OBF1 and Oct factors control the germinal center transcriptional program

Inauguraldissertation

zur

Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch-Naturwissenschaftlichen Fakultät

der Universität Basel

von

Shuang Song

von China

Basel, 2021

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät auf Antrag von

Prof. Dr. Patrick Matthias, Prof. Dr. Daniela Finke

Basel, den 2. March 2021

Prof. Dr. Marcel Mayor, Dekan

Summary

Germinal centers (GCs) are an essential structure of the humoral immune response. GC formation begins in the periphery lymphoid organs when follicular B cells encounter blood borne antigens. During the GC reaction, GC B cells undergo a series of molecular events, which ultimately lead to the generation of antibody-secreting plasma cells (PCs) and memory B cells (Bmems). Due to the mutagenic nature of GC B cells, GCs give rise to majority of non-Hodgkin lymphomas (NHLs), including Burkitt's lymphomas (BL) and most diffuse large B-cell lymphomas (DLBCL). OCT1 and OCT2 are transcription factors (TFs) of the POU family. They recognize the same set of octamer motif and derivatives thereof. OBF1 is a B cell-specific coactivator which interacts with OCT1 and/or OCT2 on binding sites. OCT2 and OBF1 have been shown to be critical for GC formation. However, so far, the functional relevance of OCT1 in B cell or GC formation is largely overlooked. Moreover, little is known about which stage of GC formation is dependent on OCT2 and OBF1. Furthermore, the mechanisms how these factors regulate GC specific target genes and GC process remain elusive.

The main part of this thesis deals with the genomic analysis of OCT1, OCT2 and OBF1 in B cells to understand their functional relevance in GC formation and GC-derived lymphoma cells. Previous work from our laboratory and other groups has demonstrated that OBF1 and OCT2 are indispensable for GC formation. We mapped the global binding pattern of these three factors, and interrogated the mechanisms of these factors in GC transcriptional program using genomic approaches and computational analysis. We found that these factors extensively co-localize with each other and with ETS factors. We confirmed that OBF1 stabilizes the genomic binding of OCT1 on chromatin. Moreover, we revealed that OBF1 maintains the GC transcriptional program by activating the expression of BCL6 and repressing the expression of IRF4. We demonstrated that the proliferation of GC-derived B lymphoma cells is dependent on OBF1, and that loss of OBF1 leads to GC exit and the initiation of post-GC differentiation program. Furthermore, we showed that OBF1 binds to the regulatory elements of GC-related genes in primary murine and human GC B cells. We discovered that depletion of OBF1 in B lymphoma cells was correlated with an upregulation of genes associated with favorable prognosis.

Therefore, OBF1 maintains the integrity of the GC regulatory network, and represents a promising therapeutic target for GC-derived B lymphoma cells.

The second part addresses the enhancer dynamics of stages in B cell differentiation. We performed and compared the ChIP-seq for H3K4me3, H3K4me1, H3K27ac, H3K27me3 and H3K9me2 in haematopoietic stem cells (HSCs), pro-B cells and splenic mature B cells. In contrast to the prevailing model – stage specific enhancer repertoire is primed in the early developmental stages prior to terminal differentiation, we found that the majority of enhancers are *de novo* established in the stage where their functions are required, without being primed in the preceding stages. Moreover, we confirmed that the H3K9me2 landscape is largely unchanged during the differentiation from HSCs to splenic mature B cells, and that enhancer dynamics are largely uncoupled with PcG-mediated silencing. In sum, our data shed new light on the epigenomic reprogramming during the progression from stem cells to differentiated cells.

Therefore, in this thesis, I demonstrate extensive genomic co-occupancy between OCT factors and OBF1, as well as ETS factors. I identify the target genes of OBF1 and functional mechanisms in regulating the GC reaction. In particular, I uncover the detailed functional dependence of OBF1 in controlling the proliferation of GC-derived B lymphoma cells, and propose it to be a novel therapeutic target for future B lymphoma treatment.

Table of Contents

Introduction

Chapter 1 Introduction

1.1 B cell development

B cell differentiation initiates from early days of embryonic development. Undifferentiated pluripotent haematopoietic stem cells (HSCs) are generated from the extra-embryonic yolk sac from E7.5. They establish contacts with blood endothelial cells. The latter produce c-kit ligand, and CXCL10, and eventually induce the homing of HSCs into developing lymphoid organ, such as fetal liver, thymus and bone marrow [1-4].

1.1.1 B cell development during embryonic stage

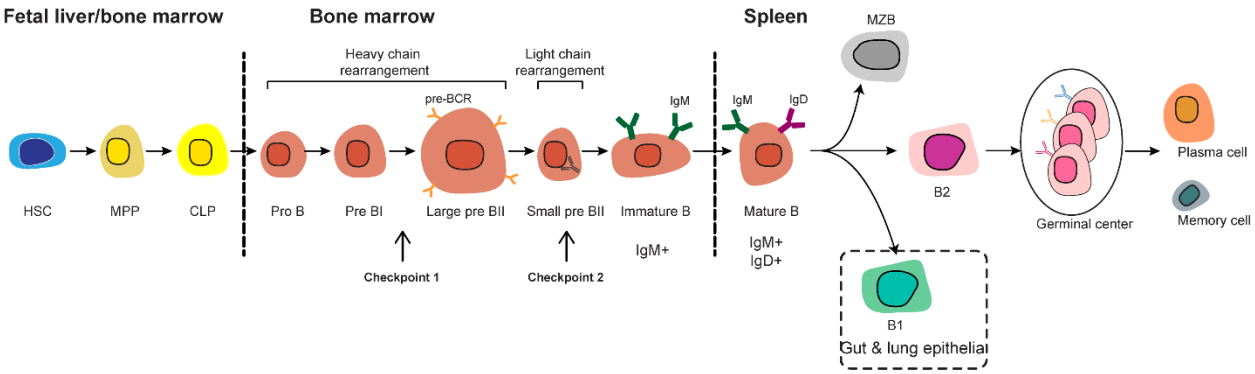


Figure 1. B cell development. The scheme shows the developmental stages of B cell development. The development of B cell originates from HSCs. B cell development initiates in the fetal liver before birth and continues in the bone marrow postnatally, while late B cell development takes place in peripheral lymphoid organs, e.g. spleen and lymph nodes. The rearrangement of heavy and light chain, the so-called VDJ recombination, represent two checkpoints during early B cell development. Pre-B cells with productive immunoglobulin heavy chain express the pre-BCR on the cellular surface and stimulate the proliferation of large pre-B II cells. Light chain V_L and J_L rearrangement takes place in small pre-B II cells. B cells with productive recombined light chain genes express the BCR on the surface and differentiate into immature B cells. B cell with non-productive or self-reactive heavy or light chain recombination undergo apoptosis. Immature B cells then migrate into peripheral lymphoid organs, in which they further differentiate into marginal zone B cells, B1 cells and B2 cells. B2 cells reside in the spleen B cell follicles and develop into germinal centers when encountering blood-born antigens. Eventually, GC B cells differentiate into antibody-secreting plasma cells and memory B cells.

During the developmental stages in fetal livers, a subset of MPPs and CLPs express IL-7Rα and commit to the lymphoid differentiation path[5]. These progenitor cells interact with mesenchymal and epithelial cells[1] and differentiate into FLT3$^+$c-kit$^+$IL-7Rα$^+$CD19$^-$ progenitors, and further generate FLT3$^+$c-kit$^+$IL-7Rα$^+$CD19$^+$ progenitors under the stimulation of Fms-related tyrosine kinase 4 ligand (FLT3L)[6]. Under current view, the beginning of the B lineage in embryonic environment is marked by the expression of the transcription factor E2A, and the recombinases RAG1/2, which constitute the V(D)J recombination machinery. Subsequently, these cells express EBF1, Igα/β, and surrogate light chain molecules (SLC, VpreB and λ5)[1]. Finally, the expression of Pax5 induces the differentiation of B cells[7,8]. These cells start immunoglobulin heavy chain (*IgH*) D-J rearrangement, which takes place on both IgH alleles[9-11] (Figure 2A). The IgH DJ rearranged CD19$^+$c-kit$^+$IL-7Rα$^+$ proB cells undergo rapid proliferation under the stimulation of IL-7, a cytokine which prevents apoptosis and further differentiation of early B cells[12]. Figure 1 illustrates the development of the B cell lineage.

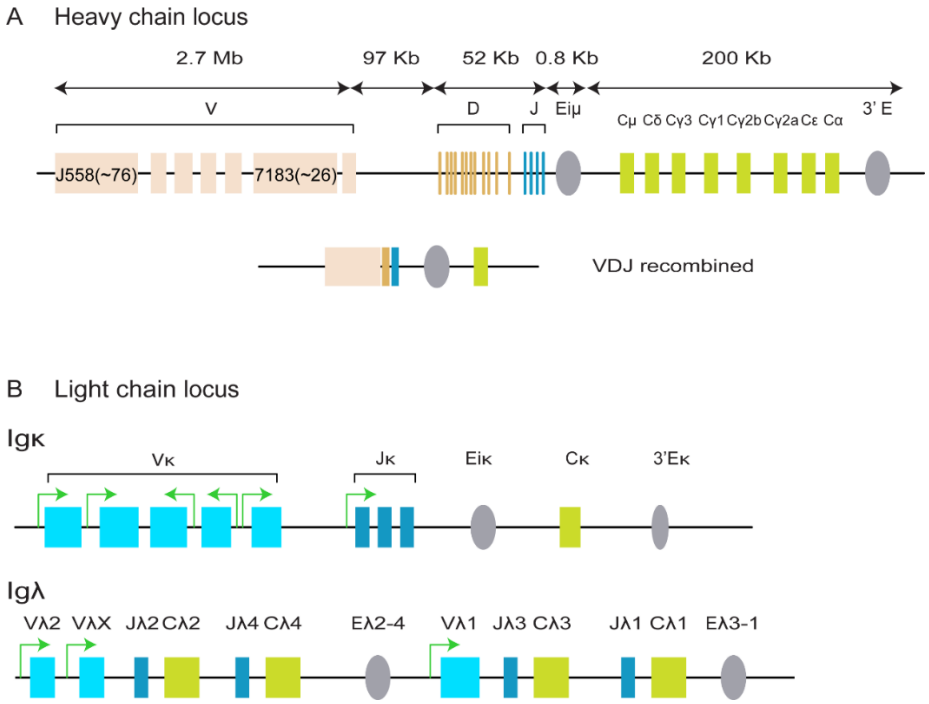


Figure 2. Structure of murine immunoglobulin heavy and light chain. (A) Organization of the murine immunoglobulin heavy (*IgH*) chain gene locus. Gene segments of variable (V) (pink), diversity (D) (orange), joining (J) (blue), and constant (C) (green) regions are depicted in color coded boxes. Enhancer clusters are shown as gray elliptical shapes. The resulting VDJ rearranged locus is shown as indicated. (B) Organization of the murine

immunoglobulin κ and λ light chain gene locus. Gene segments of variable (V) (blue), joining (J) (dark blue), and constant (C) (green) regions are depicted in color coded boxes. Enhancer clusters of the Ig light chain loci are shown as gray elliptical shapes.

1.1.2 B cell development in bone marrow

Bone marrow becomes the center of B cell differentiation after E17.5 and B cells keep being produced therein during the entire lifetime [5]. Bone marrow stromal cells promote the formation of long-term B cell niches by the expression of CXCL12 and establish close contact with HSCs. In these niches, multipotent progenitors (MPPs), common lymphoid progenitors (CLPs) and common myeloid progenitors (CMPs) are maintained and renewed. Therefore, paving the way for continual B cell production [13].

In the bone marrow, IL-7 is indispensable for the MPP/CLP to proB cell differentiation. Mice deficient of IL-7 display dramatically impaired B cell differentiation [14,15]. *IgH* DJ rearranged proB cells migrate towards and interact with mesenchymal stromal cell expressing IL-7, which establishes long lasting B lymphopoiesis [16]. However, human B cells differentiation does not require IL-7, thus it remains to be discovered which cytokine is critical for human B cell development.

ProB cells start to further differentiate when detached from IL-7 and stem cell factor (SCF) expressing stromal cells, which results in the attenuation of their fast proliferation [16]. Meanwhile, V to DJ rearrangement is induced on one of the *IgH* chain alleles in preBI stage. V-DJ recombination generates a wide spectrum of repertoires of antigen-binding domains on IgH. Successful V-DJ rearranged in-frame *IgH* is subsequently expressed on cellular surface either alone or together with the surrogate light chain (SLC) to form the pre-B cell receptor (pre-BCR) complex. Both of these two forms provide developmental critical signals to block the secondary V-DJ rearrangement on the other *IgH* locus, thus ensuring that any one B cell expresses a unique IGH molecule[9-11]. This process is called allelic exclusion. To be noted, D_H to J_H recombination almost always happens on both alleles and prior to the joining between V_H to D_H [17]. Nonetheless, V-DJ only always happens at one allele each time, therefore ensures clonal selection[18]. The pre-BCR signaling, mediated by Igα/Igβ, induces the downregulation of RAG1/2, as well as the upregulation of transcription factors (TFs) that are involved in Ig light chain rearrangement [19-21].

In addition, pre-BCR provides survival signals and stimulates the proliferation, therefore inhibits apoptosis [22]. The cells that express the pre-BCR but fail to transmit normal pre-BCR signals will undergo apoptosis. Moreover, the pre-BCR further signals to decrease the expression of SLC [23,24]. Therefore, as the proliferation goes on, new pre-BCR molecules cease to be produced and the existing ones become diluted after each round of cell cycle. In the end, large preBII cells enter the small preBII stage, which are resting cells. Small preBII cells which fail to generate enough cellular surface pre-BCR or show autoreactivity would soon be eliminated. This process is called negative selection.

In general, Immunoglobulin (*Ig*) light chain V to J rearrangement initiates after the attenuation of pre-BCR signaling in the dividing large preBII and resting small preBII cells [5]. Downregulation of pre-BCR signaling is essential for the B cell differentiation post *IgH* rearrangement; failure to silence the expression of SLC leads to continued pre-BCR signaling and preB malignancy [25,26]. V(D)J recombination is generally perceived as a highly ordered process, in which *IgH* rearrangement takes place before Ig light chain (*IgL*) recombination[18,27,28]. The *IgL* rearrangement proceeds with the re-expression of RAG1/2, increased accessibility, and sterile transcription of *IgL* loci [29,30]. However, to be noted, some studies showed that *IgL* recombination can be independent of *IgH* expression [31,32]. Therefore, it is controversial whether *IgL* rearrangement is dependent on *IgH* rearrangement or not. Moreover, *IgL* recombination has been shown to also take place in early B cell stages [31,33-35]. Figure 2B summarizes the light chain rearrangement process. Ig light chain rearrangement first takes place on one of the two alleles. If the first rearranged allele is non-productive, the second allele becomes open for another round of *IgL* V-J recombination[36] (Figure 6). In the end, rearranged *IgL* proteins are expressed, and interact with the arranged *IgH* proteins to form the B cell receptor (BCR) complexes on cellular surface[37]. B cells harboring rearranged *IgH* and *IgL* chains and surface BCRs are called immature B cells. BCRs are then tested for autoreactivity in the bone marrow microenvironment.

At the stage of immature B cells, central B cell tolerance is established which prevents the destructive autoimmune response of B cells to the host. At this stage, the majority of BCR-expressing immature B cells carry autoreactive properties, as the bone marrow microenvironment is abundant in self-antigens. RAG proteins continue to exist in autoreactive immature B cells, to facilitate receptor editing before the autoreactivity is corrected [30,38-41]. Receptor editing plays a key role in the establishment of central tolerance. Multiple evidence

from experiments with mouse models deficient in receptor editing showed that their B cells extensively undergo apoptosis[42-47]. Autoreactive immature cells, which persist without correction, will be eliminated by apoptosis[48].

1.1.3 V(D)J recombination

During V(D)J recombination, the exons encoding antigen-binding regions of Ig are randomly assembled by RAG1/2-mediated double-strand DNA break[49] and subsequent error-prone DNA repair process[50]. V(D)J recombination takes place at the B cell developmental stages in the bone marrow. During this process, RAG1/2 recognizes the recombination signal sequences (RSSs), conserved DNA sequences located in the 3'-end, 5'- and 3'-end, and 5'-end of each V, D and J segment, respectively. RSSs consist of a heptamer sequence on the 5'-end, a nonamer sequence on the 3'-end, and a 12/23 base spacer sequence in between (Figure 3). The consensus heptamer sequence is 5'-CACAGTG-3', and the consensus nonamer sequence is 5'-ACAAAACC-3' (highly conserved nucleotides are underscored)[51]. The order of recombination between each V, D and J segments is governed by the so-called "12/23 rule", in which RSSs of a 12bp spacer can only be joined with that of a 23bp spacer. In fact, 23bp RSSs flank V_H and J_H gene segments, whereas D_H gene segments are flanked by 12bp RSSs. Therefore, under the "12/23 rule", for each of the rearranged *IgH* locus, a D_H segment always occurs in between a V_H and J_H segment [52,53]. Terminal deoxynucleotidyl transferase is responsible for the diversification of V_H-D_H and D_H-J_H junctions by nucleotide deletion and non-templated addition of N-nucleotides in a random manner[54].

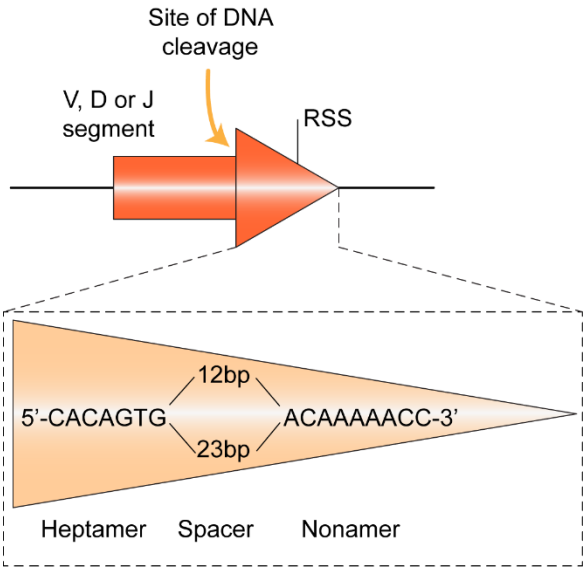


Figure 3. Structure of recombination signal sequence (RSS). The consensus RSS is shown, with the heptamer and nonamer flanking the 12/23 spacer region. RSSs, lie right next to the V, D or J gene segments, are highly conserved sequences. RSS consists of two highly conserved elements (heptamer and nonamer) which flank a spacer sequence. The first three nucleotides are the most important for the function of a RSS. The nonamer provides a binding site for the docking of RAG proteins. The length of the spacer is highly conserved with either 12 or 23 base pairs, while the sequences of spacers show more variability. This figure is partially adapted from[55].

When V(D)J recombination is finished, the germline promoter of V_H segment is placed in close range to *iEμ*, a strong enhancer located between J_H and *Cμ* exons (Figure 2A, Figure 4)[18]. This allows the transcription from rearranged V_H segment through the constant region, which is followed by RNA splicing and results in the generation of a *IgH* mRNA with variable regions connected to *Cμ* region, or *Cδ* region, in class-switched B cells[56] (figure 4).

