Ligation of membrane Ig leads to calcium-mediated phosphorylation of the proto-oncogene product, Ets-1. *J Immunol* 146: 1743-1749.
57. Tsao, H. W., T. S. Tai, W. Tseng, H. H. Chang, R. Grenningloh, S. C. Miaw, and I. C. Ho. 2013. Ets-1 facilitates nuclear entry of NFAT proteins and their recruitment to the IL-2 promoter. *Proc Natl Acad Sci U S A* 110: 15776-15781.
58. Du, C., C. Liu, J. Kang, G. Zhao, Z. Ye, S. Huang, Z. Li, Z. Wu, and G. Pei. 2009. MicroRNA miR-326 regulates TH-17 differentiation and is associated with the pathogenesis of multiple sclerosis. *Nat Immunol* 10: 1252-1259.
59. Wen, Z., L. Xu, X. Chen, W. Xu, Z. Yin, X. Gao, and S. Xiong. 2013. Autoantibody induction by DNA-containing immune complexes requires HMGB1 with the TLR2/microRNA-155 pathway. *J Immunol* 190: 5411-5422.
60. Ostuni, R., V. Piccolo, I. Barozzi, S. Polletti, A. Termanini, S. Bonifacio, A. Curina, E. Prosperini, S. Ghisletti, and G. Natoli. 2013. Latent enhancers activated by stimulation in differentiated cells. *Cell* 152: 157-171.
61. Yukawa, M., S. Jagannathan, S. Vallabh, A. V. Kartashov, X. Chen, M. T. Weirauch, and A. Barski. 2020. AP-1 activity induced by co-stimulation is required for chromatin opening during T cell activation. *J Exp Med* 217.
62. Markov, G. J., T. Mai, S. Nair, A. Shcherbina, Y. X. Wang, D. M. Burns, A. Kundaje, and H. M. Blau. 2021. AP-1 is a temporally regulated dual gatekeeper of reprogramming to pluripotency. *Proc Natl Acad Sci U S A* 118.
63. Schreiber, M., A. Kolbus, F. Piu, A. Szabowski, U. Mohle-Steinlein, J. Tian, M. Karin, P. Angel, and E. F. Wagner. 1999. Control of cell cycle progression by c-Jun is p53 dependent. *Genes Dev* 13: 607-619.
64. Sundqvist, A., O. Voytyuk, M. Hamdi, H. E. Popeijus, C. B. van der Burgt, J. Janssen, J. W. M. Martens, A. Moustakas, C. H. Heldin, P. Ten Dijke, and H. van Dam. 2019. JNK-Dependent cJun Phosphorylation Mitigates TGFbeta- and EGF-Induced Pre-Malignant Breast Cancer Cell Invasion by Suppressing AP-1-Mediated Transcriptional Responses. *Cells* 8.
65. Bakiri, L., D. Lallemand, E. Bossy-Wetzel, and M. Yaniv. 2000. Cell cycle-dependent variations in c-Jun and JunB phosphorylation: a role in the control of cyclin D1 expression. *EMBO J* 19: 2056-2068.
66. Perez-Benavente, B., A. Fathinajafabadi, L. de la Fuente, C. Gandia, A. Martinez-Ferriz, J. M. Pardo-Sanchez, L. Milian, A. Conesa, O. A. Romero, J. Carretero, R. Matthiesen, I. Jariel-Encontre, M. Piechaczyk, and R. Farras. 2022. New roles for AP-1/JUNB in cell cycle control and tumorigenic cell invasion via regulation of cyclin E1 and TGF-beta2. *Genome Biol* 23: 252.
67. Zhu, X., B. L. Schweitzer, E. J. Romer, C. E. Sulentic, and R. P. DeKoter. 2008. Transgenic expression of Spi-C impairs B-cell development and function by affecting genes associated with BCR signaling. *Eur J Immunol* 38: 2587-2599.
68. Laramée, A. S., H. Raczkowski, P. Shao, C. Batista, D. Shukla, L. Xu, S. M. M. Haeryfar, Y. Tesfagiorgis, S. Kerfoot, and R. DeKoter. 2020. Opposing Roles for the Related ETS-Family Transcription Factors Spi-B and Spi-C in Regulating B Cell Differentiation and Function. *Front Immunol* 11: 841.
69. Roes, J., B. K. Choi, and B. B. Cazac. 2003. Redirection of B cell responsiveness by transforming growth factor beta receptor. *Proc Natl Acad Sci U S A* 100: 7241-7246.

70. Cazac, B. B., and J. Roes. 2000. TGF-beta receptor controls B cell responsiveness and induction of IgA in vivo. *Immunity* 13: 443-451.

## Figure Legends:

**Figure 1. Identification of the *Ets1* TAD and sub-TAD.** Chromatin interaction maps of mouse B cell Hi-C data showing the TAD (A) and sub-TAD (B) that include the *Ets1* gene. The aqua triangles in (B) denote the sub-TAD within the TAD and the interaction hotspot discussed in the text. Below the interaction maps is CTCF ChIP-seq data from splenocytes. The aqua highlights denote the TAD and sub-TAD boundaries that overlap with strong CTCF enrichment. FIMO was used to identify CTCF motifs at the peaks of CTCF binding and are depicted below the CTCF ChIP-seq profiles. The CTCF binding site located at the downstream boundary of the sub-TAD lacks a consensus CTCF motif and the sequence shown is similar to the CTCF binding motif.

**Figure 2. Design of a BAC-derived reporter construct to identify sequences required for lymphoid *Ets1* expression.** (A) Hi-C data from mouse B cells shows that the RP23-350A20 BAC (whose extent is denoted by the light beige block flanked by dotted lines) contains the entire hotspot of most frequent interactions (smaller blue triangle in the top part of the figure). The first exon of *Ets1* was replaced in the BAC by eGFP cDNA via recombineering. eGFP reporter expression is under the control of the putative response elements contained within the BAC. (B) qPCR and (C) Western blot to show expression of *Ets1* and GFP in various tissues of BACtg mice and non-transgenic mice.