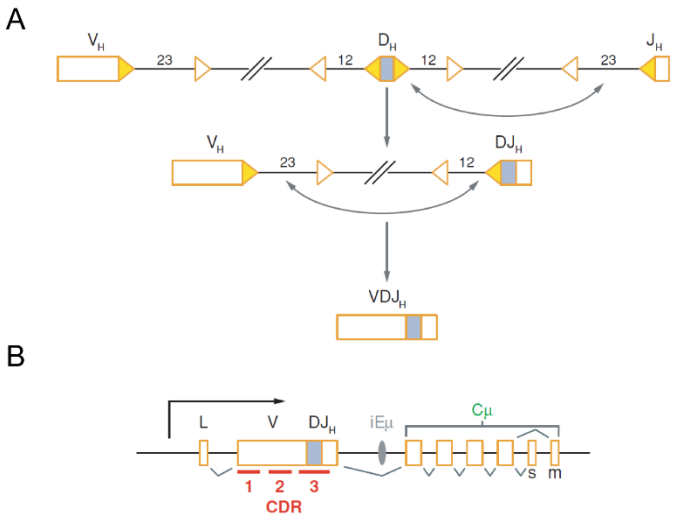


Figure 4. Recombination of IgH chain gene segments. (A) Gene segments of V, D and J regions, together with the corresponding recombination signal sequences (RSS), are shown. The heptamers are illustrated as yellow triangles, and nonamers are illustrated as empty triangles. The length of the RSS spacers is indicated. (B) The recombined *IgH* locus and transcriptional initiation site are shown. The position of the complementarity determining regions (CDRs, antigen-specific) of recombined VDJ_H segments are indicated. Possible splicing patterns are illustrated. (from Jung et al., 2006[18])

Allelic exclusion is a regulated process (Figure 5) [57,58], in which the rearranged functional *IgH* or *IgL* products prevent recombination on the other allele through a feedback inhibition mechanism. In the case of *IgH* alleles, D_HJ_H joining takes place on both alleles, whereas V_H-D_HJ_H joining takes place at one allele at a time leaving the other allele untouched. However, only B cells harboring an unproductive $V_HD_HJ_H$ assembly on the first allele will allow second V_H-D_HJ_H recombination on the second allele[18].

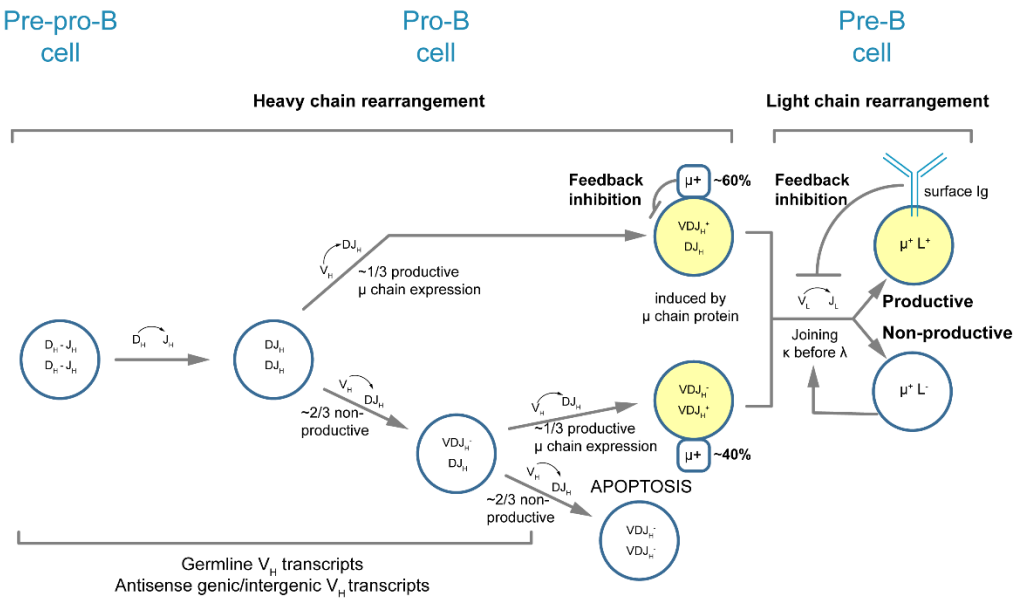


Figure 5. Model of allelic exclusion. For the *IgH* recombination, DJ_H recombination takes place on both alleles. However, V_H to DJ_H recombination only takes place on one of the alleles, the second allele initiates recombination only when the recombination is unproductive at the first allele. If recombination on both alleles is nonproductive, these B cells undergo apoptosis. Light chain recombination first starts from *Igκ* region. *Igλ* locus initiates recombination when *Igκ* region is non-productive. This figure is partially adapted from reference [18].

The configuration of light chain gene segments is different from that of *IgH* locus. Light chain genes consist of two sets of genes – *Igκ* and *Igλ* loci – present in all mammalian species[36]. Within the *Igκ* locus, similar to the *IgH* locus, Vκ genes are located on the 5' of Jκ genes, and a single *Cκ* gene is located downstream of Jκ genes. The organization of the *Igλ* locus is different from *Igκ*. Within the *Igλ* locus, *IgL*V genes are positioned at 5' of a set of *IgL*J-C pairs. The recombination of light chain genes starts from the Igκ region [58,59]. If the initial recombination of *Igκ* recombination generates unfunctional or self-reactive products, secondary recombination termed receptor editing will then take place, and leads to a higher chance of the generation of in-frame assembly[60]. Receptor editing requires the availability of unused V or J genes on the 5' and 3' of a rearranged VJ_L[36]. After all, combinations are rearranged on one allele, recombination will then occur on the second allele. If recombination on the second *Igκ* still yields unproductive recombination, then *Igλ* rearrangement starts (Figure 6).

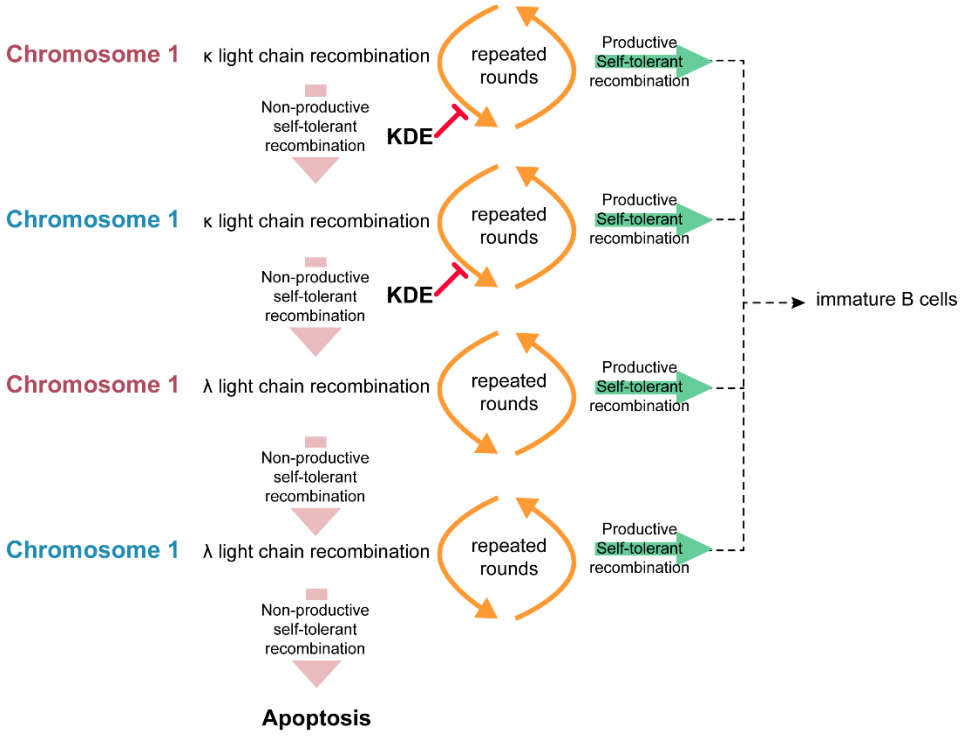


Figure 6. Light chain rearrangement. The recombination of *Ig* light chain initiates from the *Igκ* locus at one chromosome. The recombination which yields self-reactive antibodies can be modified further by repeated receptor editing. If the process of receptor editing fails to produce productive and self-tolerant antibodies, the same process will continue on the other chromosome. Kappa deleting element (KDE), a DNA sequence locates at the 3' *Igκ* locus, cancels the *Igκ* receptor editing before switching to *Igλ*. If the second chromosome still fails, then the light chain recombination switches to the λ locus of one chromosome at a time[36]. B cells which have failed to generate productive and self-tolerant light chain rearrangement after these step eventually under apoptosis. In contrast B cells having successfully undergone light chain rearrangement differentiate into immature B cells. This figure is partially adapted from reference[36].

1.1.4 Mature B cell development

Immature B cells migrate from the bone marrow into the spleen; this process is mediated by a chemoattractant, sphingosine-1-phosphate, which overrides the retention signal of CXCR4/CXCL12 in the bone marrow. Antigens in the spleen induce anergy and apoptosis of these transitional B cells, which are different from mature B cells in their response to pathogens or microbes[61]. Therefore, this represents another layer of regulation to eliminate autoreactive B cells before they finally differentiate into mature B cells. In mice, a marginal loss of transitional B cells to mature B cells is observed. Whereas roughly 20% to 40% cellular loss is observed in

human spleen[62,63]. To be noted, patients diagnosed with systemic systemic lupus erythematosus (SLE) and other autoimmune diseases show a much less cellular loss in this stage[64], indicating the importance of eliminating autoreactive B cells from transitional B cells to mature B cell differentiation. This process establishes peripheral tolerance. Immature B cells without autoreactivity eventually differentiate to mature B cells.

1.2 Germinal centers (GCs)

Antibodies are specialized proteins that specifically neutralize invading pathogens or microbes. The generation of high-affinity antigen-specific antibodies takes place in a dynamic, transient, and specialized structure called the GC. GCs play a pivotal role in acquired immunity. GCs form within peripheral lymphoid organs in response to blood-borne antigens. Within GCs, mature B cells proliferate at a rate that is unparalleled in mammalian tissues, and their antibody affinities are diversified by somatic hypermutation (SHM) mutation that specifically takes place on the Ig variable region genes. B cells expressing high-affinity antibodies subsequently exit the GC program and differentiate into antibody-secreting PCs and Bmems that mediate and sustain protection against invading pathogens for an extended period (Figure 7).

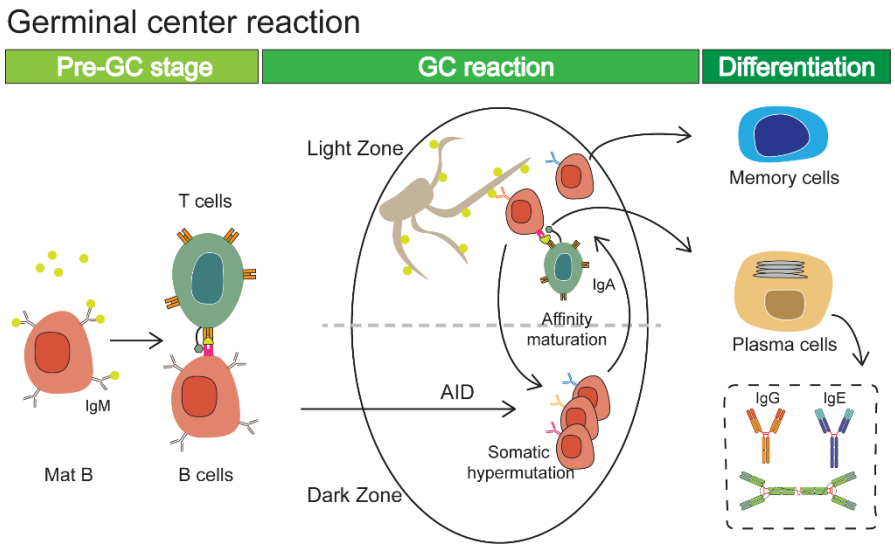


Figure 7 GC formation. In peripheral lymphoid organs, when mature B cells encounter an antigen, they proliferate rapidly, interact with T follicular helper cells (T_{FH}), and differentiate into GCs. GCs are divided into dark zones (DZ) and light zones (LZ). In DZ, GC B cells undergo fast proliferation and SHM takes place which introduces

mutations to *Ig* locus. This eventually produces a population of B cells with a repertoire of antibodies with a spectrum of affinities towards certain antigens. Then, GC B cells migrate into LZ, in which the affinities of mutated antibodies are tested through the interaction with FDCs and T_{FH} cells. This process is called affinity maturation. Finally, B cells with low affinities predominantly differentiate into Bmems, while those with high affinity differentiate into antibody-secreting PCs.

1.2.1 Initiation of GC

Under normal conditions, peripheral lymphoid organs harbor follicles which mainly contain naïve IgM^+IgD^+ B cells[65]. T cell zones exist adjacent to B cell follicles in these organs. GC reaction initiates when mature B cells in the follicles encounter antigens[66], and then antigen-stimulated B cells proliferate and form long-lasting cellular interactions with follicular T helper (T_{FH}) cells at the border region between B cell follicles and T cell zones[67,68].

Activated B cells then migrate to the GCs and differentiate to GC B cells. Recently, by using two-photon intravital microscopy to track the movements of GC precursor B cells and T_{FH} cells in lymph nodes at the initiation phase of GC reaction following immunization of NP-OVA[69,70], researchers found that already one day post immunization, activated B cells and T_{FH} cells interact with each other and migrate into interfollicular regions of lymph nodes. BCL6 level is elevated in GC B cells two days post stimulation[71]. At day three, pre-GC B and T cells eventually migrate into B cell follicles, and early GC structures are formed at day four. It is worth mentioning that different antigens and peripheral lymphoid organs might result in distinct kinetics of GC reaction[72]. In general, these different studies agree on the observation that the GC reaction is quickly initiated once B and T cells encounter antigen. On day four, B cells proliferate rapidly in the center of the B follicle within the underlying network of follicular dendritic cells (FDCs) and form secondary follicles. B cells proliferate rapidly in the GCs, and the size of GCs keeps growing until five or six days post immunization. On day seven, GCs become fully established and contain two compartments, the dark zone (DZ) and light zone (LZ)[65,73] (Figure 8), in which GC B cells undergo distinct biological processes.

Instead of entering GCs, activated B cells can also migrate to the extrafollicular region, and proliferate and differentiate there into short-lived plasmablasts which produce low-affinity antibodies [74,75]. Although the mechanism that determine this fate choice is unclear, it is possible

that the antigen affinity of the BCR, the intensity of antigen-BCR interaction, and costimulatory signals from T_{FH} might be involved[76-79,80L,81].

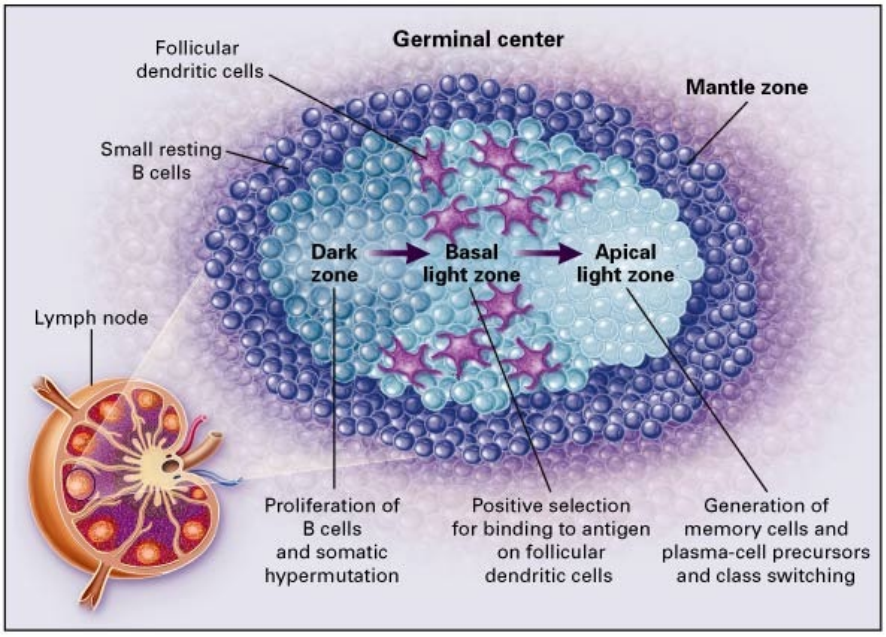


Figure 8 Scheme of GC. During humoral immune response, GCs are formed in peripheral lymphoid tissues. GCs consist of DZ and LZ. In DZ, GC B cells undergo rapid proliferation and localized mutation on *Ig* variable regions. In LZ, the affinities of the surface BCRs of GC B cells are selected, which is called affinity maturation. Eventually, GC B cells differentiate into plasma cells and memory B cells. This scheme is adapted from [82].

1.2.2 Dark zone (DZ) and light zone (LZ)

GCs can be divided into DZ and LZ[74]. GC B cells in the DZs are called centroblasts and proliferate at fast rate[74] within an interconnected network of reticular cells, similar to FDCs morphologically, that express CXC-chemokine ligand 12 (CXCL12)[83]. Centroblasts are large, proliferating cells without cellular surface Ig. These cells express high level of surface CXCR4 which is important for the positioning of centroblast in DZs [84]. GC B cells in the LZs are called centrocytes, are small in size, and express surface BCR with various affinities towards antigen[74].

The cell surface markers which distinguish DZ and LZ populations are under debate. RNA-seq analysis showed that DZ and LZ GC B populations exhibit differences in CD83 and CD86

expression[85]. Victora et al proposed that CD83 and CD86 together with CXCR4 could be used to determine DZ and LZ population, with GC B cells in DZ being $CXCR4^{hi}CD83^{lo}CD86^{lo}$, while those in LZ being $CXCR4^{lo}CD83^{hi}CD86^{hi}$ [85]. CD77 has been widely used to distinguish centroblasts and centrocytes. It is reported to be expressed on the cellular surface of fast proliferating GC B cells in DZ compartment, corresponding to centroblasts[86]. However, the functional differences between $CD77^{+}$ and $CD77^{-}$ population is poorly reproducible between labs[87-89]. In recent years, CXCR4 emerged as a functional cellular surface marker for centroblasts[84,90]. GC B cells express high and low level showed different transcriptional program which related to cell proliferation in DZ and activation in LZ[91]. Moreover, fast proliferating $CXCR4^{+}$ DZ cells fail to show a commensurate level of CD77 expression.

Under the classic model, the DZ is the site where B cell clonal expansion and SHM-mediated BCR diversification take place, while the LZ is the site where antigen affinities of BCRs are tested and selected. Centroblasts in DZs proliferate fast and undergo SHM which specifically introduces DNA mutations to the coding sequence of Ig variable regions, which leads to the generation of a huge repertoire of antibody affinities. Then, GC B cells migrate into LZs, with BCRs expressed on the cellular surface, and minimize cycling and become centrocytes. Centrocytes undergo affinity maturation by competing with each other for the binding to antigen-bound immune complexes presented on the cellular surface of FDCs or T cells and are selected to re-enter the DZ to further mutate their Ig genes or differentiate into PCs or Bmems[73,92].

However, the classic model has recently been under challenge. RNA-seq analysis of centroblasts and centrocytes showed that the transcriptome of GC B cells in DZs and LZs are very similar, and only minor differences could be detected[85]. Moreover, cellular division of GC B cells was found in both DZs and LZs, which is in striking contrast to the conventional concept – GC B cells proliferate fast in DZs and exit cell cycle once migrated to LZs[90]. Furthermore, apoptosis also takes place in both the DZ and LZ compartments. Besides, GC B cells from DZ and LZ are similar in terms of size, morphology, and migration speed[90,93]. These observations suggest that DZ and LZ B cells are more similar than previously thought. However, these findings are hardly commensurate with the functions of GC B cells in these two GC compartments, as centroblasts are BCR negative and undergo somatic hypermutation while centrocytes express BCR [94] and under affinity maturation [95].