**Figure 3. Expression of the BACtg eGFP in lymphoid cells.** (A) Flow cytometry analysis of intracellular *Ets1* and eGFP in gated B220<sup>+</sup>CD19<sup>+</sup> B cells from spleen of wild-type non-transgenic (WT) and BAC transgenic (BACtg) mice. The top figure shows GFP staining while the bottom figure shows intracellular *Ets1* staining. (B) Flow cytometry plots of intracellular *Ets1* versus eGFP staining in splenic B cell subsets: follicular B cells (B220<sup>+</sup>CD19<sup>+</sup> CD23<sup>hi</sup>CD21<sup>lo</sup>), marginal zone B cells (B220<sup>+</sup>CD19<sup>+</sup> CD23<sup>lo</sup>CD21<sup>hi</sup>), germinal center B cells (B220<sup>+</sup>CD19<sup>+</sup> Fas<sup>+</sup>PNA<sup>+</sup>), memory B cells (B220<sup>+</sup>CD19<sup>+</sup>CD80<sup>+</sup>PDL2<sup>+</sup>), plasma cells (B220<sup>lo</sup>CD138<sup>+</sup>CD98<sup>+</sup>), CD4 T cells (CD3<sup>+</sup>CD4<sup>+</sup>) and CD8 T cells (CD3<sup>+</sup>CD8<sup>+</sup>).

**Figure 4. The *Ets1* sub-TAD region contains several putative response elements highly enriched for markers of active response elements.** (A) Genome Browser image depicting B cell epigenetic profiles within the *Ets1* sub-TAD. Datasets shown include mouse B cell ATAC-seq (accessible chromatin) and ChIP-seq for H3K27ac (active response elements), H3K4me1 (active or poised response elements), CTCF (denoting the upstream sub-TAD boundary), and H3K4me3 (active promoter). Highlighted in light blue are sequences with ATAC-seq peaks and enrichment of H3K27ac and H3K4me1. Several transcript variants of the mouse *Ets1* gene are shown, but the major isoform expressed in B cells is *Ets1*/NM\_011808.3 (B) The human *ETS1* gene locus in regions homologous to the mouse BACtg. DNase I hypersensitivity is shown below. Also shown is H3K27ac ChIP-seq (active response elements) from GM12878 B lymphoma cells, which identifies seven discrete regions of enrichment (green highlights labeled 1-7). Two transcript variants are shown, but the major isoform expressed in B cells is *ETS1*/NM\_005238.4.

**Figure 5. Differences in *Ets1* expression in different B cell differentiation stages correlate with changes in chromatin accessibility in the mouse *Ets1* locus.** (A) Relative expression levels of *Ets1* in sort-purified cells at various stages in the differentiation from hematopoietic stem cells to mature B cells and to plasma cells based on RNA-seq data from the ImmGen immunocyte RNA-seq project (GSE109125). (B) ATAC-seq profiles for follicular (FO B) versus plasma cells (PC), based on ImmGen ATAC-seq data (GSE100738), within the sequences included in the RP23-350A20 BAC. Blue highlighted regions mark six areas with differential ATAC-seq peaks in FO B cells versus PC.

**Fig. 6. Identification of differentially-accessible regions (DARs) in response to BCR stimulation.** (A) Schematic of the experimental layout. Total splenic B cells from wild-type C57BL/6 mice were isolated and either left unstimulated or stimulated for 2 hours with BCR crosslinking antibody (anti-IgM). ATAC-seq was performed and DARs were identified. (B) Top seven biological processes identified by analyzing genome-wide DARs that either gained or lost accessibility upon BCR stimulation. Gene ontology analysis

was performed on genes within 10 kb of DARs using the DAVID software. (C) Transcription factor binding motifs enriched in DARs that gained (left) or lost (right) accessibility with BCR stimulation. Genome-wide DAR sequences were analyzed by HOMER for *de novo* transcription factor motif discovery. Shown are the top 3 *de novo* motifs enriched in each set of DARs and the best matching known consensus motif.

**Fig. 7. DARs in the *Ets1* gene locus contain motifs for B cell-relevant transcription factors.** (A) Shown are the ATAC-seq profiles in the *Ets1* locus (within the region of the BAC transgene) in unstimulated (NT) and stimulated ( $\alpha$ IgM) B cells. The third track shows the difference between the two datasets, with the positive peaks indicating sites that are more accessible in unstimulated and negative peaks indicating sites that are more accessible in stimulated B cells. DARs showing at least 2-fold changes in accessibility are highlighted blue for those that gained accessibility and highlighted in red for those that lost accessibility. (B) Close-up views of the DARs and transcription factor motifs they contain. FIMO was used to scan for transcription factor motifs within the four DARs. Depicted are the five highest scoring motifs. (C) ATAC-seq profiles compared with AP-1 subunits (c-Jun and JunD) ChIP-seq data from CH12 cells. The blue highlight denotes the +49 kb DAR.

**Fig. 8. Individual putative response elements do not mediate strong activation or repression in transient transfection assays.** (A) ATAC-seq profiles of follicular (FO) B cells and plasma cells (PC) in the *Ets1* locus (the region contained within the BAC transgene). Highlighted in light blue are larger fragments equivalent to Sites 1-3 and 5-7 that were incorporated into luciferase constructs. Highlighted in bright green are small regions containing differentially-accessible regions comparing follicular B cells to plasma cells (as shown in Figure 5B). Highlighted in pink is the region surrounding the proximal promoter and part of first intron that were previously tested in transgenic mice and shown to be insufficient for mediating lymphocyte-specific expression. (B) General schematic of the design of the luciferase constructs. (C-D) A20 cells were transfected with firefly luciferase plasmids containing the *Ets1* promoter alone or with the indicated response elements, or with empty vector, along with eF1 $\alpha$  promoter Renilla luciferase

internal control plasmid. Luciferase activity was measured 24 hours after transfection. Firefly luciferase activity was normalized to Renilla luciferase activity, then values were set relative to the plasmid with the Ets1 promoter alone. Significance was determined by one-way ANOVA. N=2-5 replicates for each transfection.

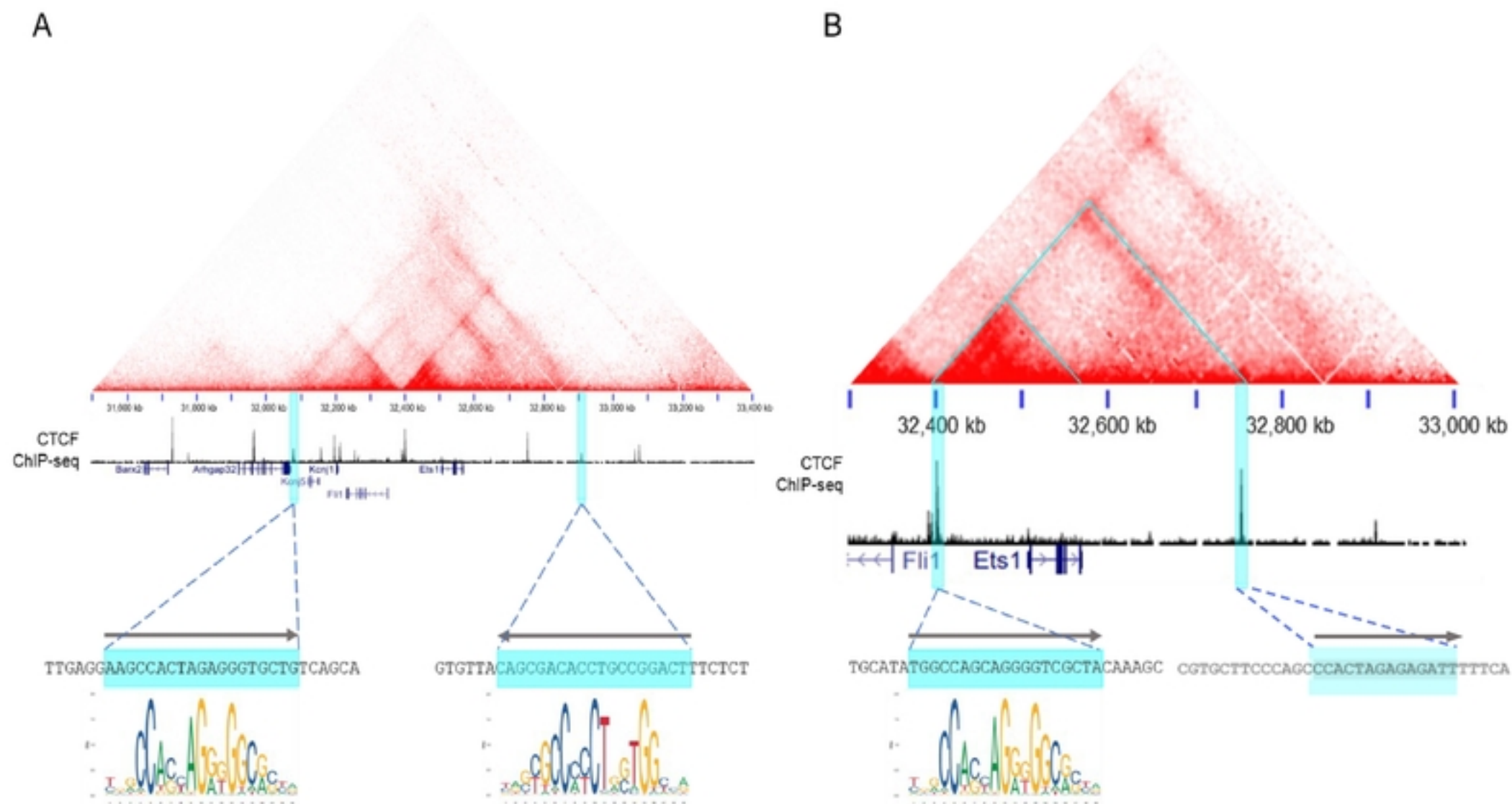
## Supplemental Figure Legends:

**Supplemental Fig. S1. Sequences within the RP23-350A20 BAC can mediate the BCR-induced decrease in transcription from the *Ets1* promoter.** Splenic B cells were isolated from wild type and BACtg mice, and were incubated with BCR crosslinking antibody for 2 hrs. *Ets1* mRNA and pre-mRNA levels as well as GFP mRNA levels were measured using qPCR and were normalized to beta-actin. N=2 for each condition.

**Supplemental Fig. S2: ATAC-seq shows that T cells contain regions of open chromatin that lie outside the region of the BACtg.** (B) ATAC-seq profiles for B cells, CD4 T cells, and CD8 T cells within the *Ets1* sub-TAD. The genomic region contained in the BAC transgene is highlighted in yellow. Highlighted in aqua are accessible sites outside of the BAC that may be T cell-specific *Ets1* response elements.