Taken together, the controversial findings and models suggest that GCs are highly dynamic and complex. More studies are needed to better define the differences between centrocytes and centroblasts molecularly.

1.2.3 FDC network

FDCs form a reticular network within the LZ of GC structures[74,93,96-98]. They trap and present antigens on the cellular surface for a long period of time[99]. Antigens presented on the cell surface of FDCs are in the form of immunes complex (iccosomes)[100]. Moreover, FDCs secrete chemoattractants that are critical in positioning GC B cells into LZ. CXCL13, a CXCR5 ligand, is a chemoattractant produced by FDCs for the homing of $CXCR5^+$ GC B cells to the LZ compartment for affinity maturation and downstream differentiation [101-103], and also plays an important role in the polarization of the GC structure[84] together with $CXCR4^+$ centroblasts. Also, FDCs are responsible for the maintenance of the GC structure by producing cytokines and cell-cell adhesion molecules, such as IL-6[104,105], BAFF[106], ICAM-1 and VCAM-1[107]. Taken together, FDCs are critical for GC maintenance and affinity maturation.

1.2.4 Somatic hypermutation (SHM)

Somatic hypermutation (SHM), a process that introduces localized mutations in the Ig variable regions, takes place in DZs. It happens most frequently in variable regions exons and the proximal surroundings[108,109].

Activation-induced cytidine deaminase (AID), encoded by the *Aicda* gene, is mainly expressed in centroblasts[110,111]. It deaminates cytidine (C) residues to uracil (U) in the VDJ segments, which is the main reaction in SHM[112]. C to U mutations results in U:G lesions in the DNA. Although U:G lesions can directly introduce mutations after DNA replication, GC B cells tend to use low-fidelity DNA repair mechanisms to generate mutations[113]. The MRE11/RAD50/NBS1 (MRN) complex is involved in the DNA repair process in SHM. MRN can be recruited to the DNA in an AID-dependent manner. MRE11 specifically binds to rearranged V_H regions in mutating B cells, and MRE11/RAD50 cleaves single-strand DNA at abasic sites, which coincides with AID activity, and leaves SSBs on DNA. Then, low-fidelity DNA polymerase is

recruited to augment the mutation frequency[114]. MSH2/MSH6 (MMR) complex can be recruited to U:G pairs, and also leads to error-prone DNA synthesis to mutate DNA[115].

In SHM, the mutation rate in the V regions is ca. 10^{-3} mutations per base per cell cycle[116]. The mechanism underlying the specific recruitment of AID to the mutation regions is yet to be elucidated, nonetheless, the deamination process of AID takes place specifically on single-stranded ssDNA[117,118]. Therefore, AID activity seems to be coupled with transcription. Evidence showed that the distance from the promoter negatively correlates with SHM frequency[119]. Therefore, SHM associates closely with the transcriptional activity[120,121]. Although the distance to the promoter associates with SHM frequency, the mechanism about the specific targeting of AID in SHM remains unknown[122,123].

To be noted, SHM mediated by AID can also take place outside of Ig loci, which then leads to mutations in oncogenes, or even chromosomal translocations, therefore increasing the chance of B lymphomagenesis[124].

1.2.5 Class Switch recombination (CSR)

CSR is a process by which GC B cells exchange the constant region of *IgH* through a specific recombination event. CSR leads to the production of isotype-switched antibodies[125] (Figure 9), such as IgG, IgE and IgA, and a cessation of the expression of IgM[126]. The antigen affinities of resulting antibodies remain unchanged during CSR[127]. The usage of specific *IgH* regions is determined mainly by cytokines of the microenvironment and signals from T_{FH} cells[128,129]. These signals lead to transcription of the selected unrearranged switch (S) region and production of germline transcripts (GLTs). GLTs are produced immediately before CSR, and thus have been used as an indicator of CSR activity[130]. CSR in GC B cells is mediated by AID[131], however, a different domain of AID is involved than for SHM [132,133]. UNG and apurinic-apyrimidinic endonuclease 1 (APE1) is required for the DNA repair process for AID-mediated DNA mutation in S regions[101,134,135]. These DNA repair enzymes generate DNA breaks that result in the recombination between the variable region of VDJ segment and an alternative *IgH* constant region (C_H)[127].

Traditionally, CSR is thought to take place mainly in centrocytes. However, CSR has recently been proposed to initiate before GC commitment, while attenuating quickly in GC B cells prior to SHM[126]. The transcription of GLTs begins at 1.5 days post immunization and reaches its peak level at day 2.5, which coincides with the expression of *Aicda*. However, on day 3.5, B cells start to commit to the GC pathway with elevated level of BCL6[136]. Coincidentally, from day 3.5 onwards, GLTs level starts to decrease and are dramatically reduced before the establishment of GCs[126]. In addition, AP endonuclease 1, an enzyme engaged in DNA repair during CSR[137], is also reduced in GC B cells. Therefore, the window of CSR might only occur in the early days post immunization before GC commitment.

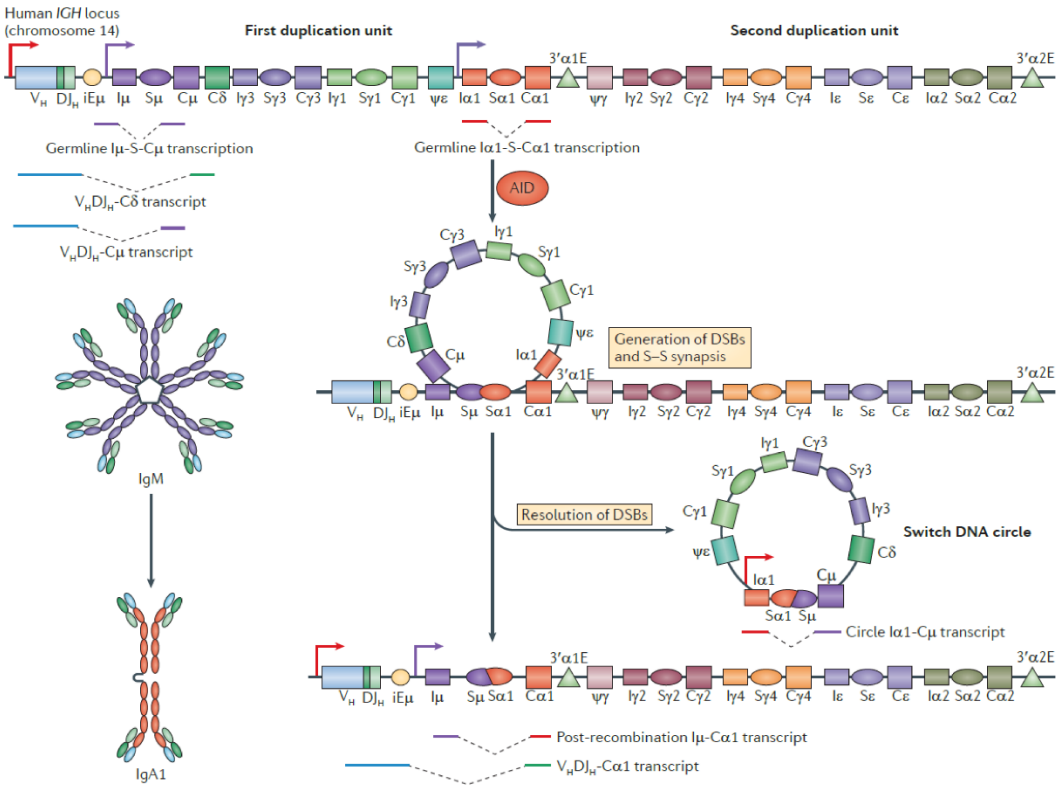


Figure 9 Scheme of IgM to IgA1 class switch recombination (CSR). This scheme illustrates the CSR between Sμ and Sα1 in the human *IGH* locus**.** The coding regions of *IGH* was switched with one of a set of downstream *IGH* constant regions (C_H). AID generates double-strand break in the switch (S) regions, upstream of C_H gene, and DNA repair process leads joining of newly arranged variable regions with the corresponding downstream constant regions [138]. Scheme taken from [138].

1.2.6 Affinity maturation

Antibody affinities in the serum after immunization increase strongly throughout the response; this phenomenon is designated affinity maturation[139].

The strength of BCR signaling and T_{FH} help play fundamental roles in promoting the increase of BCR affinity. BCRs of high-affinity GC B cells interact with antigens much more efficiently than the low-affinity counterparts do. Therefore, the majority of antigens in the microenvironment are predominantly bound by GC B cells expressing high-affinity BCRs, which results in lower affinity B cells not having access to antigens. Subsequently, the insufficient antigens binding fails to activate BCR signals in GC B cells of low-affinity antibodies, which ultimately leads to the removal of these cells by apoptosis. The survival signals from T_{FH} cells are critical for the survival of apoptosis-prone GC B cells. Evidence shows that T_{FH} preferentially forms synapses with B cells which present the highest level of antigen peptide-MHC (pMHC) molecules on the surface[140]. The affinity of BCR positively correlates with the antigen presentation ability of the cell. Therefore, the higher the BCR affinity, the more pMHC molecules are presented on the cell surface. In fact, the amount of pMHC becomes a readout of BCR affinity[141,142]. Thereby, in this way, B cells with high-affinity BCRs outcompete the low-affinity counterparts to interact with TFH cells and survive.

1.2.7 Transcriptional regulation of GC formation

The transcriptional regulation of the initiation, maintenance and downstream differentiation of GC has been reviewed by our lab[92].

Transcriptional control of germinal center (GC) and post-GC differentiation

(Published, 2018, Frontiers in Immunology)

frontiers in Immunology

REVIEW
published: 05 September 2018
doi: 10.3389/fimmu.2018.02026

The Transcriptional Regulation of Germinal Center Formation

Shuang Song[1,2] *and Patrick D. Matthias*[1,2*]

[1] *Friedrich Miescher Institute for Biomedical Research, Basel, Switzerland,* [2] *Faculty of Sciences, University of Basel, Basel, Switzerland*

Germinal centers (GCs) are essential structures of the humoral immune response, which form in the periphery in response to T cell dependent antigens. During the GC reaction, B cells undergo critical differentiation steps, which ultimately lead to the generation of antibodies with altered effector function and higher affinity for the selected antigen. Remarkably, many of the B cell tumors have their origin in the GCs; thus, understanding how the formation of these structures is regulated or deregulated is of high medical importance. This review gives an overview of the transcription factors that have been linked to the generation of GCs, and of their roles in the process.

Keywords: hematopoiesis, transcription factors, B cell development, germinal center (GCs), transcriptional regulation, germinal center development, germinal center maintenance, plasma cell and memory B cell differentiation

OPEN ACCESS

Edited by:
Rhodri Ceredig, National University of Ireland Galway, Ireland

Reviewed by:
Antony Basten, Garvan Institute of Medical Research, Australia
Kai-Michael Toellner, University of Birmingham, United Kingdom

***Correspondence:**
Patrick D. Matthias
patrick.matthias@fmi.ch

Specialty section:
This article was submitted to B Cell Biology, a section of the journal Frontiers in Immunology

Received: *30 June 2018*
Accepted: *16 August 2018*
Published: *05 September 2018*

Citation:
Song S and Matthias PD (2018) The Transcriptional Regulation of Germinal Center Formation. Front. Immunol. 9:2026. doi: 10.3389/fimmu.2018.02026

BACKGROUND TO B CELL DEVELOPMENT

B (and T) cells represent a unique model of cellular development, in which cells of multiple differentiation stages can be identified based on surface markers and readily isolated. Owing to these advantages, the lymphoid system has been used widely, beyond immunology, as a developmental paradigm in which the role of transcription factors (TFs) or signaling molecules can be tested experimentally. Several excellent reviews exist that describe in detail how B lymphocytes develop, what regulatory circuits are critical, or the details of GC development (1–11). We will therefore not discuss these aspects in detail, but will only give a high-level overview, and then focus this review on the transcriptional control of GCs formation.

B cells originate and develop in the bone marrow from hematopoietic stem cells (HSCs) that differentiate into progenitor stages of increasingly restricted potential. Once committed to the B lineage, B cell progenitors go through several successive stages, at which key events of their developmental fate take place. In particular, the PreB stage represents the phase during which immunoglobulin (Ig) genes, which code for the antibody molecules, rearrange their DNA segments in order to produce functional genes. The heavy chain rearranges first at the ProB stage, followed by the light chain at the small PreB-II stage. Immature B cells then express IgM at their surface and exit the bone marrow to enter the circulation and move to peripheral lymphoid organs such as the spleen or the lymph nodes. There, marginal zone (MZ) B cells play vital functions in T cell-independent humoral immune responses against blood-borne pathogens, follicular B cells can capture antigen presented by Follicular Dendritic Cells (FDCs) and present it to $CD4^+$ follicular T helper cells (T_{FH}) that are located around the B cell zone of the developing GC. This is the time during which critical signals, sent by the T_{FH} cells, induce isotype switching (so-called class switching, which exchanges IgM for IgG) and expansion of B cell clones starts. These B cells are called centroblasts and form the dark zone (DZ) of the GC. After several rounds of proliferation, somatic hypermutation begins, a process by which the Ig DNA becomes mutated under the

action of activation-induced cytidine deaminase (AID), leading to the generation of diverse clones expressing antibodies with different, potentially higher, affinity for antigen. From there, the B cells (centroblasts) move to the adjacent region called the light zone (LZ), where they express their antibody on the cell surface. GC B cells in the light zone are called centrocytes and are in a near apoptotic state. It is there that selection for the quality (affinity) of the antibody takes place: based on the affinity of the antibody for the antigen, the B cell can be eliminated or rescued and sent back to the dark zone as centroblast for an additional round of mutations, followed by renewed entry into the light zone and further antibody affinity testing. At some point in this dark zone–light zone selection cycle, the B cell expresses a high affinity antibody and can now exit the GC as a plasma cell that secretes high amounts of the antibody, or as a memory B cell that is ready to be reactivated upon future encounter with the antigen.

The rest of this review will put the emphasis on the transcriptional control of the formation and function of GCs, and highlight in particular TFs that are essential.

TRANSCRIPTION FACTORS REGULATING GC FORMATION

GC Initiation

Initiation of the GC reaction involves activation of the B cell receptor (BCR) by antigen engagement, followed by interaction of these B cells with antigen presenting cells and T_{FH} cells, which provide further activation signals (2, 3). **Figure 1** summarizes the molecular networks regulating initiation and function of the germinal centers reaction.

Transcription factors that are downstream of the BCR, such as the transcription coactivator OBF1 (a.k.a. OCA-B, or Bob1), a B cell-specific coactivator for the octamer transcription factors OCT1 and OCT2, are critical for GC formation (15–18). Mice deficient in *Pou2f2* (encoding OCT2), *Pou2af1* (encoding OBF1) or both showed complete lack of GCs (19). The underlying molecular mechanism is not clear yet, and the target genes of OBF1/OCT2 in the context of the germinal center reaction are largely unknown, although Spi-B which itself is required for GCs (20, 21) has been identified as a downstream target of OBF1 (22). Moreover, in $CD4^+$ T cells OBF1 and OCT1/OCT2 directly bind to the promoter region of *Bcl6* and activate its transcription, thereby promoting the development of T_{FH} cells (23). The putative role of these factors in regulating *Bcl6* expression in early GC B cells remains to be investigated.

BCL6 is a zinc finger TF that is essential for germinal center formation, as *Bcl6*-null mice completely lack GCs and affinity maturation (3, 24). During the early phase of the GC response, antigen stimulated B cells rely on T_{FH} cells for differentiation into GC B cells, and interaction between T_{FH} and B cells leads to the upregulation of BCL6 (25). Moreover, the upregulation of BCL6 leads to stabilized conjugation between B and T_{FH} cells, creating a positive feedback loop that enhances the GC formation program (3, 25). Failure in BCL6 upregulation prevents B cells from entering GC clusters and impairs the upregulation of CXCR4, a chemokine receptor expressed on germinal center DZ B cells that is critical for the maintenance of GC structural integrity (25).

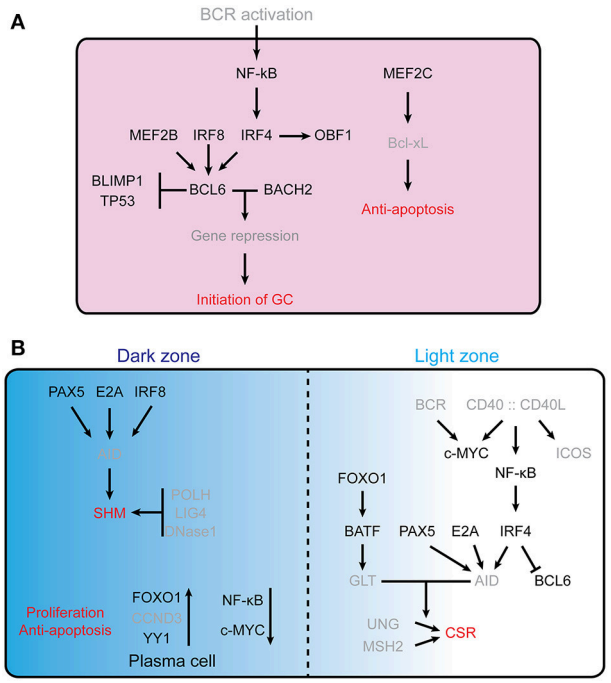


FIGURE 1 | Transcription factors controling GC formation. **(A)** Initiation of the GC reaction in follicular B cells. For clarity, TFs are indicated in black, while other molecules (e.g., receptors, cytokines, etc…) are in gray. B cell lymphoma 6 (BCL6) is essential for the initiation of germinal center, MEF2B, IRF8, IRF4, BLIMP1, and TP53 are involved in regulating the expression of *Bcl6*. BCL6 and Bach2 cooperatively (12) repress gene expression and thus allow the establishment of the germinal center B cell program. MEF2C is required for B cell survival post-antigen stimulation by upregulating the Bcl2l1. **(B)** Schematic representation of the dark zone and light zone of the GC. The different TFs involved are indicated, as well as some of the processes regulated (SHM, Proliferation, anti-apoptosis, CSR), see text for further details. In the GC DZ, AID is a key enzyme for SHM; its expression is controlled by PAX5, E2A, and IRF8. POLH, LIG4, and DNaseI are required for SHM and are highly expressed in DZ B cells. FOXO1 is a key factor for maintaining the GC DZ B cell program, CCND3 is preferentially expressed in GC DZ B cells and YY1 is required for GC DZ B cell proliferation and survival. NF-κB signaling and c-Myc are not essential for GC DZ B cells. In the GC LZ, CD40 signaling stimulated NF-κB further stimulates IRF4 expression, which suppresses *Bcl6* gene (13, 14). PAX5, E2A, and IRF4 are key factors in regulating AID level. BATF, a downstream target of FOXO1, regulates germline transcripts (GLTs) in centrocytes. GLT levels are highly correlated with accessibility of AID in CSR.

IRF4 is required at the early stage of GC formation. In transplantation experiments, $Irf4^{-/-}$ B cells fail to differentiate into GC B cells (26). Conditional knockout of *Irf4* by CD19cre which deletes from early B cells onwards leads to impaired GC formation (26). In contrast, once GCs have formed or initiated, IRF4 is no longer needed, as conditional knockout by Cγ1cre which deletes in already formed GC cells has minimal effects on GC differentiation (27). These results suggest that IRF4 is required for the very early phase upon T-cell-dependent antigen stimulation. Additional evidence supporting this idea is the rapid upregulation of IRF4 following BCR stimulation (28). Moreover, IRF4 is involved in modulating the expression of BCL6 and

OBF1, which both are key factors for GC initiation (3, 26). Taken together, IRF4 plays an important role in the early initiation phase of GC formation, possibly by regulating the induction of *Bcl6* and *Pou2af1*.

IRF8 was reported to upregulate BCL6 and AID levels in GC B cells (29, 30), and it was shown to promote GC B cells survival by regulating the expression level of MDM2 (31). However, deletion of IRF8 in B cells did not affect GC formation (32). Moreover, IRF8 is involved in the regulation of the BCL6-related transcriptional program in GC cells by directly interacting with BCOR (B cell lymphoma 6 corepressor) and BCL6. In transactivation assays, IRF8 augments the transcription repressive activity of BCL6 (33).

MEF2C is required for the proliferation and survival of B cells upon antigen receptor stimulation by upregulating the expression level of *Bcl2l1* (encoding the Bcl-xL protein) and several cell cycle related genes (34). Specific deletion of *Mef2c* in B cells leads to reduced proliferation and increased cell apoptosis upon anti-IgM stimulation. However, the responses are normal in the case of LPS, CD40, IL4, BAFF and RP105 stimulations. By histological examination, reduced number of GC follicules are observed in the spleens of *Mef2C*$^{fl/fl}$-CD19cre mice immunized with sheep red blood cells (SRBC) (34). MEF2B, another member of the MEF2 family, has been found to be mutated in ca. 11% of diffuse large B cell lymphoma (DLBCL), which are GC-derived tumors (35). MEF2B directly activates *Bcl6* transcription by binding to the regulatory region 1 kb upstream of the *Bcl6* gene transcription start site (35). Mutation of the MEF2B binding motif in the *Bcl6* gene promoter abrogates *Bcl6* transcription activity in cotransfection assays in 293T cells. Furthermore, knockdown of MEF2B protein by shRNAs leads to downregulation of BCL6 and upregulation of BCL6 target genes. These data suggest that MEF2B plays an important role in early GC formation by modulating *Bcl6* expression (35, 36).

BATF is a transcription factor of the AP-1 family, which is involved in GC structure establishment and class switch recombination. *Batf*$^{-/-}$ mice failed to develop normal GC structures when immunized with SRBC, as characterized by a lack of CD95 or GL7 positive B cells (37). *Batf*-null T_{FH} cells lack expression of the chemokine receptor CXCR5, which is essential for GC structure integrity. Additionally, the expression of *Bcl6* and *c-Maf*, both of which are important factors for T_{FH} cells development, is downregulated in absence of BATF (37).

c-MYC is another TF indispensable during the early phase of germinal center formation. Its expression is induced already 1–2 days after immunization (38) and it is required for GC maintenance, as conditional deletion of *c-Myc* by Cγ1cre leads to impaired GCs (39).

GC Development

The dark zone and the light zone of the GC are organized by the expression of the chemokine receptors CXCR4 and CXCR5, respectively (40). Thus, one can expect that TFs critical for CXCR4 and CXCR5 expression will be important for GCs.

GC Dark Zone

The germinal center DZ is characterized by an interconnected network of CXCL12 expressing reticular cells and compactly filled with rapidly proliferating centroblasts (41).

FOXO1 is highly expressed in human and mouse GC B cells, and its expression is largely specific to DZ B cells (with also some expression in naïve B cells) (42). Like in *Cxcr4*$^{-/-}$ mice, GCs from *Foxo1*$^{fl/fl}$-Cγ1cre mice completely lack a DZ structure, while the differentiation of plasma cells is normal (42, 43). *Foxo1*-null GCs lack proper structural polarization and show an even distribution of the FDC network (42). FOXO1, together with BCL6, represses the expression of B lymphocyte induced maturation protein 1 (BLIMP1), a key factor promoting differentiation of GC B cells into plasma cells, which is encoded by the *Prdm1* gene. By binding to the *Prdm1* promoter region, FOXO1 and BCL6 maintain the germinal center DZ program (42).

Bcl6-null GC precursor B cells fail to upregulate the expression of CXCR4 (25), which is a crucial chemokine receptor for GC DZ B cells. c-MYC is required throughout the early and late initiation phases of GC formation, but is not expressed in the proliferating DZ B cells (3), where it is repressed by BCL6 (38).

YY1 is required for GC B cell proliferation and GC development at least partly by modulating cell apoptosis (44). Deletion of *Yy1* specifically in GC B cells leads to a significant decrease in the number of DZ B cells, and elevated cell apoptosis (44).

Somatic hypermutation (SHM)

SHM generates a wide repertoire of affinities toward specific antigens, and mainly takes place in the DZ (45), although some extrafollicular SHM has been reported in transgenic mice deficient in the ability to establish GCs (46). AID, encoded by the *Aicda* gene, is the enzyme responsible for SHM and class switch recombination (47, 48). AID deaminates cytidines in DNA (49–54), followed by error-prone repair involving different DNA repair factors and ultimately leading to the introduction of somatic mutations (55). Thus, transcription factors which affect the expression of *Aicda* and DNA-damage tolerance related genes should be important for SHM. E proteins (56), PAX5 (57) and IRF8 (29) have been associated with positive regulation of *Aicda* transcription.

FOXO1 is involved in SHM by affecting the protein level of AID: *Foxo1*-null GC B cells show reduced level of AID enzyme, while mRNA level of *Aicda* is unchanged. Therefore, *Foxo1*-null GC B cells carry lower level of mutations in Ig locus than control cells (58).

Irf8 mRNA level peaks in centroblasts, and IRF8 regulates SHM by modulating the expression of *Aicda* and *Bcl6*: knockdown of IRF8 by siRNA leads to decreased transcription of *Aicda* and *Bcl6* (29). By ChIP, IRF8 binds to the promoter regions of *Aicda* and *Bcl6* in both human and mouse B cells. Furthermore, luciferase assays showed that IRF8 directly regulates the transcription of *Aicda* and *Bcl6* in HeLa cells cotransfected with an IRF8 expression vector and a reporter containing promoter regions of *Aicda* or *Bcl6* (29). Moreover,

IRF8 promotes GC B cells survival by regulating the expression level of MDM2 in the case of DNA damage (31).

Aicda expression is significantly reduced in activated B cells in which the helix-loop-helix factor ID3 is ectopically expressed. The possible mechanism is that ID3 inhibits the DNA-binding activity of E-proteins which activate the expression of *Aicda* (56, 59).

Light Zone

Three crucial B-cell developmental processes take place in the GC light zone: (i) selection of B cells that produce high-affinity antibodies, (ii) CSR, and (iii) initiation of centrocytes differentiation into plasma cells or memory B cells (9).

After rapid expansion in the DZ, B cells migrate to the LZ where those carrying high affinity B cell receptor genes are selected. The BCR pathway plays a fundamental role in this process: BCR signaling leads to phosphorylated AKT, and activated AKT further phosphorylates FOXO1 which then relocates from the nucleus to the cytoplasm (60). CD40 stimulation leads to NF-κB-mediated upregulation of IRF4 (13), which in turn represses *Bcl6* transcription (61). Together, these coordinated actions terminate the dark zone-associated transcriptional programme and allow establishment of the LZ transcriptome (13).

c-MYC is absent in most GC B cells, however, its expression is induced in high affinity BCR presenting GC B cells when receiving help from T_{FH} cells, in a process that requires both BCR and CD40 signaling (60). In addition to the requirement of c-MYC activity during the initial stage of GC formation, c-MYC is needed for GC maintenance in the late GC response (38). With the help from T_{FH} cells, c-MYC is transiently induced and upregulated in a small fraction of high affinity BCR expressing GC B cells within the LZ compartment. The Omomyc protein inhibits c-MYC function by antagonizing its DNA binding activity (62). Specific inhibition of c-MYC function by Doxycyclin-induced Omomyc expression in late GC B cells (10 days post-immunization by SRBC) leads to reduced GC size, indicating that c-MYC is required for GC maintenance once GCs are established (38).

Affinity maturation

FOXO1 is necessary for effective antibody affinity maturation: SHM frequency is comparable between WT and *Foxo1*-null GC B cells, but a severely decreased number of GC B cells harboring high affinity antibodies is observed in *Foxo1*-null GCs (42). Furthermore, *Foxo1*-null GC B cells have a lower level of cell surface BCR and are ineffective in activating T_{FH} cells in the LZ; this leads to lower stimulation of T_{FH} cells in the GC microenvironment and reduced production of IL-21, a cytokine that is vital for antibody affinity maturation (58). Thus, FOXO1 regulates antibody affinity maturation through both antigen presentation and T_{FH} cell activation (42, 58).

c-MYC is transiently induced in LZ B cells after receiving help from T_{FH} cells (38); selected GC B cells with induced c-MYC present high affinity BCR on the cell surface, and subsequently migrate into the DZ for the next iteration of proliferation and SHM (38, 39). However, in *Foxo1*-null GC B cells, the expression level of c-MYC is downregulated even under the help from T_{FH} cells. BATF, which controls the expression of *Aicda*, is another TF downregulated in the absence of *Foxo1* (58).

Class Switch Recombination

Like SHM, CSR also requires the expression of *Aicda* (47, 48). However, CSR depends on a different domain of the AID protein (63, 64). It is worthwhile mentioning that CSR already takes place before the GCs are formed, following B cell activation (65, 66). Much of the knowledge about CSR and the required factors originates from *in vitro* B cell activation experiments.

IRF4 was shown to regulate CSR in CD40 and IL4 stimulated B cells (27, 67). In the absence of IRF4, the *Aicda* expression level is decreased (67), and CSR is impaired (67). However, the expression of other genes important for CSR, such as Ung or Msh2 remains normal in $Irf4^{-/-}$ B cells. Therefore, the CSR defects in *Irf4*-null B cells seem to mainly reflect the impaired *Aicda* expression (67).

FOXO1 deficiency results in impaired class-switching: the compartment of *Igg1*-switched B cells in *Foxo1*-null GCs is heavily reduced, with accumulation of IgM^+ GC B cells. Yet, the expression level of *Aicda* is similar between WT and $Foxo1^{fl/fl}$-Cγ1cre GC B cells (42). In addition, *Foxo1*-null GC B cells display significantly lower expression of germline transcripts (GLTs) across the Ig locus. GLTs coincide with open chromatin and allow the exposure of switch regions to AID, which in turn induces single-strand DNA breaks through which class switch recombination is accomplished (68, 69). Thus, lower levels of GLTs correlate with reduced accessibility of AID toward the class switch regions. At the molecular level, FOXO1 possibly modulates GLT and post-switch transcripts by binding to I-mu, the 3′ IgH enhancer and a super-enhancer (70) in the Ig locus. Moreover, the transcription factor BATF, which is necessary for the expression of Ig GLTs and subsequent CSR (37), is downregulated upon FOXO1 depletion (42).

PAX5 binds to the promoter region of *Aicda* and activates its expression. Overexpression of PAX5 in a ProB cell line induces the expression *Aicda*, while ID2 has an antagonizing effect on this induction. Moreover, PAX5, E2A, and AID directly interact with each other and form a complex, which contributes to directing AID to the *Igh* locus for CSR (71). In addition, ID2 and ID3 negatively regulate CSR by repressing *Aicda* expression (56, 57).

BATF directly controls the expression of *Aicda*. By ChIP-seq and EMSA, BATF was shown to bind to the regulatory region of *Aicda*. In line with this, the expression level of *Aicda* is downregulated in $Batf^{-/-}$ mice (37). Consequently, production of isotype switched antibodies is almost completely missing, although IgM production upon T-cell-dependent or -independent antigen stimulation is still normal in $Batf^{-/-}$ mice (37). Moreover, $Batf^{-/-}$ mice display a reduction in GLTs from different isotypes, except those from μ-chain. Germline transcription initiated by switch region (I) region promoters, located upstream of the different constant heavy chain exons, is required for AID targeting and successful CSR (8, 54). Taken together, BATF regulates CSR by modulating the expression of *Aicda* and GLTs from the Ig locus.

TFs Controlling the Migration of Cells Between Dark and Light Zone

The LZ-to-DZ transition is mainly driven by high affinity antibody presentation on the surface of GC B cells and the subsequent help from T_{FH} cells toward high affinity antibody LZ cells (58, 72–74).

Foxo1-null GC B cells showed reduced level of surface BCR and Igβ when compared to WT cells (58). Therefore, reduced antigen presentation on *Foxo1*$^{-/-}$ GC LZ B cells fails to effectively activate T_{FH} cells, resulting in lower number of T_{FH} cells and decreased production of IL-21, an important cytokine for GC B cell differentiation and affinity maturation promoted by T_{FH} cells (75, 76). Moreover, *Foxo1*-null GC LZ B cells express less IL21R, further abrogating the ability to receive help from T_{FH} cells (BCR and CD40 signal). Therefore, the LZ-to-DZ migration of selected cells is impaired in the absence of *Foxo1* (58) and *Foxo1*-null GC B cells are trapped in the LZ compartment. Furthermore, *Foxo1*-null GC B cells showed lower proliferation rate in the LZ compartment, in spite of harboring a high affinity BCR, indicating defects in cyclic reentry mediated by help from T_{FH} cells which is coupled with the LZ-to-DZ migration (58).

Interaction between high affinity BCR expressing LZ B cells and T_{FH} cells leads to activation of c-MYC expression (60), which promotes cyclic reentry and LZ-to-DZ migration (58).

Differentiation: Memory/Plasma Cell Fate Decision

Figure 2 summarizes the molecular networks involved in plasma cell and memory B cell differentiation.

Transient and low expression of *Irf4* leads to the expression of *Bcl6* and *Pou2af1* during the early phase of GC formation, while sustained and high level expression of *Irf4* is required for plasma cell differentiation (26). *Irf4* expression is induced to a high level in the fraction of LZ B cells which present high affinity antibodies (26, 28). BLIMP1 is a transcriptional repressor that is essential for plasma cell development: *Prdm1*-deficient mice cannot produce plasma cells and overexpression of BLIMP1 is sufficient to induce plasmablast differentiation (77, 78). *Prdm1* is a downstream target of IRF4, and BLIMP1 can further increase the level of IRF4, thus reinforcing the plasma cell differentiation program in a feed forward loop (26, 79). Morever, BLIMP1 represses the expression of *Aicda* (79) and *Abf-1* (80).

BLIMP1, together with IRF4, acts upstream of X-box binding protein 1 (XBP1), a transcription factor that is essential for upregulation of the secretory apparatus required for antibody production in plasma cells (27, 81, 82). PAX5 represses XBP1 and thus prevents plasma cell differentiation (83). Conversely, downregulation of *Pax5* by BLIMP1 is a necessary step for induction of *Xbp-1* and activation of the plasma cell program (81, 84). However, in the absence of IRF4, BLIMP1 alone is not sufficient for plasma cell differentiation. Moreover, BLIMP1 is required for post-transcriptional regulation of XBP1 mRNA, by which the active form XBP1 protein is generated (85). XBP1 directly regulates *Pou2af1* expression in plasma cells (86), which might be important for IgG production, since OBF1 and OCT2 are required for normal IgG expression (87). *Irf4*-deficient

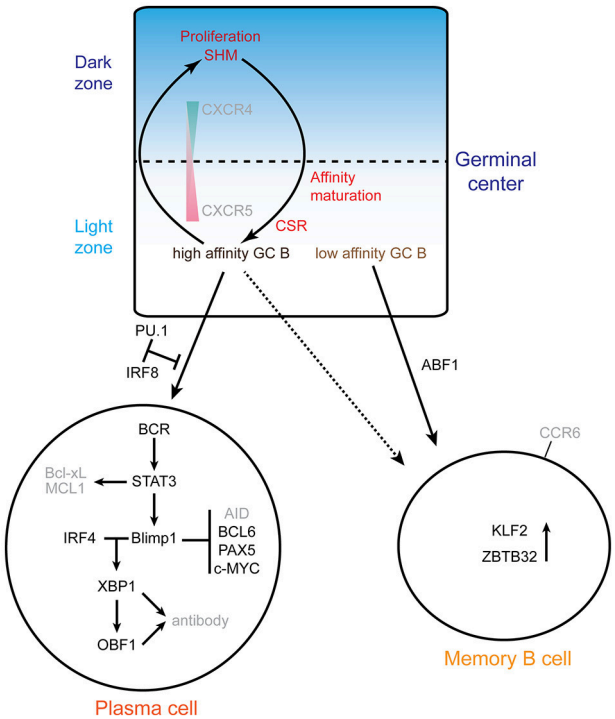


FIGURE 2 | Exit from the GC: memory vs. plasma cells. The top scheme highlights the cycling of B cells between the dark zone and the light zone. The gradients of the chemokine receptors CXCR4 and 5 are indicated, as well as the processes involved. GC B cells proliferate and undergo somatic hypermutation (SHM) in the GC DZ, and then migrate to the light zone where the affinity of mutated BCRs is selected. B cells with high affinity BCR are selected to reenter the DZ for several rounds of SHM or can differentiate into plasma cells. Memory B cells are mainly originated from B cells with low affinity BCR. The alternative fates of B cells exiting the GC are depicted at the bottom: plasma cells or memory B cells. The main networks that have been identified are indicated. In the plasma cell differentiation pathway, PU.1 and IRF8 negatively regulate plasma cell differentiation. IRF4 and BLIMP1 form the central axis in establishing plasma cells. BLIMP1 further suppresses the expression of *Aicda*, *Bcl6*, *Pax5* and *c-Myc* genes and finally terminates the GC program. BLIMP1 positively regulates *Xbp1* expression, makes the cells ready for antibody production and secretion. ABF1 promotes memory B cell differentiation and inhibits plasma cell differentiation. ZBTB32 and KLF2 are highly expressed in memory B cells.

(Cγ1cre) mice completely lack CD138^{+} (aka Syndecan1^{hi}) plasma cells in spleen, peripheral blood and bone marrow (27). In addition to the important yet mechanistically unclear role of OBF1 in the early development of GC formation, this factor is required for antibody production (both unswitched and switched isotypes) in T-dependent antigen stimulation and normal antibody secreting cell differentiation, as the number of antibody-secreting Syndecan1^{hi} cells is dramatically reduced in absence of OBF1 (88). Moreover, OBF1 is required for the induction of *Prdm1* (88). Taken together, IRF4 and BLIMP1 function together to drive the transition from GC B cells to plasma cells by repressing the GC program and enhancing the plasma cell differentiation program.

In contrast, IRF8 together with PU.1 inhibit the GC B cell differentiation toward plasma cells, suggesting that the balance

between IRF4 and IRF8 may be critical for the fate of B cells at this developmental transition (89). In addition, STAT3 regulates the differentiation of plasma cells possibly by promoting cell survival through activating the expression of pro-survival genes such as *Bcl2l1* and *Mcl1* (90, 91).

It is not resolved yet which transcription factor(s) play a major role in regulating the differentiation of memory B cells (92); however, recent evidence suggests that memory B cells originate from the low affinity compartment of the LZ (92, 93). Comparison of the transcriptome profiles from high- and low-affinity BCR expressing GC B cells in the LZ compartment showed that genes involved in DZ maintenance and cyclic reentry are downregulated in low-affinity fractions (with switched isotype), where *Bach2* and *Pax5* are upregulated (93). As indicated above, upregulation of these transcription factors blocks plasma cell differentiation (94–96).