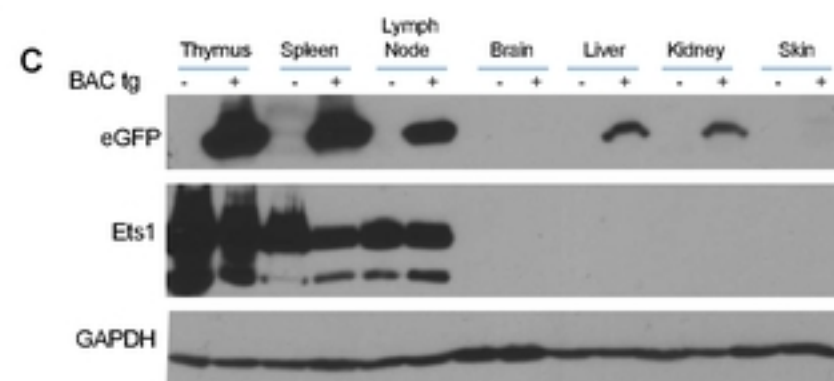
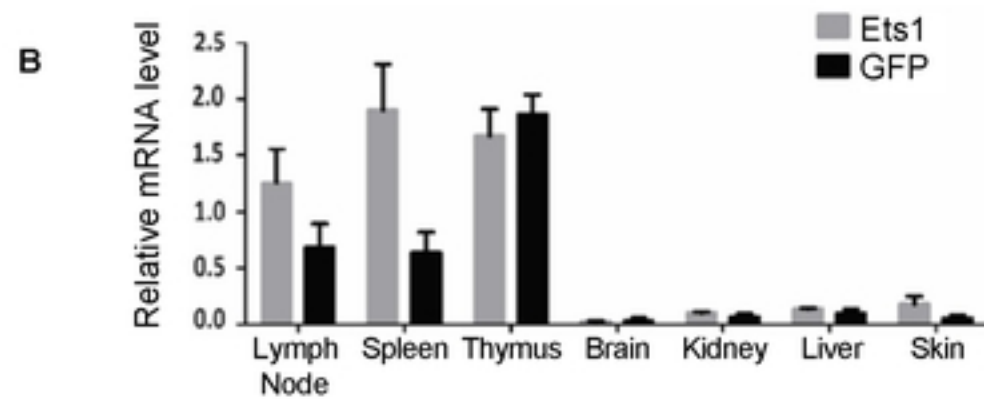
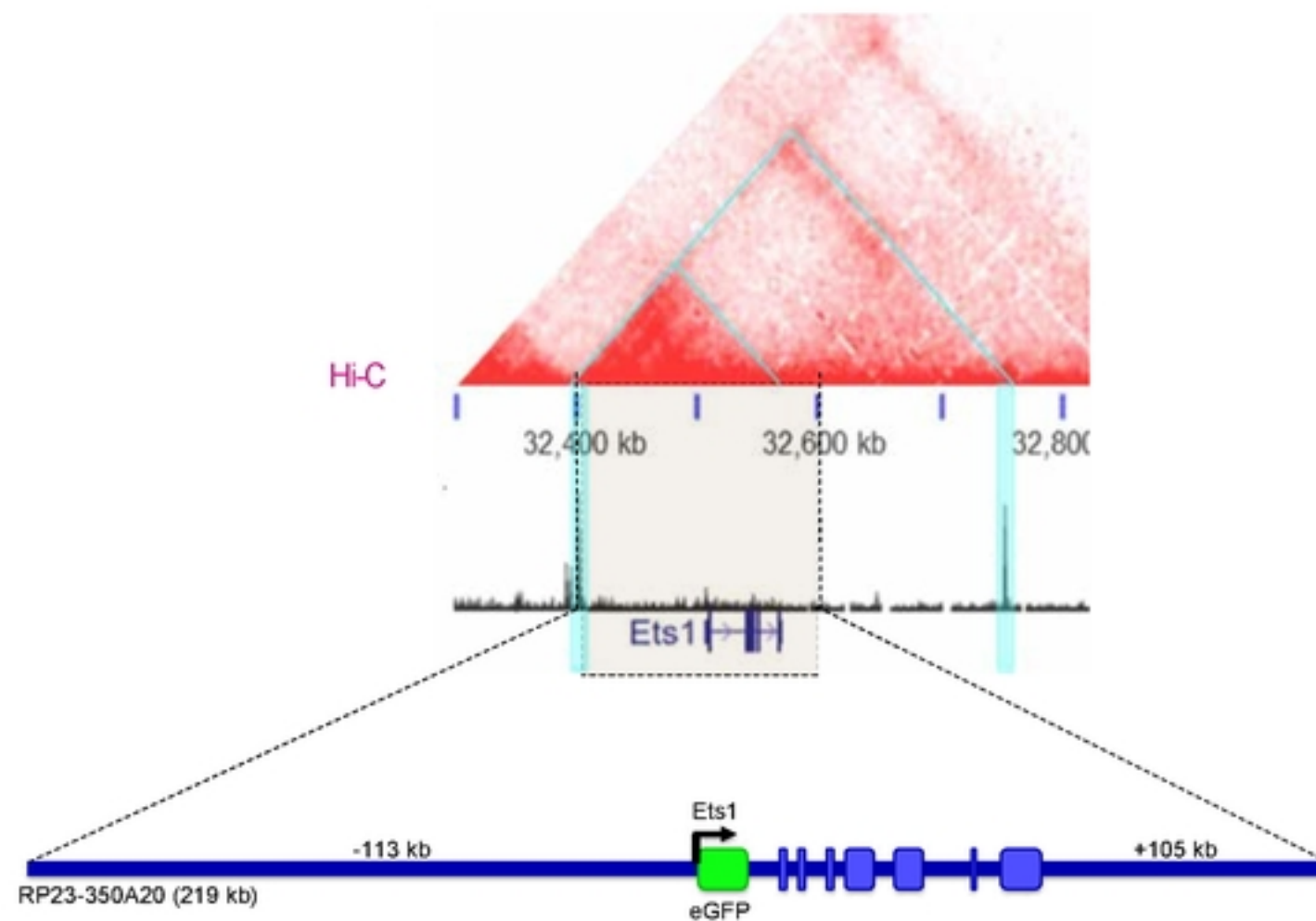
**Supplemental Figure S3: Sequence conservation between mouse and human genes encoding *Ets1* in regions with putative response elements.** Genome Browser image depicting the ATAC-seq peaks in follicular (FO) B cells and plasma cells (PC). Orange highlighted regions are regions within Sites 1-7 that show strong sequence homology between mouse the human. Underneath are shown Blat alignments for these segments. DNA sequence homologies outside Sites 1-7 are not shown. Also shown are the locations of the four DAR regions identified in Figure 7A, with the +39 and +49 kb DARs showing significant sequence conservation, while the -85 and +36 kb DARs lack such conservation.