BACH2 regulates LZ B cells to commit to the memory B cell differentiation path in a dosage dependent manner, as complete knockout or haploinsufficiency of *Bach2* lead to lower memory B cell differentiation (93). Furthermore, T_{FH} cell interaction and affinity maturation in LZ compartment are negatively correlated with *Bach2* expression, thus confirming that memory B cells are generated from low affinity fraction in the LZ compartment (92, 93).

Activated B cell Factor 1 (ABF-1), a helix-loop-helix TF predominantly expressed in memory B cells, blocks the plasma cell differentiation program (80). An inducible ABF-1-ER mouse model demonstrated that induction of ABF-1 promotes GC formation and memory B cell differentiation (80, 97). ZBTB32 and KLF2 are two factors expressed in memory B cells and which inhibit the GC response (98). They have been associated with memory B cells (99), but further mechanistic studies are needed to understand their specific role. BCL6 is a known inhibitor of plasma cell differentiation which directly represses the expression of *Prdm1* (100, 101). STAT5 directly modulates the expression level of *Bcl6*, which in turn directly represses *Prdm1* expression, and thus promotes memory B cell differentiation.

OUTLOOK AND OPEN QUESTIONS

A central theme is that often multiple distinct TFs act in concert to promote, or repress, specific steps of GC development. Some factors, such as BCL6 or BLIMP1, are considered to be master regulators for GC or plasma cell development, respectively. Yet, many other factors have been shown to be important (FOXO1, IRF4) or sometimes essential (OBF1, OCT2). In most cases, however, additional mechanistic studies are required to precisely understand how these factors fit in the overall regulatory circuitry. Moreover, it is often not clear how the TFs perform their function in this biological paradigm: which co-activators or co-repressors are involved and what epigenetic regulators are required?

Finally, intensive investigations have been conducted to understand the interaction between high affinity BCR expressing GC B cells and T_{FH} cells, yet little is known about how GC B cells are determined by transcriptional regulators to proceed with the LZ-to-DZ migration, or commit to the plasma cell differentiation cascade.

AUTHOR CONTRIBUTIONS

SS wrote the manuscript and PM corrected the manuscript.

ACKNOWLEDGMENTS

This work was supported by the Novartis Research Foundation. We thank all lab members and colleagues for useful discussions. This review is dedicated to our friend and colleague Antonius (Ton) Rolink, with whom we have had a very long-standing collaboration; it was always fun.

REFERENCES

1. Matthias P, Rolink AG. Transcriptional networks in developing and mature B cells. *Nat Rev Immunol.* (2005) 5:497–508. doi: 10.1038/nri1633
2. Victora GD, Nussenzweig MC. Germinal centers. *Annu Rev Immunol.* (2012) 30:429–57. doi: 10.1146/annurev-immunol-020711-075032
3. De Silva NS, Klein U. Dynamics of B cells in germinal centres. *Nat Rev Immunol.* (2015) 15:137–48. doi: 10.1038/nri3804
4. Mesin L, Ersching J, Victora GD. Germinal center B Cell Dynamics. *Immunity* (2016) 45:471–82. doi: 10.1016/j.immuni.2016.09.001
5. Shapiro-Shelef M, Calame K. Regulation of plasma-cell development. *Nat Rev Immunol.* (2005) 5:230–42. doi: 10.1038/nri1572
6. Wagner SD, Neuberger MS. Somatic hypermutation of immunoglobulin genes. *Annu Rev Immunol.* (1996) 14:441–57. doi: 10.1146/annurev.immunol.14.1.441
7. Allen CD, Okada T, Cyster JG. Germinal-center organization and cellular dynamics. *Immunity* (2007) 27:190–202. doi: 10.1016/j.immuni.2007.07.009
8. Stavnezer J, Guikema JE, Schrader CE. Mechanism and regulation of class switch recombination. *Annu Rev Immunol.* (2008) 26:261–92. doi: 10.1146/annurev.immunol.26.021607.090248
9. Klein U, Dalla-Favera R. Germinal centres: role in B-cell physiology and malignancy. *Nat Rev Immunol.* (2008) 8:22–33. doi: 10.1038/nri2217
10. Pillai S, Cariappa A. The follicular versus marginal zone B lymphocyte cell fate decision. *Nat Rev Immunol.* (2009) 9:767–77. doi: 10.1038/nri2656
11. Hardy RR, Hayakawa K. B cell development pathways. *Annu Rev Immunol.* (2001) 19:595–621. doi: 10.1146/annurev.immunol.19.1.595
12. Huang C, Geng H, Boss I, Wang L, Melnick A. Cooperative transcriptional repression by BCL6 and BACH2 in germinal center B-cell differentiation. *Blood* (2014) 123:1012–20. doi: 10.1182/blood-2013-07-518605
13. Saito M, Gao J, Basso K, Kitagawa Y, Smith PM, Bhagat G, et al. A signaling pathway mediating downregulation of BCL6 in germinal center B cells is blocked by BCL6 gene alterations in B cell lymphoma. *Cancer Cell.* (2007) 12:280–92. doi: 10.1016/j.ccr.2007.08.011
14. Heise N, De Silva NS, Silva K, Carette A, Simonetti G, Pasparakis M, et al. Germinal center B cell maintenance and differentiation are controlled by distinct NF-kappaB transcription factor subunits. *J Exp Med.* (2014) 211:2103–18. doi: 10.1084/jem.20132613
15. Kim U, Qin XF, Gong S, Stevens S, Luo Y, Nussenzweig MC, et al. The B-cell-specific transcription coactivator OCA-B-OBF-1-Bob-1 is essential for normal production of immunoglobulin isotypes. *Nature* (1996) 383:542–7.
16. Schubart DB, Rolink A, Kosco-Vilbois MH, Botteri F, Matthias P. B-cell-specific coactivator OBF-1/OCA-B/Bob1 required for immune response and germinal centre formation. *Nature* (1996) 383:538–42. doi: 10.1038/383538a0

17. Luo Y, Roeder RG. Cloning, functional characterization, and mechanism of action of the B-cell-specific transcriptional coactivator OCA-B. *Mol Cell Biol.* (1995) 15:4115–24. doi: 10.1128/MCB.15.8.4115
18. Strubin M, Newell JW, Matthias P. OBF-1, a novel B cell-specific coactivator that stimulates immunoglobulin promoter activity through association with octamer-binding proteins. *Cell* (1995) 80:497–506. doi: 10.1016/0092-8674(95)90500-6
19. Schubart K, Massa S, Schubart D, Corcoran LM, Rolink AG, Matthias P. B cell development and immunoglobulin gene transcription in the absence of Oct-2 and OBF-1. *Nat Immunol.* (2001) 2:69–74. doi: 10.1038/83190
20. Garrett-Sinha LA, Su GH, Rao S, Kabak S, Hao Z, Clark MR, et al. PU.1 and Spi-B are required for normal B cell receptor-mediated signal transduction. *Immunity* (1999) 10:399–408. doi: 10.1016/S1074-7613(00)80040-0
21. Su GH, Chen HM, Muthusamy N, Garrett-Sinha LA, Baunoch D, Tenen DG, et al. Defective B cell receptor-mediated responses in mice lacking the Ets protein, Spi-B. *EMBO J.* (1997) 16:7118–29. doi: 10.1093/emboj/16.23.7118
22. Bartholdy B, Du Roure C, Bordon A, Emslie D, Corcoran LM, Matthias P. The Ets factor Spi-B is a direct critical target of the coactivator OBF-1. *Proc Natl Acad Sci USA.* (2006) 103:11665–70. doi: 10.1073/pnas.0509430103
23. Stauss D, Brunner C, Berberich-Siebelt F, Hopken UE, Lipp M, Muller G. The transcriptional coactivator Bob1 promotes the development of follicular T helper cells via Bcl6. *EMBO J.* (2016) 35:881–98. doi: 10.15252/embj.201591459
24. Crotty S, Johnston RJ, Schoenberger SP. Effectors and memories: Bcl-6 and Blimp-1 in T and B lymphocyte differentiation. *Nat Immunol.* (2010) 11:114–20. doi: 10.1038/ni.1837
25. Kitano M, Moriyama S, Ando Y, Hikida M, Mori Y, Kurosaki T, et al. Bcl6 protein expression shapes pre-germinal center B cell dynamics and follicular helper T cell heterogeneity. *Immunity* (2011) 34:961–72. doi: 10.1016/j.immuni.2011.03.025
26. Ochiai K, Maienschein-Cline M, Simonetti G, Chen J, Rosenthal R, Brink R, et al. Transcriptional regulation of germinal center B and plasma cell fates by dynamical control of IRF4. *Immunity* (2013) 38:918–29. doi: 10.1016/j.immuni.2013.04.009
27. Klein U, Casola S, Cattoretti G, Shen Q, Lia M, Mo T, et al. Transcription factor IRF4 controls plasma cell differentiation and class-switch recombination. *Nat Immunol.* (2006) 7:773–82. doi: 10.1038/ni1357
28. Sciammas R, Li Y, Warmflash A, Song Y, Dinner AR, Singh H. An incoherent regulatory network architecture that orchestrates B cell diversification in response to antigen signaling. *Mol Syst Biol.* (2011) 7:495. doi: 10.1038/msb.2011.25
29. Lee CH, Melchers M, Wang H, Torrey TA, Slota R, Qi CF, et al. Regulation of the germinal center gene program by interferon (IFN) regulatory factor 8/IFN consensus sequence-binding protein. *J Exp Med.* (2006) 203:63–72. doi: 10.1084/jem.20051450
30. Qi CF, Li ZY, Raffeld M, Wang HS, Kovalchuk AL, Morse HC. Differential expression of IRF8 in subsets of macrophages and dendritic cells and effects of IRF8 deficiency on splenic B cell and macrophage compartments. *Immunol Res.* (2009) 45:62–74. doi: 10.1007/s12026-008-8032-2
31. Zhou JX, Lee CH, Qi CF, Wang H, Naghashfar Z, Abbasi S, et al. IFN regulatory factor 8 regulates MDM2 in germinal center B cells. *J Immunol.* (2009) 183:3188–94. doi: 10.4049/jimmunol.0803693
32. Feng J, Wang H, Shin DM, Masiuk M, Qi CF, Morse HC III. IFN regulatory factor 8 restricts the size of the marginal zone and follicular B cell pools. *J Immunol.* (2011) 186:1458–66. doi: 10.4049/jimmunol.1001950
33. Yoon J, Feng X, Kim YS, Shin DM, Hatzi K, Wang H, et al. Interferon regulatory factor 8 (IRF8) interacts with the B cell lymphoma 6 (BCL6) corepressor BCOR. *J Biol Chem.* (2014) 289:34250–7. doi: 10.1074/jbc.M114.571182
34. Wilker PR, Kohyama M, Sandau MM, Albring JC, Nakagawa O, Schwarz JJ, et al. Transcription factor Mef2c is required for B cell proliferation and survival after antigen receptor stimulation. *Nat Immunol.* (2008) 9:603–12. doi: 10.1038/ni.1609
35. Ying CY, Dominguez-Sola D, Fabi M, Lorenz IC, Hussein S, Bansal M, et al. MEF2B mutations lead to deregulated expression of the oncogene BCL6 in diffuse large B cell lymphoma. *Nat Immunol.* (2013) 14:1084–92. doi: 10.1038/ni.2688
36. Pon JR, Marra MA. MEF2 transcription factors: developmental regulators and emerging cancer genes. *Oncotarget* (2016) 7:2297–312. doi: 10.18632/oncotarget.6223
37. Ise W, Kohyama M, Schraml BU, Zhang T, Schwer B, Basu U, et al. The transcription factor BATF controls the global regulators of class-switch recombination in both B cells and T cells. *Nat Immunol.* (2011) 12:536–43. doi: 10.1038/ni.2037
38. Dominguez-Sola D, Victora GD, Ying CY, Phan RT, Saito M, Nussenzweig MC, et al. The proto-oncogene MYC is required for selection in the germinal center and cyclic reentry. *Nat Immunol.* (2012) 13:1083–91. doi: 10.1038/ni.2428
39. Calado DP, Sasaki Y, Godinho SA, Pellerin A, Kochert K, Sleckman BP, et al. The cell-cycle regulator c-Myc is essential for the formation and maintenance of germinal centers. *Nat Immunol.* (2012) 13:1092–100. doi: 10.1038/ni.2418
40. Allen CD, Ansel KM, Low C, Lesley R, Tamamura H, Fujii N, et al. Germinal center dark and light zone organization is mediated by CXCR4 and CXCR5. *Nat Immunol.* (2004) 5:943–52. doi: 10.1038/ni1100
41. Bannard O, Horton RM, Allen CD, An J, Nagasawa T, Cyster JG. Germinal center centroblasts transition to a centrocyte phenotype according to a timed program and depend on the dark zone for effective selection. *Immunity* (2013) 39:912–24. doi: 10.1016/j.immuni.2013.08.038
42. Dominguez-Sola D, Kung J, Holmes AB, Wells VA, Mo T, Basso K, et al. The FOXO1 transcription factor instructs the germinal center dark zone program. *Immunity* (2015) 43:1064–74. doi: 10.1016/j.immuni.2015.10.015
43. Sander S, Chu VT, Yasuda T, Franklin A, Graf R, Calado DP, et al. PI3 Kinase and FOXO1 transcription factor activity differentially control b cells in the germinal center light and dark zones. *Immunity* (2015) 43:1075–86. doi: 10.1016/j.immuni.2015.10.021
44. Trabucco SE, Gerstein RM, Zhang H. YY1 Regulates the germinal center reaction by inhibiting apoptosis. *J Immunol.* (2016) 197:1699–707. doi: 10.4049/jimmunol.1600721
45. Oropallo MA, Cerutti A. Germinal center reaction: antigen affinity and presentation explain it all. *Trends Immunol.* (2014) 35:287–9. doi: 10.1016/j.it.2014.06.001
46. Toyama H, Okada S, Hatano M, Takahashi Y, Takeda N, Ichii H, et al. Memory B cells without somatic hypermutation are generated from Bcl6-deficient B cells. *Immunity* (2002) 17:329–39. doi: 10.1016/S1074-7613(02)00387-4
47. Muramatsu M, Kinoshita K, Fagarasan S, Yamada S, Shinkai Y, Honjo T. Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. *Cell* (2000) 102:553–63. doi: 10.1016/S0092-8674(00)00078-7
48. Revy P, Muto T, Levy Y, Geissmann F, Plebani A, Sanal O, et al. Activation-induced cytidine deaminase (AID) deficiency causes the autosomal recessive form of the Hyper-IgM syndrome (HIGM2). *Cell* (2000) 102:565–75. doi: 10.1016/S0092-8674(00)00079-9
49. Petersen-Mahrt SK, Harris RS, Neuberger MS. AID mutates E. coli suggesting a DNA deamination mechanism for antibody diversification. *Nature* (2002) 418:99–103. doi: 10.1038/nature00862
50. Bransteitter R, Pham P, Scharff MD, Goodman MF. Activation-induced cytidine deaminase deaminates deoxycytidine on single-stranded DNA but requires the action of RNase. *Proc Natl Acad Sci USA.* (2003) 100:4102–7. doi: 10.1073/pnas.0730835100
51. Chaudhuri J, Tian M, Khuong C, Chua K, Pinaud E, Alt FW. Transcription-targeted DNA deamination by the AID antibody diversification enzyme. *Nature* (2003) 422:726–30. doi: 10.1038/nature01574
52. Di Noia J, Neuberger MS. Altering the pathway of immunoglobulin hypermutation by inhibiting uracil-DNA glycosylase. *Nature* (2002) 419:43–8. doi: 10.1038/nature00981
53. Dickerson SK, Market E, Besmer E, Papavasiliou FN. AID mediates hypermutation by deaminating single stranded DNA. *J Exp Med.* (2003) 197:1291–6. doi: 10.1084/jem.20030481
54. Ramiro AR, Stavropoulos P, Jankovic M, Nussenzweig MC. Transcription enhances AID-mediated cytidine deamination by exposing single-stranded DNA on the nontemplate strand. *Nat Immunol.* (2003) 4:452–6. doi: 10.1038/ni920

55. Chahwan R, Edelmann W, Scharff MD, Roa S. AIDing antibody diversity by error-prone mismatch repair. *Semin Immunol*. (2012) 24:293–300. doi: 10.1016/j.smim.2012.05.005
56. Sayegh CE, Quong MW, Agata Y, Murre C. E-proteins directly regulate expression of activation-induced deaminase in mature B cells. *Nat Immunol*. (2003) 4:586–93. doi: 10.1038/ni923
57. Gonda H, Sugai M, Nambu Y, Katakai T, Agata Y, Mori KJ, et al. The balance between Pax5 and Id2 activities is the key to AID gene expression. *J Exp Med*. (2003) 198:1427–37. doi: 10.1084/jem.20030802
58. Inoue T, Shinnakasu R, Ise W, Kawai C, Egawa T, Kurosaki T. The transcription factor Foxo1 controls germinal center B cell proliferation in response to T cell help. *J Exp Med*. (2017) 214:1181–98. doi: 10.1084/jem.20161263
59. Quong MW, Harris DP, Swain SL, Murre C. E2A activity is induced during B-cell activation to promote immunoglobulin class switch recombination. *EMBO J*. (1999) 18:6307–18. doi: 10.1093/emboj/18.22.6307
60. Luo W, Weisel F, Shlomchik MJ. B cell receptor and CD40 signaling are rewired for synergistic induction of the c-Myc transcription factor in germinal center B cells. *Immunity* (2018) 48:313–26 e5. doi: 10.1016/j.immuni.2018.01.008
61. Allman D, Jain A, Dent A, Maile RR, Selvaggi T, Kehry MR, et al. BCL-6 expression during B-cell activation. *Blood* (1996) 87:5257–68.
62. Soucek L, Jucker R, Panacchia L, Ricordy R, Tato F, Nasi S. Omomyc, a potential Myc dominant negative, enhances Myc-induced apoptosis. *Cancer Res*. (2002) 62:3507–10.
63. Barreto V, Reina-San-Martin B, Ramiro AR, McBride KM, Nussenzweig MC. C-terminal deletion of AID uncouples class switch recombination from somatic hypermutation and gene conversion. *Molecular Cell* (2003) 12:501–8. doi: 10.1016/S1097-2765(03)00309-5
64. Ta VT, Nagaoka H, Catalan N, Durandy A, Fischer A, Imai K, et al. AID mutant analyses indicate requirement for class-switch-specific cofactors. *Nat Immunol*. (2003) 4:843–8. doi: 10.1038/ni964
65. Chan TD, Gatto D, Wood K, Camidge T, Basten A, Brink R. Antigen affinity controls rapid T-dependent antibody production by driving the expansion rather than the differentiation or extrafollicular migration of early plasmablasts. *J Immunol*. (2009) 183:3139–49. doi: 10.4049/jimmunol.0901690
66. Zhang Y, Garcia-Ibanez L, Toellner KM. Regulation of germinal center B-cell differentiation. *Immunol Rev*. (2016) 270:8–19. doi: 10.1111/imr.12396
67. Sciammas R, Shaffer AL, Schatz JH, Zhao H, Staudt LM, Singh H. Graded expression of interferon regulatory factor-4 coordinates isotype switching with plasma cell differentiation. *Immunity* (2006) 25:225–36. doi: 10.1016/j.immuni.2006.07.009
68. Chaudhuri J, Basu U, Zarrin A, Yan C, Franco S, Perlot T, et al. Evolution of the immunoglobulin heavy chain class switch recombination mechanism. *Adv Immunol*. (2007) 157–214. doi: 10.1016/S0065-2776(06)94006-1
69. Pavri R, Nussenzweig MC. AID targeting in antibody diversity. *Adv Immunol*. (2011) 110:1–26. doi: 10.1016/B978-0-12-387663-8.00005-3
70. Qian J, Wang Q, Dose M, Pruett N, Kieffer-Kwon KR, Resch W, et al. B cell super-enhancers and regulatory clusters recruit AID tumorigenic activity. *Cell* (2014) 159:1524–37. doi: 10.1016/j.cell.2014.11.013
71. Hauser J, Grundstrom C, Kumar R, Grundstrom T. Regulated localization of an AID complex with E2A, PAX5 and IRF4 at the Igh locus. *Mol Immunol*. (2016) 80:78–90. doi: 10.1016/j.molimm.2016.10.014
72. Allen CD, Okada T, Tang HL, Cyster JG. Imaging of germinal center selection events during affinity maturation. *Science* (2007) 315:528–31. doi: 10.1126/science.1136736
73. Victora GD, Schwickert TA, Fooksman DR, Kamphorst AO, Meyer-Hermann M, Dustin ML, et al. Germinal center dynamics revealed by multiphoton microscopy with a photoactivatable fluorescent reporter. *Cell* (2010) 143:592–605. doi: 10.1016/j.cell.2010.10.032
74. Liu D, Xu H, Shih C, Wan Z, Ma X, Ma W, et al. T-B-cell entanglement and ICOSL-driven feed-forward regulation of germinal centre reaction. *Nature* (2015) 517:214–8. doi: 10.1038/nature13803
75. Linterman MA, Beaton L, Yu D, Ramiscal RR, Srivastava M, Hogan JJ, et al. IL-21 acts directly on B cells to regulate Bcl-6 expression and germinal center responses. *J Exp Med*. (2010) 207:353–63. doi: 10.1084/jem.20091738
76. Zotos D, Coquet JM, Zhang Y, Light A, D'Costa K, Kallies A, et al. IL-21 regulates germinal center B cell differentiation and proliferation through a B cell-intrinsic mechanism. *J Exp Med*. (2010) 207:365–78. doi: 10.1084/jem.20091777
77. Turner CA Jr, Mack DH, Davis MM. Blimp-1, a novel zinc finger-containing protein that can drive the maturation of B lymphocytes into immunoglobulin-secreting cells. *Cell* (1994) 77:297–306. doi: 10.1016/0092-8674(94)90321-2
78. Shapiro-Shelef M, Lin KI, McHeyzer-Williams LJ, Liao J, McHeyzer-Williams MG, Calame K. Blimp-1 is required for the formation of immunoglobulin secreting plasma cells and pre-plasma memory B cells. *Immunity* (2003) 19:607–20. doi: 10.1016/S1074-7613(03)00267-X
79. Minnich M, Tagoh H, Bonelt P, Axelsson E, Fischer M, Cebolla B, et al. Multifunctional role of the transcription factor Blimp-1 in coordinating plasma cell differentiation. *Nat Immunol*. (2016) 17:331–43. doi: 10.1038/ni.3349
80. Chiu YK, Lin IY, Su ST, Wang KH, Yang SY, Tsai DY, et al. Transcription factor ABF-1 suppresses plasma cell differentiation but facilitates memory B cell formation. *J Immunol*. (2014) 193:2207–17. doi: 10.4049/jimmunol.1400411
81. Reimold AM, Iwakoshi NN, Manis J, Vallabhajosyula P, Szomolanyi-Tsuda E, Gravallese EM, et al. Plasma cell differentiation requires the transcription factor XBP-1. *Nature* (2001) 412:300–7. doi: 10.1038/35085509
82. Shaffer AL, Shapiro-Shelef M, Iwakoshi NN, Lee AH, Qian SB, Zhao H, et al. XBP1, downstream of Blimp-1, expands the secretory apparatus and other organelles, and increases protein synthesis in plasma cell differentiation. *Immunity* (2004) 21:81–93. doi: 10.1016/j.immuni.2004.06.010
83. Recaldin T, Fear DJ. Transcription factors regulating B cell fate in the germinal centre. *Clin Exp Immunol*. (2016) 183:65–75. doi: 10.1111/cei.12702
84. Lin KI, Angelin-Duclos C, Kuo TC, Calame K. Blimp-1-dependent repression of Pax-5 is required for differentiation of B cells to immunoglobulin M-secreting plasma cells. *Mol Cell Biol*. (2002) 22:4771–80. doi: 10.1128/MCB.22.13.4771-4780.2002
85. Tellier J, Shi W, Minnich M, Liao Y, Crawford S, Smyth GK, et al. Blimp-1 controls plasma cell function through the regulation of immunoglobulin secretion and the unfolded protein response. *Nat Immunol*. (2016) 17:323–30. doi: 10.1038/ni.3348
86. Shen Y, Hendershot LM. Identification of ERdj3 and OBF-1/BOB-1/OCA-B as direct targets of XBP-1 during plasma cell differentiation. *J Immunol*. (2007) 179:2969–78. doi: 10.4049/jimmunol.179.5.2969
87. Ren X, Siegel R, Kim U, Roeder RG. Direct interactions of OCA-B and TFII-I regulate immunoglobulin heavy-chain gene transcription by facilitating enhancer-promoter communication. *Mol Cell* (2011) 42:342–55. doi: 10.1016/j.molcel.2011.04.011
88. Corcoran LM, Hasbold J, Dietrich W, Hawkins E, Kallies A, Nutt SL, et al. Differential requirement for OBF-1 during antibody-secreting cell differentiation. *J Exp Med*. (2005) 201:1385–96. doi: 10.1084/jem.20042325
89. Carotta S, Willis SN, Hasbold J, Inouye M, Pang SH, Emslie D, et al. The transcription factors IRF8 and PU.1 negatively regulate plasma cell differentiation. *J Exp Med*. (2014) 211:2169–81. doi: 10.1084/jem.20140425
90. Fornek JL, Tygrett LT, Waldschmidt TJ, Poli V, Rickert RC, Kansas GS. Critical role for Stat3 in T-dependent terminal differentiation of IgG B cells. *Blood* (2006) 107:1085–91. doi: 10.1182/blood-2005-07-2871
91. Ding C, Chen X, Dascani P, Hu X, Bolli R, Zhang HG, et al. STAT3 signaling in B cells is critical for germinal center maintenance and contributes to the pathogenesis of murine models of lupus. *J Immunol*. (2016) 196:4477–86. doi: 10.4049/jimmunol.1502043
92. Suan D, Krautler NJ, Maag JLV, Butt D, Bourne K, Hermes JR, et al. CCR6 Defines memory B cell precursors in mouse and human germinal centers, revealing light-zone location and predominant low antigen affinity. *Immunity* (2017) 47:1142–53.e4. doi: 10.1016/j.immuni.2017.11.022
93. Shinnakasu R, Inoue T, Kometani K, Moriyama S, Adachi Y, Nakayama M, et al. Regulated selection of germinal-center cells into the memory B cell compartment. *Nat Immunol*. (2016) 17:861–9. doi: 10.1038/ni.3460
94. Basso K, Dalla-Favera R. Roles of BCL6 in normal and transformed germinal center B cells. *Immunol Rev*. (2012) 247:172–83. doi: 10.1111/j.1600-065X.2012.01112.x

95. Igarashi K, Ochiai K, Muto A. Architecture and dynamics of the transcription factor network that regulates B-to-plasma cell differentiation. *J Biochem*. (2007) 141:783–9. doi: 10.1093/jb/mvm106
96. Cobaleda C, Schebesta A, Delogu A, Busslinger M. Pax5: the guardian of B cell identity and function. *Nat Immunol*. (2007) 8:463–70. doi: 10.1038/ni1454
97. Zuccarino-Catania GV, Sadanand S, Weisel FJ, Tomayko MM, Meng H, Kleinstein SH, et al. CD80 and PD-L2 define functionally distinct memory B cell subsets that are independent of antibody isotype. *Nat Immunol*. (2014) 15:631–7. doi: 10.1038/ni.2914
98. Wang Y, Shi J, Yan J, Xiao Z, Hou X, Lu P, et al. Germinal-center development of memory B cells driven by IL-9 from follicular helper T cells. *Nat Immunol*. (2017) 18:921–30. doi: 10.1038/ni.3788
99. Bhattacharya D, Cheah MT, Franco CB, Hosen N, Pin CL, Sha WC, et al. Transcriptional profiling of antigen-dependent murine B cell differentiation and memory formation. *J Immunol*. (2007) 179:6808–19. doi: 10.4049/jimmunol.179.10.6808
100. Scheeren FA, Naspetti M, Diehl S, Schotte R, Nagasawa M, Wijnands E, et al. STAT5 regulates the self-renewal capacity and differentiation of human memory B cells and controls Bcl-6 expression. *Nat Immunol*. (2005) 6:303–13. doi: 10.1038/ni1172
101. Tunyaplin C, Shaffer AL, Angelin-Duclos CD, Yu X, Staudt LM, Calame KL. Direct repression of prdm1 by Bcl-6 inhibits plasmacytic differentiation. *J Immunol*. (2004) 173:1158–65. doi: 10.4049/jimmunol.173.2.1158

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

1.3 General introduction on OCT1, OCT2 and OBF1

1.3.1 Octamer motif

The conserved octamer motif [143,144], consisting of eight nucleotides with a consensus sequence of 5'-ATGCAAAT-3' (Figure 10A), has been found in every promoters and most enhancers in Ig regions [143,145-148]. The activity of *Ig* promoter and high level of B cell specific expression of *Ig* genes is mainly dependent on octamer motif[143,145,149]. Deletion of element in the *Ig* promoters leads to a substantial reduction of *Ig* transcription [150]. In addition to *Ig* genes, octamer motif is also critical for the transcription of other B cell specific genes [151-154]. Insertion of octamer element confers B cell specificity to a minimal promoter containing beta-globin TATA box [148]. Moreover, multimerization of octamer motif functions as a strong enhancer that is able to activate transcription from distal position in a B cell specific manner [146]. To be noted, the local regulatory context also plays a role in determining B cell specific, as octamer motifs are also present in several house-keeping genes such as histone *H2B* and small nuclear RNA genes [155,156].

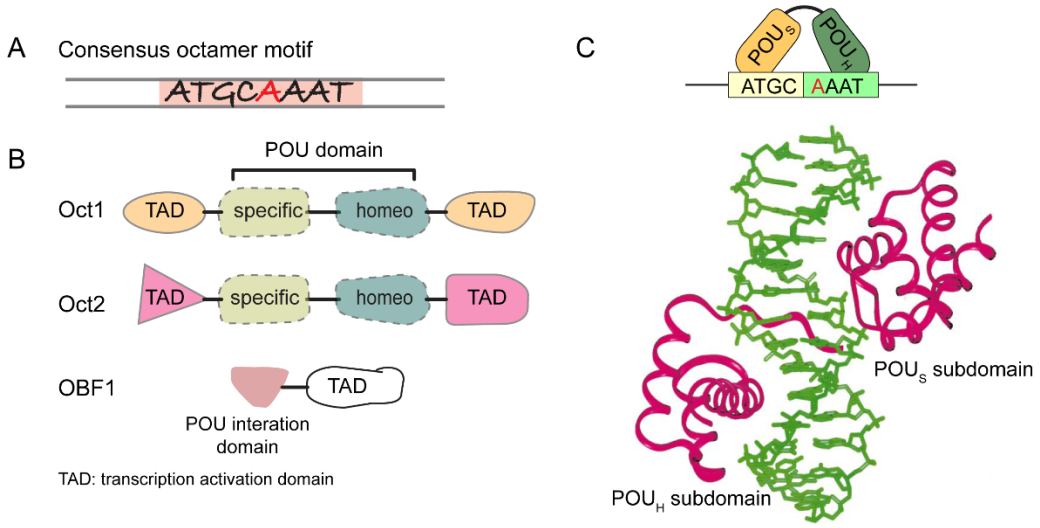


Figure 10 Oct factors and octamer motif. (A) The consensus sequence of octamer motif. (B) Scheme of the functional domains of OCT1, OCT2 and OBF1. OCT1 and OCT2 harbor two TADs in N- and C-terminus flanking the POU domain in the middle. TAD, transcription activation domain. (C) The upper panel shows the paradigm of the binding between POU_S and POU_H subdomains, with POU_S binding to ATGC and POU_H binding to AAAT. Lower panel, crystal structure of POU domain from OCT1 and octamer site from H2B promoter[157]. POU_S and POU_H subdomains are indicated.

1.3.2 OCT1 and OCT2

OCT2 and OCT1 are TFs founding members of the POU family[158] (Figure 10B). They bind to the octamer motif 5'-ATGCAAAT-3' [159] with similar specificity and affinity[160-163]

OCT1, encoded by gene *Pou2f1*, having a molecular weight of ~100 kDa, was purified from HeLa cells in 1987. It was found to be able to bind to the octamer motif, and was initially shown to activate the transcription of *H2b*[164]. OCT1 is ubiquitously expressed in all human and mouse cells[160]. OCT2, encoded by gene *Pou2f2*, having a molecular weight of 59-63 kDa, is a lymphoid restricted TF that also specifically recognizes and binds to the octamer motif. It initially had been shown to activate the transcription of Igκ chain promoter under *in vitro* condition[165,166].

These two POU domain TFs bind to octamer via the POU domain[160,161]. The POU domain is 160 amino acid long lies in the central part of OCT1, spanning from 280 to 439 residue[160,161]. Similarly, POU domain of OCT2 also lies in the central part of the OCT2 molecule[161]. The POU domains of OCT1 and OCT2 share 89% sequence similarity[160].

Structurally, POU domain is a unique DNA binding domain. It consists of two separate subdomains, termed POU-specific (POU_S) and POU-homeodomain (POU_H). Both of the POU_S and POU_H subdomains contain helix-turn-helix (HTH) motifs for DNA binding, and they are connected by a linker region[167]. Both of the POU_S and POU_H subdomains are capable to fold into independent structures and contact DNA in a sequence specific manner[167-169]. Crystallography of the complex of OCT1-POU and H2B octamer site showed that POU_S and POU_H subdomains bind to the opposite sides of octamer motif DNA by interacting with the major groove of DNA double helix. The POU_S subdomain interacts with 5'-ATGC-3', and the POU_H domain interacts with the 5'-AAAT-3' [157] (Figure 10C). To be noted, when binding to DNA, the two subdomains establish no direct protein-protein contact with each other, it seems that the POU_S and POU_H subdomains bind the two halves of octamer motif in an independent manner[157].

1.3.3 Transcriptional activity of OCT1 and OCT2

OCT1 is an essential gene, as *Pou2f1* knockout (KO) mice die in mid-late gestation stage for unknown reasons[170].

OCT1 is not essential for B cell development and function. B cell development and surface Ig expression are comparable between Oct1-deficient B cells and wild type (WT) counterparts. Late B cell functions, such as CSR and responses to antigens stimulation, are also unaffected. Furthermore, serum levels of IgM and IgA are only slightly reduced when OCT1 is ablated[171]. Molecularly, OCT1 has been shown to bind to regulatory elements of interleukin, Ig and histone genes. However, it is not essential for the transcription of endogenous *H2B*, *IgH* and *Igκ* genes[170,171]. Therefore, the OCT1 protein seems to be dispensable for B cell lineage development and function.

Although the role of OCT1 is unclear, OCT1 has been associated with the transcriptional regulation under genotoxic and oxidative stress. Upon treatment by ionizing irradiation, Set/Thr residues in the N-terminal transcriptional regulatory domain of OCT1 is phosphorylated by DNA-PK[172]. Phosphorylation of N-terminal domain of OCT1 is critical to the survival of cells under DNA damage stress. In line with this, OCT1 also interacts with BRAC1, which induces the expression of GADD45 through interaction with OCT1 during DNA damage stress response[173,174]. Moreover, OCT1 has been shown to interact with poly(ADP-ribose) synthetase (PARS), a protein critical for stress signaling regulation, and the interaction with PARS stabilizes the binding of OCT1 to target promoters during the DNA repair process.

OCT1 is ubiquitously expressed in all cell types, whereas OCT2 is largely restricted to the lymphoid system and also cells in the nervous system[162,175]. Because of the distinct expression pattern of OCT2, it was initially thought to play critical role in regulating the expression of *Ig* genes and early B cell development. Surprisingly, B cell develop normally in OCT2 deficient mice, and the rearrangement of Ig genes are similar to WT littermates. The transcriptional level of IgH and IgL is also normal in the absence of OCT2. However, late development B cell development and function are perturbed, as *Pou2f2*$^{-/-}$ B cells respond poorly to LPS stimulation, and GC formation is severely impaired[176,177]. The transactivation domains of OCT2 lie at both the amino- and carboxyl-terminus[178]. Luciferase assays with various truncated forms of OCT2 proteins showed that the N-terminal domain harbors low transactivation activity, whereas the C-terminal transactivation domain contains the major transcriptional activity and functions of OCT2. Transgenic mice expressing C-terminally truncated OCT2 (OCT2ΔC) showed similar phenotypes as *Pou2f2-/-* mice[179]. These mutant mice die shortly after birth, exhibit impaired B cell maturation, a complete lack of B1 cell population, decreased responses to stimulation, and a

dramatic reduction of serum IgG level. OCT2 target genes, such as *Cd36*, were significantly downregulated in OCT2ΔC mice. However, although carboxy-terminal domain (CTD) seems to play the most significant transcriptional functions of OCT2, the cooperation between N- and C-terminus of OCT2 remain unclear. To be noted, the C-terminal regions of OCT1 and OCT2 are highly different[160], indicating their distinct functions in activating gene transcription.

1.3.4 Interaction partners of OCT1 and OCT2

Even though OCT1 and OCT2 contain very similar POU domains and bind to the same set of octamer and octamer-like motifs [143,144,158,180,181], these two TFs exhibit distinct transactivation properties. Therefore, they serve as a paradigm for studying how two TFs, with highly similar DNA binding specificity and ability, exert differentiation functions in regulating the transcriptional activities of specific target genes[158]. Indeed, the differential transcriptional regulatory roles of these two TFs lies in the unique abilities to recruit distinct sets of TFs or co-factors that confer the selectivity in the transactivation activity from specific loci.

OBF1 enhances the transcriptional activity of OCT1 and OCT2 from the IgH promoter, however not from the H2B promoter[182], suggesting that an additional cofactor might be needed in order to stimulate H2B transcription. Indeed, a cofactor, termed OCT1 CoActivator in S phase (OCA-S), was discovered to interact with OCT1 in activating H2B transcription during S phase[183].

HMG2 (High mobility group protein 2) has been shown to interact with OCT1 and OCT2 and enhance the transactivation activity of these two TFs on promoters containing octamer motifs[184].

Retinoid X receptors (RXR), are responsible for controlling development, differentiation and homeostasis by translating signals from specific ligands into transcriptional activity. RXR has been implicated to interact with the POU domain of OCT1 and OCT2, which was believed to be essential for nuclear hormone receptors in regulating gene transcription[185].The DNA binding domain (DBD) of glucocorticoid receptor (GR) binds directly to the POU_H of OCT1 and OCT2[186]. These findings may provide possible explanations why OCT1- or OCT2-decifient mice die during embryo development or shortly after birth[186]. In addition, androgen receptor (AR) and progesterone receptor (PR) also bind to the POU domains of OCT1 and OCT2, which enhances the binding of OCT1 and OCT2 in promoter of mouse mammary tumor virus (MMTV)[187].