**Figure 1**



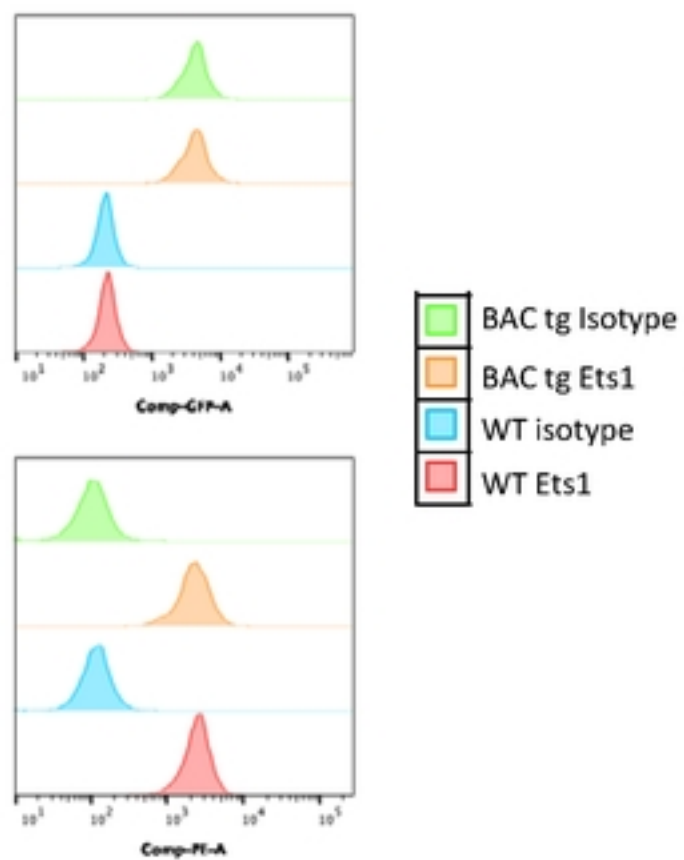


A

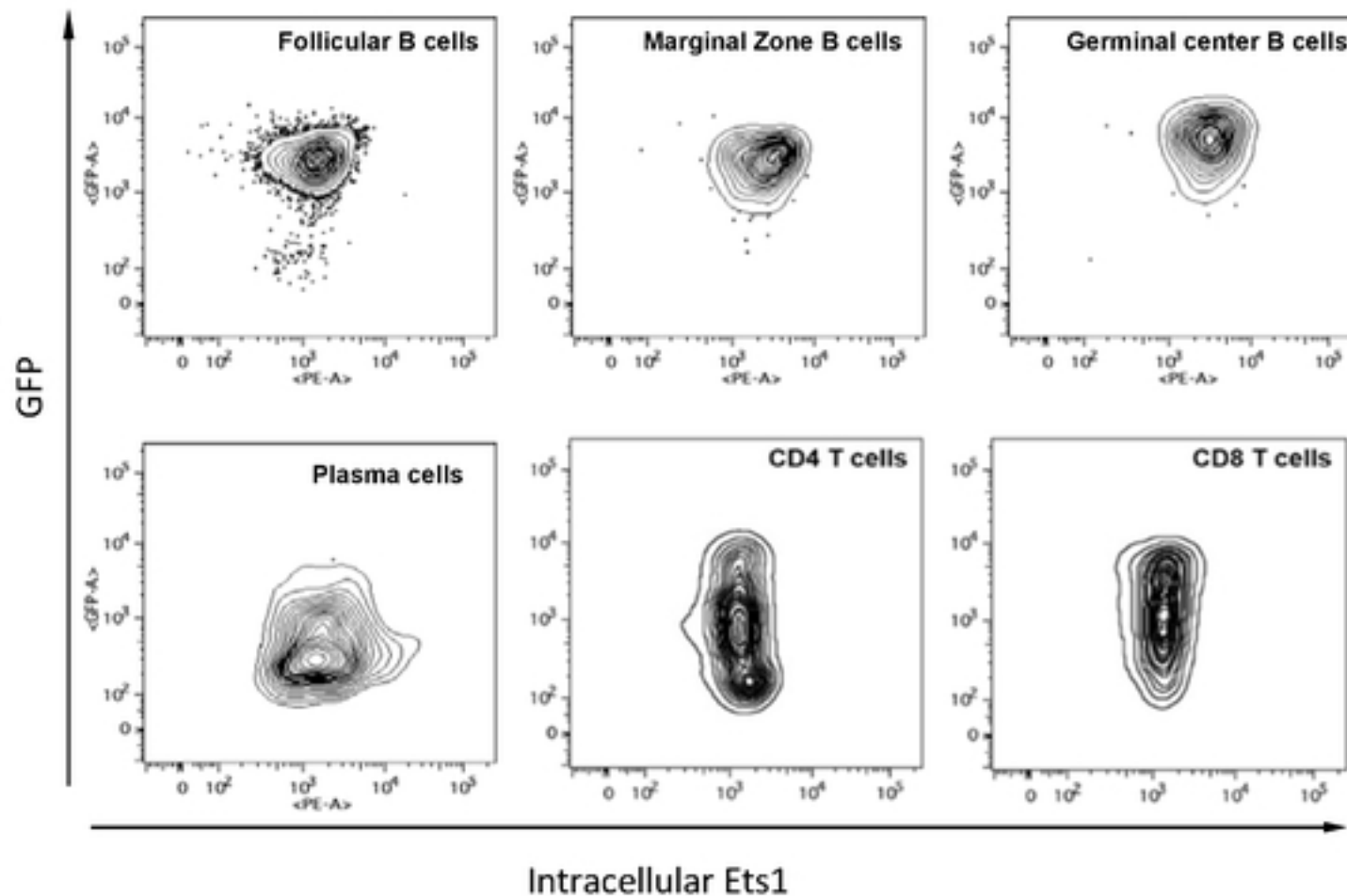


**Figure 3**

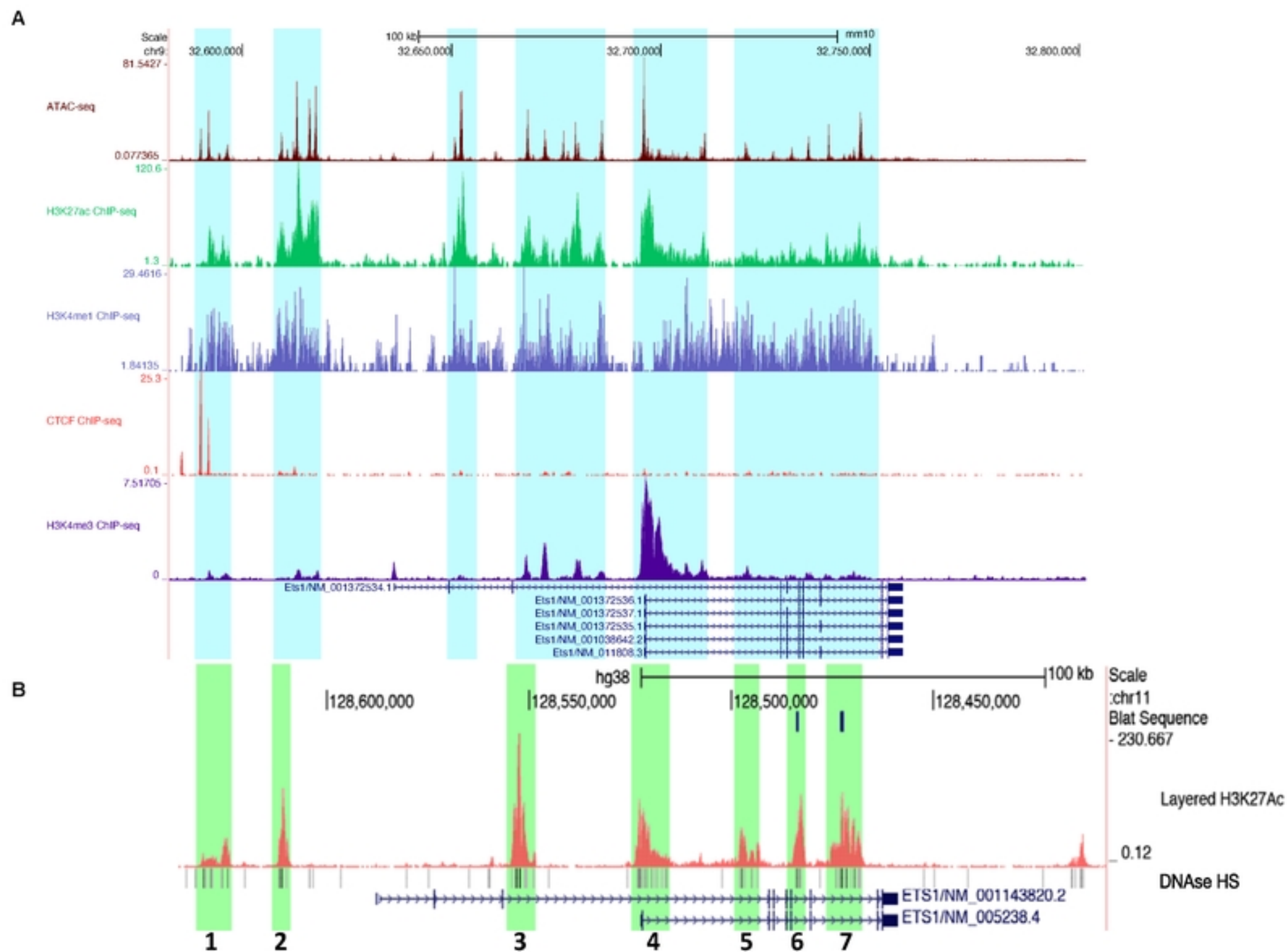
**A**



**B**

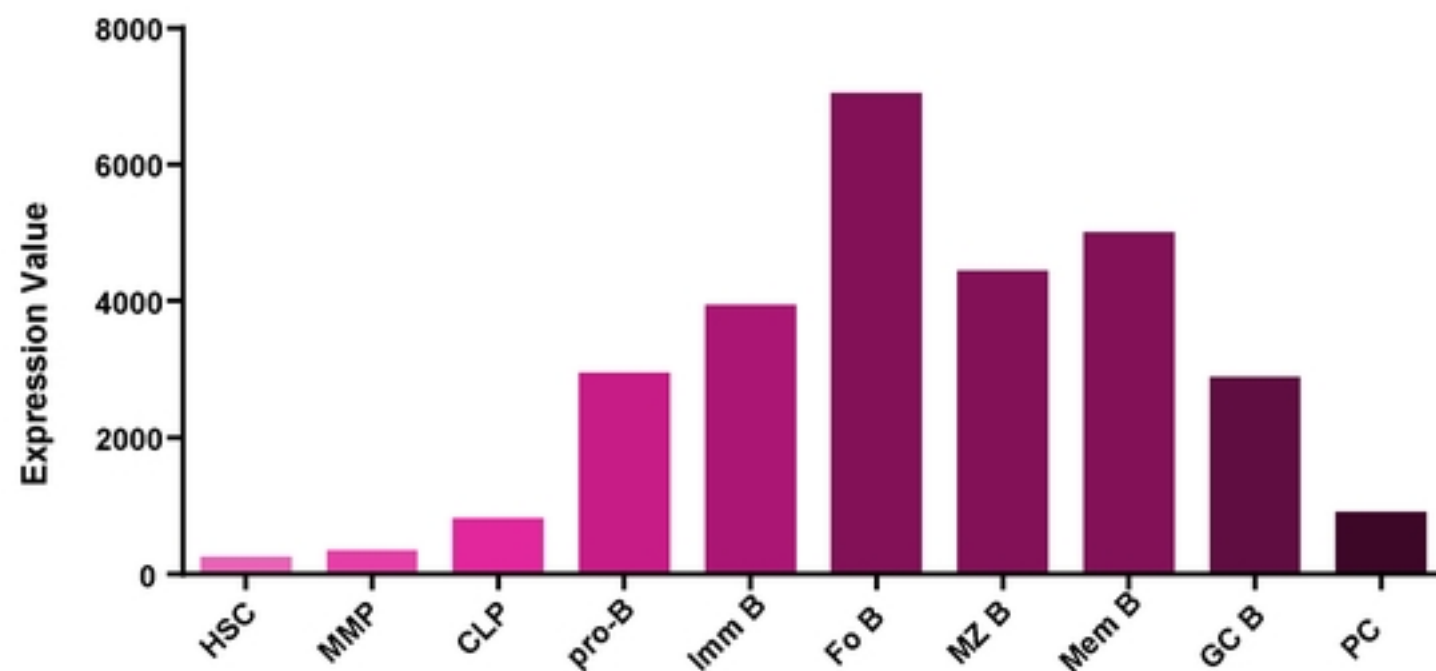