MAT1, a subunit of CDK-activating kinase (CAK), interacts with OCT1 and OCT2 through POU domain in a DNA-independent manner[188], and is important to enhance the phosphorylation of POU domain of OCT1 by CAK. The interaction between OCT-MAT1-TFIIH might be important for the formation of preinitiation complex and the ensuing elongation reactions. Furthermore, C-terminal domain of TBP, a TATA box binding subunit of TFIID factor, also interacts with POU domains of OCT1 and OCT2[189]. Therefore, OCT1 and OCT2 exert direct activities on RNA transcription through interaction with the transcription preinitiation complex.

VP16 is a very potent viral transcription cofactor that only interacts with OCT1 but not OCT2[190], therefore OCT2 fails to stimulate the expression of genes regulated by OCT1-VP16 complex (Figure 11). Incorporation of VP16 enhances the binding of OCT1 on an alternative motif 5'-(OCTA-)TAATGARAT-3'. OCT1 or OCT2 bind weakly to this low affinity motif, thereby are unable to effectively activate promoters containing this motif. However, the specific interaction of VP16 stabilizes the binding of OCT1 on this motif, and subsequently stimulates target genes of OCT1-VP16 complex [191]. This illustrates one mechanism of how OCT1 and OCT2 exert different transactivational activities.

Moreover, the tissue specific expression patterns of OCT1 and OCT2 leads to selective cellular presences and relative abundance of the two TFs, which limits the possible factors with which they can interact and the intensity in competing binding sites, which eventually leads to differential transactivation activities.

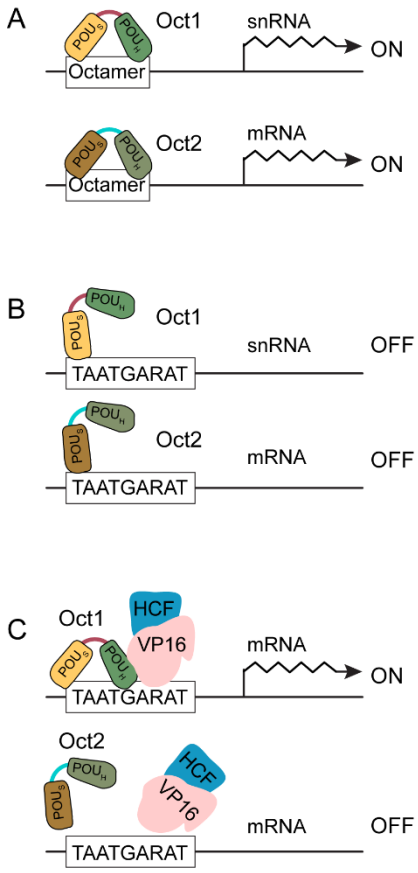


Figure 11 Selective transcriptional regulation of OCT1 and OCT2 by interaction with VP16 protein. (A) OCT1 and OCT2 bind to consensus octamer motif with high affinity and subsequently drive the expression of genes downstream of the octamer site. (B) OCT1 and OCT2 are unable to drive transcription of the downstream genes from low affinity site, 5'-TAATGARAT-3'. (C) OCT1, but not OCT2, binds efficiently with VP16 and HCF, and effectively drives the transcription of downstream genes. This figure is adapted from[158].

1.3.5 OBF1

As mentioned above, OCT1 is a ubiquitous TF and OCT2 is largely B cell specific, therefore OCT2 was initially thought as the main factor that control the Ig transcription[192]. However, Ig expression was not perturbed in OCT2 deficient B cells[176,193,194]. In addition, the expression level of Ig in B cell extracts does not correlate with protein amount of OCT2[195].

OBF1 (aka OCA-B/Bob.1), a cofactor expressed mainly in B cells[196], was originally discovered biochemically as a protein present in B cell nuclear extract using OCT1 affinity chromatography[193,197], which was able to stimulate *Igh* promoter transcription in vitro through interaction with OCT1 and OCT2 molecules[193]. It was subsequently cloned by several

groups[182,196,198]. Based on their B cell specific expression, OCT2 and OBF1 have been thought to be critical factors for Ig expression[162,182,196]. However, *Ig* transcription is still normal in OBF1 ablated B cells, as well as in cells lacking both OCT2 and OBF1 [194,199,200] In contrast, the expression of antigen specific class-switched Ig genes and GC formation are dependent on OBF1, suggesting that OBF1 is important for late B cell function and development[200-202].

Electrophoretic mobility shift assays (EMSA) showed that OBF1 specifically interacts with OCT1 and OCT2 through their POU domain, whereby it forms a ternary complex on DNA[155,196,203-207]. OBF1 contains no intrinsic DNA binding activity and accesses the DNA in an OCT1/OCT2-dependent manner. In addition, biochemical studies with truncated forms of OBF1 showed that the N-terminus of OBF1 is responsible in mediating the interaction with OCT1 and OCT2. However, it is insufficient in enhancing transcriptional activity of OCT factors, therefore suggesting the existence of a transactivation domain[196,203]. The C-terminal part of OBF1 contains transcriptional activation activity. OBF1 (Δ1-122) truncated shows a 90% trans-activation activity compared to full length OBF1. OBF1 has weak transcription activation independently of Oct factors: a fusion protein of GAL4 and C-terminal of OBF1 is able to activate transcription of a target promoter over 3-fold compared to the GAL4 DNA binding domain alone[203].

The direct interaction between OBF1, POU domain and DNA was confirmed by crystallography of OCT1-POU/OBF1/octamer complex[203,208] (Figure 12). This interaction is highly specific[209], as OBF1 could not recognize and bind to POU domains of OCT3, OCT4, OCT6, Brn1 or Brn2, even though these factors could bind to octamer motifs as OCT1 and OCT2 [196,198].

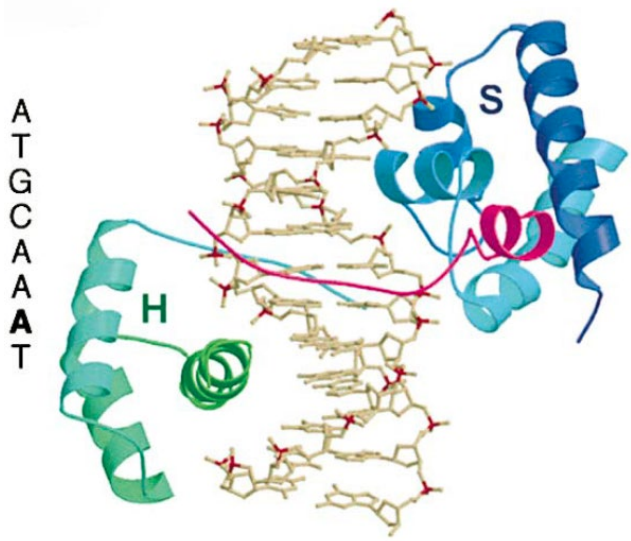


Figure 12 Tenary structure of OCT1/OBF1/octamer complex. POU domain of OCT1 binds to octamer motif as monomer, POU_H and POU_S subdomains are indicated as H and S. OBF1, in magenta, interacts with POU domain on octamer motif. This figure is adapted from [208,210].

Two isoforms of OBF1 are produced by alternative splicing, which yields two proteins of 34 and 35 kDa, termed p34 and p35, respectively[182]. Both of these two isoforms are able to initiate octamer-dependent transcription; intriguingly, the N-terminus of p35 can be modified by N-myristoylation and lead to a distribution of the protein to the cell membrane with unclear function[211]. Furthermore, the protein level of OBF1 was shown to be also regulated by interaction with the E3 ligase Siah-1 protein[212,213].

1.3.6 The interaction between OBF1 and OCT1/OCT2

1.3.6.1 Interaction on Octamer motif

OCT1 and OCT2 bind to the DNA as either monomers or dimers, depending on the binding sites [214,215]. POU_S and POU_H subdomains bind directly to DNA[157]. POU_S binds mainly to ATGC sequence from the 5' half of the octamer motif, and POU_H establishes contact with the AAAT sequence at the 3' half. Crystal structure of the POU-DNA complex showed that these two POU subdomains bind to DNA independent from each other without protein-protein interaction between them; POU_S and POU_H occupy opposite sides of the octamer double helix. POU_S and POU_H are connected by an unstructured linker peptide. Each of the two subdomains recognizes and binds to DNA via helix-turn-helix (HTH), and helix 3 of each subdomain is positioned in the

major groove and engaged in extensive interaction with corresponding halves of the octamer motif. Residues Arg20, Gln27, Gln44 and Glu51 from helix 3 of POU_S subdomain are conserved in all known POU_S domains that are critical in contacting nucleotides and strengthening the protein-DNA interaction. The structure and DNA-binding character of POU_H are conserved in homeodomain DNA complexes[216-218]. POU_H subdomain interacts with DNA in both major and minor groove. POU_H subdomain contacts the DNA major groove through helix 3, and also minor groove with the N-terminal residues. It interacts with AAAT of the 3' half of the octamer motif, and residues Val47, Cys50, Asn51 and Gln54 in POU_H subdomain are responsible for establishing contacts with octamer motifs of major groove. Residue Arg5 of POU_H subdomain makes contacts with adenine in the fifth position of octamer motif in minor groove.

POU_S and POU_H subdomains bind DNA in a cooperative manner[157]. Individual POU_H binds to DNA efficiently, while POU_S binds poorly to DNA on its own. Also, the spacing between ATGC (POU_S binding site) and AAAT (POU_H binding site) predominantly affect the binding of POU domain. Elongating the spacing by one or more base pairs leads to the rapid reduction in associating constant [157]. Surprisingly, these two subdomains are structurally independent and form no protein-protein contact while binding on DNA. However, crystal structure shows that binding of the POU domain results in alterations in DNA structure[157]. Together with the flexibility of linker peptide, these two subdomains bind to DNA in a cooperative manner, even though protein-protein contacts are missing.

Although OBF-1 exhibits no independent DNA-binding capacity, the formation of OBF1-OCT complex on DNA is sequence-specific. OBF1 interacts with OCT1 and OCT2 on canonical octamer motifs [182,196,198,219]. However, this interaction is sequence dependent, particularly the fifth positon of octamer motif. OBF1 does not form a complex with OCT1-POU on HSV αIE element which contains an octamer-like motif, 5'-<u>ATGCTAAT</u>GATATT-3', which differs from canonical octamer motif at the fifth nucleotide. Switching the fifth thymine to adenine restores the OBF1 incorporation. By EMSA and luciferase assays, OBF1 forms ternary complex with OCT factors on octamer containing elements (Figure 13A and 13B), and drives the transcription from these promoters. Mutation in the octamer sequence, mutating the fifth adenine to thymine, specifically abrogated the binding of OBF1 and diminished the transcription [196,198]. Therefore, it is the fifth position of octamer motif that is required for the interaction between OBF1 (Figure 13A), OCT factors and DNA, regardless of the surrounding DNA sequences[204].

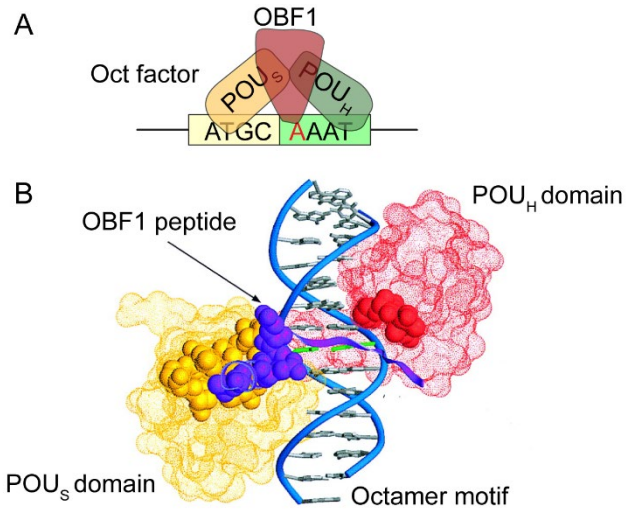


Figure 13 Interaction between OBF1 and OCT factors. (A) Paradigm of the binding between OBF1 and POU domain on octamer motif. POU_S subdomain binds to ATGC and POU_H binds to AAAT. OBF1 interacts with both POU_S and POU_H subdomains and the fifth A of the octamer motif. (B) The configuration of POU_S subdomain (yellow), POU_H subdomain (red) and OBF1 peptide (purple) are displayed. This figure is adapted from[208].

Amino acids 1-44 of OBF1 (OBF1N1-44) exhibit high affinity in binding to OCT1 and display selective requirement for an adenine at the fifth position of octamer motif in order to form a complex with OCT factors [208]. Crystal structure of OCT1-POU/octamer/ OBF1N1-44 shows that OBF1 peptide covers the major groove and interacts with A/T base pair at the fifth position. Adenine at fifth position interacts with Val22 from OBF1 via two hydrogen bonds, therefore conferring the specificity observed in EMSA studies. OCT1 POU_S and POU_H subdomains bind the opposite sides of the DNA double helix with N-terminus of POU_H subdomain positioned in the minor groove near the center of octamer motif. And, POU_S and POU_H subdomains bind each half of the octamer motif in a cooperative manner flanking the fifth nucleotide[157,204]. Inserting two base pairs in the middle of two subsites of consensus octamer motif leads to decreased binding affinity of OCT1, however OBF1 could still interact with OCT1 on the elongated motif, as long as adenine exists in the fifth position[204].

These results demonstrated that the DNA contact is needed by OBF1 to interact with OCT factors, although OBF1 has no ability to bind DNA independently.

1.3.6.2 Interaction on tandem octamer motifs, PORE and MORE

Initially, OCT1 and OCT2 were identified to bind to the octamer motif as monomers. Later, homo- and heterodimers of these two TFs were discovered on the *IgH* promoters[214,215,220] and in the first intron of the OPN gene[221], which indicated cooperative DNA binding between these two proteins. These motifs were designated PORE (Palindromic Oct factor Recognition Element) and MORE (More palindromic Oct factor Recognition Element), and they are different in terms of the binding of OBF1.

The PORE motif was first discovered in the first intron of *OPN* gene[221]. It is composed of a palindromic pair of POU_H subdomain binding sites separated by a 5-base-pair spacer (ATTTG+N_5+CAAAT) (Figure 14A). Mutations on PORE motifs showed that the ATTTG and CAAAT are vital for POU dimer formation; each half of the PORE motif is sufficient to support the binding of monomeric POU domain. However, the spacing of the central 5 nucleotides is also important in supporting the POU dimer formation, as revealed by the EMSA experiments using PORE derivative sequences with various spacer lengths[221]. POU domains bind to the PORE motif in a configuration with POU_S positioned near the center of the PORE motif and POU_H at the distal sides[222]. On the PORE motif, two subdomains of each POU molecule bind along the DNA helix (Figure 14C), which supports the finding that the binding efficiency of the POU domain on PORE motifs is dependent on the spacing between two halves of the PORE. Two subdomains of the POU domain form two different protein-protein interfaces, due to the asymmetric nature of the PORE motif. In the smaller interface, the N-terminus of the POU_H subdomain interacts with several residues of POU_S. In the larger interface, POU_S and POU_H interact with each other via the contact with the same nucleotide, both of POU_S and POU_H domains interact with nucleotide at position 10 in the minor groove. Furthermore, POU_S and POU_H subdomains also interact each other when binding to DNA[222].

The MORE motif is found in an array of promoters, with a consensus sequence of 5'-ATGCATATGCAT-3'[210] (Figure 14A). Upon binding on MORE motif, the POU_S is situated in the middle of the motif and interacts with ATGC sequence, while the POU_H binds to the distal part [222] and interacts with AT sequence[210]. On MORE motifs, each POU domain binds half the motif across the double helix (Figure 14B), therefore, insertion of the spacer in the center of MORE motif does not influence the POU dimer formation. Indeed, POU dimerization can

tolerate two base pair insertion in the middle of MORE motif. Due to the symmetric trait of the MORE motif, two subdomains of the POU domain form identical protein-protein interface on each half of the MORE motif. In this interface, Ile159 from POU_H interact with hydrophobic pocket in POU_S subdomain, and residues 157-160 of POU_H subdomain interacts with helix 3/4 of POU_S subdomain.

In both cases of PORE and MORE, POU_H interacts with major groove, and residues Asn151 and Gln154 makes base contacts[222].

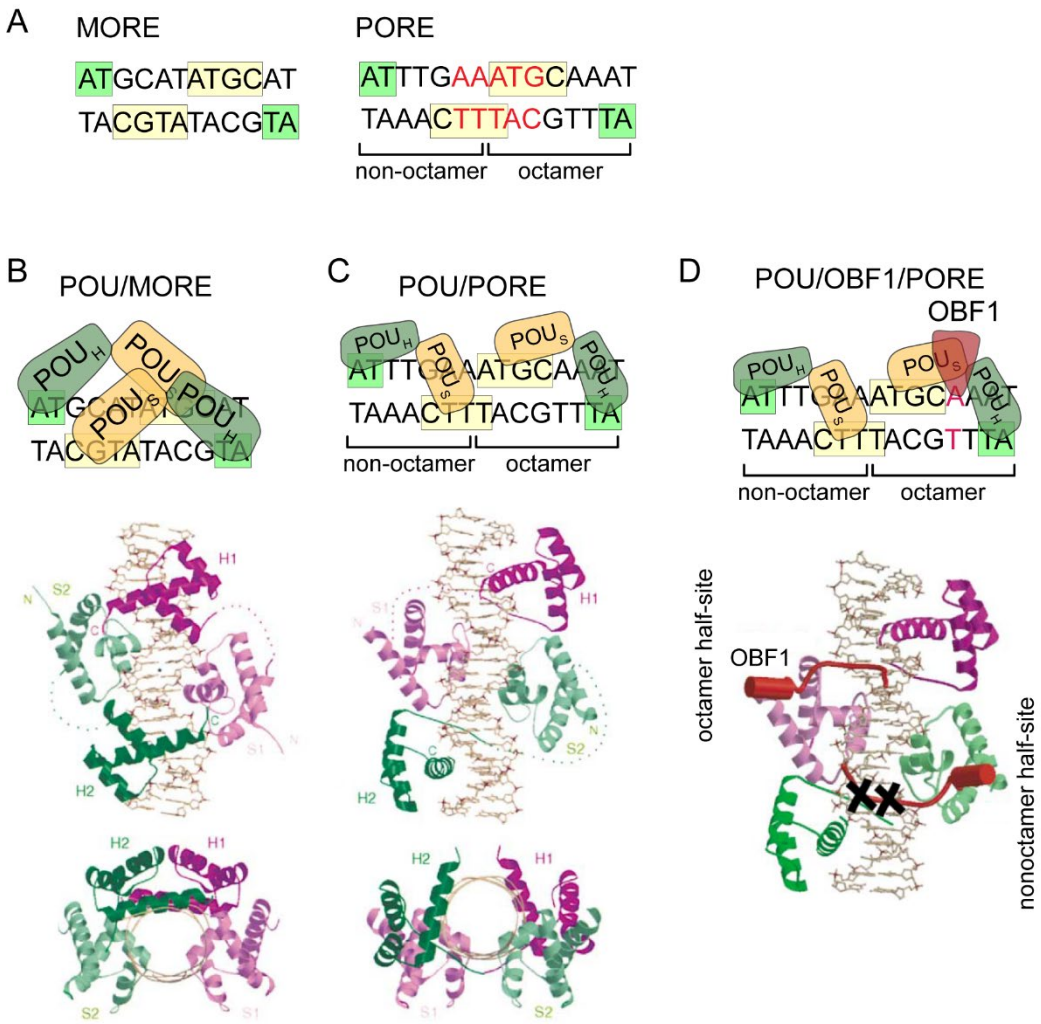


Figure 14 Interaction between OBF1 and OCT factors on tandem sites. (A) Structure of PORE and MORE motifs; spacers of PORE motif is indicated in red; the binding sequences for POU_H (5'-AT-3') and POU_S (5'-ATGC-3', 5'-TTTC-3') subdomains are indicated in green and yellow, respectively. (B) Upper panel, paradigm of the OCT factor dimer formation on MORE motif. POU_S and POU_H are colored in yellow and green, respectively. Binding sites for POU_S and POU_H are indicated in yellow and green, respectively. Lower panel, the bipartite OCT1

POU domain binds to the MORE motif as a homodimer. POU_S subdomain (green) comprises four α-helices, and POU_H subdomain (magenta) comprises three α-helices. The structure of the complex from the view of DNA axis is also shown. The palindromic center of the MORE motif is indicated with an asterisk. N terminus and C terminus of each POU subdomains are indicated with N and C. The dotted lines indicate flexible linker regions. (C) Upper panel, paradigm of the OCT factor dimer formation on PORE motif. POU_S and POU_H are colored in yellow and green, respectively. Binding sites for POU_S and POU_H are indicated in yellow and green, respectively. Lower panel, the bipartite OCT1 POU domain binds to PORE motif as a homodimer. The structure of the complex from the view of DNA axis is also shown. N terminus and C terminus of each POU subdomains are indicated with N and C. The dotted lines indicate flexible linker regions. (D) Upper panel, paradigm of the OCT factor dimer and OBF1 complex on PORE motif. POU_S, POU_H and OBF1 are colored in yellow, green and red, respectively. Binding sites for POU_S, POU_H and OBF1 are indicated in yellow, green and red, respectively. The lower panel shows the superimposition of the OBF1 with the POU/PORE complex. OBF1 is only allowed to interact with the POU domain occupying the octamer half site of PORE motif due to the different geometric parameters (crossed). This figure is adapted from[222].