**Figure 4**



**Figure 5**

**A**



**B**

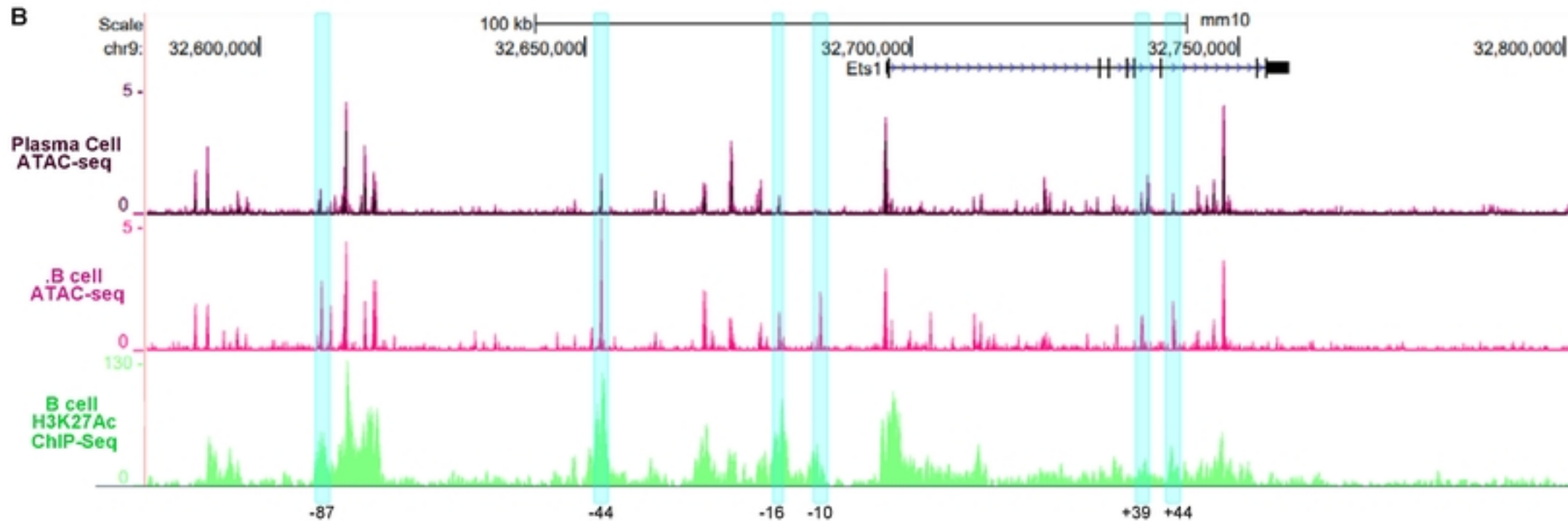
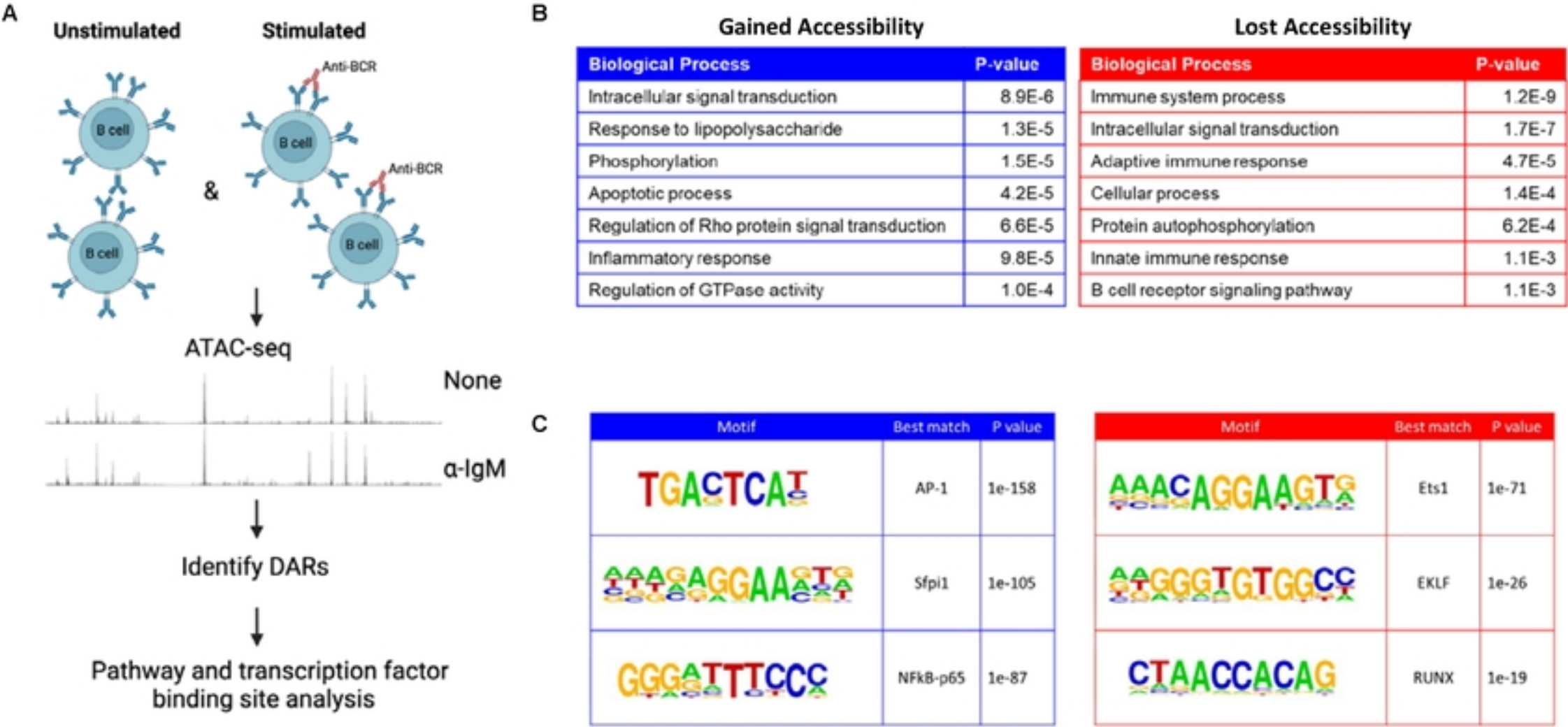
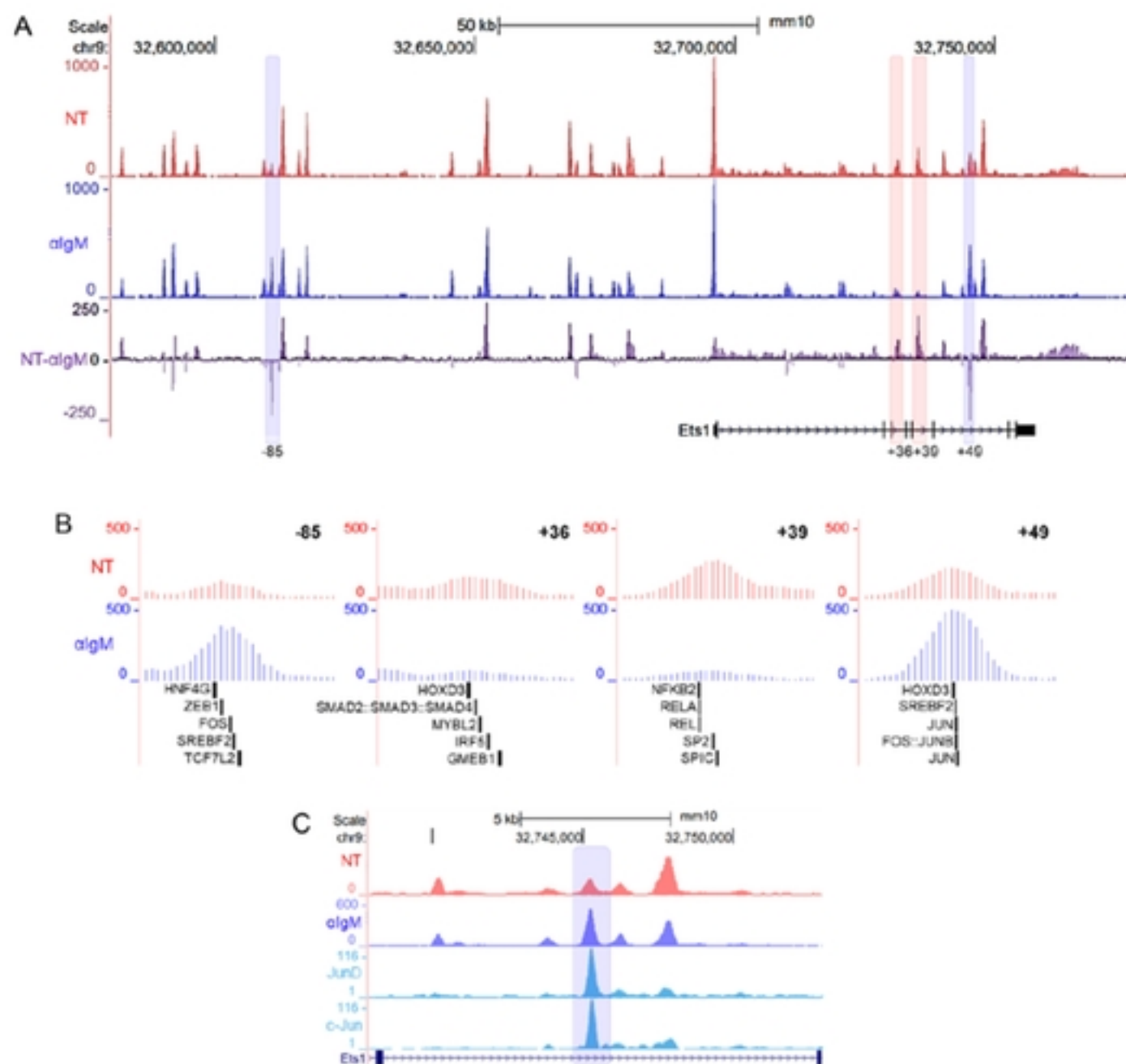


Figure 6



**Figure 7**





**Figure 8**