OBF1 specifically interacts with a homo- or heterodimer of OCT1 or OCT2 on PORE motif, and augments the transcriptional activity[210], whereas it fails to interact with an OCT1 dimer in MORE. Therefore, the distinct configuration of the OCT1 dimer on PORE determines the recruitment of OBF1. Compared to the OCT1/octamer/OBF1 structure[208], the OCT1 molecule in the dimer occupies the octamer-like half-site 5'-ATTTGAAAGGCAAAT-3' in PORE and exhibits a structural configuration allowing the binding of OBF1 (Figure 14D).Therefore, for each OCT dimer on PORE motif, one OBF1 molecule is allowed in the complex. OBF1 interacts with the C-terminus of helix 3 of the POU_H subdomain, which is critical for the OBF1 recruitment[208]. In the other non-octamer half-site, POU_H subdomain is distorted, therefore precluding the binding of OBF1[222]. This finding has been corroberated by EMSA assays showing that only one OBF1 molecule incorporates with OCT1 dimer on the probe containing PORE[222].

1.3.7 OCT factors and OBF1 are dispensable for early B cell development

Two checkpoints exist in early B cell development in the bone marrow: the rearrangement of *IgH* and *IgL*. *IgH* rearrangement starts in proB cells at *IgH* locus. *IgH* chain rearrangement initiates with the assembly of diversity (D) and joining (J) gene segments followed by the joining between variable regions (V) to rearranged DJ segment. Once the functional *IgH* rearrangement is finished, IGH protein is produced, which then forms the pre-BCR with surrogate light chain (Igα/Igβ). pre-

BCR processes signals from the microenvironment and orchestrates proB cells to differentiate into preB cells. At small preBII stage, light chain V-J joining takes place. Following the production of functional light chain, the rearranged IGH and IGL form the BCR, and B cells further differentiate into immature B cells[223].

At *Ig* loci, transcription and V(D)J recombination of *Ig* gene segments are linked. Transcription increases the accessibility of unrearranged Ig gene segments, therefore facilitates the accessibility of the recombination machinery. The octamer motif is highly conserved in the promoters of V regions, and this motif is important for regulating the transcriptional activity of *Ig* enhancers and promoters[56]. OCT1, OCT2 and OBF1 have been shown to activate *Ig* genes transcription by binding to the octamer motifs in the *Ig* promoters and enhancers[199,201]. Mutation in the octamer motif or deletion leads to dramatic decrease of the transcriptional activities from these promoters in both preB and mature B cell lines[223]. OBF1 interacts with OCT1 or OCT2 on octamer motif, and stimulated the transcription from Igκ light chain gene promoters[224] in luciferase assays[196]. Because of the connection between Ig transcription and V(D)J rearrangement, which is further coupled with B cell development, these factors were initially considered to play vital roles in early B cell development.

However, results from mouse models surprisingly showed that these factors are less important during early B cell development than previously anticipated. *Pou2f2*$^{-/-}$, *Pou2af1*$^{-/-}$ and even the double mutant mice showed normal transcription of Ig genes and early B cell development.

Kim et al found that *Pou2af1*$^{-/-}$ mice exhibit normal IgM mRNA and protein levels, as well as sterile IgM transcripts compared with WT mice. FACS analysis showed that the number of early stages B cells in bone marrow and mature B cells in spleen are also normal[200-202] Therefore, IgM expression and B cell development were not perturbed by the deficiency of OBF1. Meanwhile, *Pou2af1*$^{-/-}$ mice showed normal CD4/CD8 and TCR-αβ positive T cell compartments, indicating that normal T cell development is not dependent on OBF1[200-202] In the fetal liver of *Pou2f2*$^{-/-}$ embryos, the numbers of preB cells harboring rearranged Ig genes are comparable with wild-type micea and the transcription of heavy and light is also normal in preB cells of *Pou2f2*$^{-/-}$ mice[176]. Most interestingly, Ig production is still normal even in the absence of both OCT2 and OBF1[199].

Results from *Pou2f2*$^{-/-}$ or *Pou2af1*$^{-/-}$ mice suggested that OCT1 might play an essential role in driving the B cell specific transcription of the Ig locus. The functional relevance of OCT1 in B cell

development was studied in a transplantation model by injecting fetal liver cells of wild-type and *Pou2f1*$^{-/-}$ mice into irradiated RAG-1 deficient mice[171]. B cell development in the bone marrow and spleen was normal in *Pou2f1*$^{-/-}$ and WT mice[171]. Furthermore, the expression levels of Ig was also normal in *Pou2f1*$^{-/-}$ mice. Unlike *Pou2f2*$^{-/-}$ B cells, which respond poorly to LPS stimulation, OCT1 depleted B cell proliferated normally when stimulated by anti-IgM + IL4, anti-CD40, anti-CD40 + IL4 or LPS. Moreover, the frequency of Vκ light chain usage was comparable between OCT1 deficient preB cells[171].

In summary, OCT1, OCT2 and OBF1 are not necessary for the development of early B cells.

1.3.8 OCT factors and OBF1 are required for late B cell development and function

Although OCT factors and OBF1 are largely dispensable for the development of early B cells, they are necessary for the functions and development of B cells at late stages, including the development of peritoneal B1 cells, splenic marginal zone (MZ) B cells, responses to stimulations, GC formation and post-GC differentiation.

The expression level of OCT2 peaks in peritoneal B1 cells, and becomes reduced along the terminal differentiation path to antibody-secreting cells (ASCs)[225]. In line with this, using an adoptive transfer model it was found that peritoneal B1 cells, which reside mainly in the peritoneal cavity and are different from splenic B cells (B2 cells) [226-228], are completely lost in the absence of OCT2[229]. In contrast, the peritoneal B1 compartment is enlarged in OBF1-deficient mice[230]. In addition to the peritoneal B1 population, splenic MZ B cells are missing in both OCT2- and OBF1-deficient mice[229,231].

OCT2 deficient mature B cells showed severely reduced proliferation under LPS or anti-IgM stimulation [176,177], and *Pou2f2*$^{-/-}$ B cells were arrested in G1 phase of the cell cycle[177]. Although OCT2 and OBF1 are dispensable for Ig expression in early B cells, they are vital for the expression of class switched antibodies, particularly those that are T cell dependent. The IgG serum level is dramatically reduced in the absence of OCT2 or OBF1[177,200-202].

Furthermore, mice deficient for OCT2, OBF1 or both genes show a complete lack of GC formation [200-202,232], and the post-GC PCs generation[225,233,234].

1.3.8.1 GC formation and downstream differentiation

The role of OCT2 in GC formation is unclear, and seems to be antigen dependent. OCT2-deficient mice lacked GCs following 4-hydroxy-3-nitro-phenyl-acetyl–ovalbumin (NP-OVA) immunization[199]. In contrast to the GC deficiency when challenged with NP-OVA, normal GCs formation occurred in *Pou2f2*$^{-/-}$ mice after influenza virus infection [235]. However, GC formation, as well as GC B cell generation, is severely impaired in influenza virus infected *Pou2af1*$^{-/-}$ mice[235]. Later, Hodson et al found that NP-KLH challenge successfully elicited GC formation with normal GC B cell number in *Pou2f2*$^{-/-}$ mice. However, these GC B cells showed retarded proliferation, and reduced levels of serum antigen-specific antibodies, and PC production is also impaired. However, the number of Bmems is normal in these mice, and the *Pou2f2*$^{-/-}$ Bmems respond and expand normally when re-immunized with the same antigen. To sum up, OCT2 seems to be dispensable for the development and immune responses of Bmems[236]. And, OCT2 is required for GC formation in an antigen-specific manner, but the normal functions of GC B cells require OCT2 expression.

OBF1 is indispensable for the GC formation in each antigen tested so far, and it is essential for the development and functions of GC B cells, as well as post-GC differentiation. GC formation in OBF1 deficient mice is completely blocked when immunized with NP-OVA[199], NP-KLH[236], NP-CGG[237], influenza virus[235], SRBC ([238] and this study) and DNP-KLH (this study). Although GC formation is blocked, T_{FH} cell differentiation is normal in OBF1-deficient mice. Therefore, the impaired GC formation is due to the intrinsic defects in *Pou2af1*$^{-/-}$ B cells[235]. In addition to the compromised GC formation, serum level of secondary Ig isotypes, including IgGl, IgG2a, IgG2b, IgG3, IgA and IgE, are greatly reduced in *Pou2af1*$^{-/-}$ mice, while IgM level is unchanged[200,201]. Under *in vitro* stimulation, IgG production is severely reduced in *Pou2af1*$^{-/-}$ B cells under the stimulation of CD40L and IL4[239]. Moreover, PC differentiation is also impaired, as master regulators, such as *Prdm1* and *Xbp1* fail to be induced in *Pou2af1*$^{-/-}$ cells[239]. In another study, OBF1 was found to specifically control the production of IgG in a plasmacytoma cell line, MPC11, by mediating long-range chromatin interaction between *Igh* promoter and 3' *Igh* enhancer. ChIP-seq revealed that upon the knockdown of OBF1 in MPC11 cells, the level of H3K4me3, which is associated with transcriptional activation is significantly reduced[240]. Therefore, OBF1 is indispensable for IgG production in PCs. Elevated level of OBF1 suppresses the generation of plasmablasts, presumably through repression of MCL-1[241]. And, in our study, we found that

downregulation of OBF1 in GC B cells is essential for activating post-GC differentiation, by activating genes involved in PC development, including *IRF4* and *PRDM1*. OBF1 is also required for Bmem differentiation. Overexpression of OBF1 in peripheral naïve B cells when stimulated with CD40L and IL21, leads to the suppression of BCL6, and uregulation of CD20 and ABF1, which eventually promote the differentiation of Bmems [242].

Therefore, GC B cells, PCs and Bmems are strictly dependent on OBF1 but not OCT2. OBF1 functions at multiple stages of GC reaction, and studies that focus on the molecular mechanisms underlying OBF1 function in GC formation are highly relevant.

1.3.8.2 Humoral immune response depends on OCT factors and OBF1

Due to the impaired expression of class-switched Ig isotypes[201], impaired GC formation[200,201] and abnormal Ig repertoire[243], the humoral immune response is affected in *Pou2af1*$^{-/-}$ and/or *Pou2f2*$^{-/-}$ mice. One study showed that OBF1 null mice showed dramatically decreased IgM and IgG titer and increased susceptibility to viral infection when challenged with live Vesicular stomatitis virus (VSV)[232]. Similar to VSV, T-dependent lymphocytic choriomeningitis virus (LCMV) fails to induce humoral immunity in mice with the depletion of OBF1[232]. In addition, OBF1 and OCT2 are needed for the production of IL6 which is important for the maturation of T_{FH} cells during B:T interaction[244,245]. Furthermore, OBF1 is required for CD40L and IL4 signaling in GC B cells[237].

Moreover, OBF1 has also been implicated in the autoimmune responses. Systemic lupus erythematosous (SLE) is an autoimmune disease. Its symptoms are highly variable among patient cohorts, however anti-nuclear antibodies seem to be the common symptom between patients [246]. Although the exact cause of this disease is not entirely clear, B cells seem to be the most important cause[247]. In addition, Aiolos$^{-/-}$ mice naturally develop the symptoms of SLE[248], therefore serves as animal model of human-like SLE. The SLE phenotype of Aiolos deficient mice are at least partially dependent on OBF1, as concomitant ablation of both OBF1 along with Aiolos reversed the SLE phenotype and blocked the spontaneous GC formation[249]. Another study showed that the expression of OBF1 is silenced in samples from SLE patients[246]. Therefore, OBF1 emerges to be a key factor in inhibiting SLE development. Sanroque mice bear a mutation in the *Roquin1/Rc3h1* gene, which produces a mutant Roquin1 protein conferring constant signaling of ICOS, a costimulatory receptor highly expressed on follicular T helper cells (T_{FH}). These mice display

autoimmune symptoms, producing autoreactive Igs and displaying spontaneous GC formation driven by T cell intrinsic factors[250]. Chevrier et al found that, although *sanroque*-mediated phenotypes were not completely rescued, OBF1 deficiency did inhibit the $CD4^+$ T cell to T_{FH} differentiation and GC formation in these mice[230]. Therefore, these *Pou2af1*$^{-/-}$ *sanroque* mice developed autoimmune pathology in a GC-independent way. Autoimmune complexes of autoreactive IgG were absent in blood or kidneys`, however anti-DNA or anti-nuclear IgM complexes abundantly occur in the kidneys of mutant mice which leads to glomerulonephritis[230]. These results coincide well with the findings that OBF1 is essential for the GC formation and expression of isotype-switched Igs[201], and it is essential for the development of GC B cells into Bmems and PCs[239,242]. Thus, these results provide a functional mechanism of OBF1 in autoimmune diseases.

1.4 Aim of this thesis

In the lab of Prof. Patrick Matthias, one focus has been on the transcriptional regulatory roles of OCT1, OCT2 and OBF1. The TFs OCT1 and OCT2 and their cofactor OBF1 had been found to bind to the same set of DNA motifs and to play key roles in late B cell function and development. OCT2 and OBF1 were the first B-cell-specific TFs identified and have been shown by the Matthias lab[199,200] and other groups[201,202] to be essential for the GC reaction and humoral immune responses. More recently, OBF1 was identified as a critical gene for DLBCL proliferation[251,252] and the importance of OCT2 for DLBCL was also evidenced[236]. Yet, no mechanism was available to explain how these factors connect to GCs and the humoral immune response. Therefore, it is highly relevant to address their functional mechanisms in B cells and humoral immune response.

We first decided to map the genome-wide binding patterns of these factors comparatively. To this end, we established mice in which these factors are endogenously tagged with an AviTag; this allowed us to perform Bio-ChIP-seq for these three factors and gave rise to data sets that are more comparable than antibody-based ChIP-seq (also because we would have needed to use three different antibodies). The genome-wide binding map of these factors allowed us to interrogate their detailed binding characteristics, between themselves and in relation to histone modifications. Moreover, comparing the binding enrichment of OCT factors in WT and *Pou2af1*$^{-/-}$ B cells gave us conclusive evidence as to whether OBF1 stabilizes the chromatin binding of OCT1 or OCT2. In addition, motif analyses of the binding regions of these factors allowed us to identify potential co-localizing factors that would have been difficult to discover otherwise.

We also aimed to uncover the mechanisms by which these factors control the GC reaction and GC-derived B lymphoma cells. We wanted to establish a regulatory hierarchy between OCT1, OCT2 and OBF1 in controlling the GC program and for this used using GC-derived B lymphoma cells. Genomic binding maps of these factors in primary murine and human GC B cells further facilitated the interrogation of the target genes of these factors in a physiological GC context. In this way, we wanted to dissect the fundamental mechanism explaining the essential functions of these factors in GC formation. As OBF1 and OCT2 have been implicated to be critical for the proliferation of GC-derived B lymphoma cells, including Burkitt's lymphomas and DLBCLs, we

wanted to interrogate the relevance of these factors in B cell lymphomas. Based on our results, we propose OBF1 or OCT2 to be novel therapeutic targets of GC-derived B cell lymphomas.

Results

Chapter 2 Results

2.1 OBF1 and Oct factors control the germinal center transcriptional program

(Accepted, 2021, Blood)

American Society of Hematology
2021 L Street NW, Suite 900,
Washington, DC 20036
Phone: 202-776-0544 | Fax 202-776-0545
editorial@hematology.org

OBF1 and Oct factors control the germinal center transcriptional program

Tracking no: BLD-2020-010175R1

Shuang Song (Friedrich Miescher Institute for Biomedical Research, University of Basel, Switzerland) Chun Cao (Friedrich Miescher Institute for Biomedical Research, University of Basel, Switzerland) Mohamed-Amin Choukrallah (Friedrich Miescher Institute for Biomedical Research, University of Basel, Switzerland) Fengyuan Tang (Department of Biomedicine, University of Basel, Switzerland) Gerhard Christofori (University of Basel, Switzerland) Hubertus Kohler (Friedrich Miescher Institute for Biomedical Research, University of Basel, Switzerland) Fabian Wu (Novartis Institutes for Biomedical Research, Switzerland) Barna Fodor (Novartis Institutes for Biomedical Research, Switzerland) Mathias Frederiksen (Novartis Institutes for Biomedical Research, Switzerland) Simon Willis (Department of Medical Biology, University of Melbourne, Australia) Jacob Jackson (Department of Medical Biology, University of Melbourne, Australia) Stephen Nutt (Department of Medical Biology, University of Melbourne, Australia) Stefan Dirnhofer (pathology, Switzerland) Michael Stadler (Swiss Institute of Bioinformatics, Switzerland) Patrick Matthias (Faculty of Sciences, University of Basel, Switzerland)

Abstract:
OBF1 is a specific coactivator of the POU family transcription factors OCT1 and OCT2. OBF1 and OCT2 are B cell-specific and indispensable for germinal center (GC) formation, but their mechanism of action is unclear. Here, we show by ChIP-seq that OBF1 extensively colocalizes with OCT1 and OCT2. We found that these factors also often colocalize with transcription factors of the ETS family. Furthermore, we showed that OBF1, OCT2 and OCT1 bind widely to the promoters or enhancers of genes involved in GC formation in mouse and human GC B cells. shRNA knockdown experiments demonstrated that OCT1, OCT2 and OBF1 regulate each other and are essential for proliferation of GC-derived lymphoma cell lines. OBF1 downregulation disrupts the GC transcriptional program: genes involved in GC maintenance -such as *BCL6*- are downregulated, while genes related to exit from the GC program -such as *IRF4*- are upregulated. Ectopic expression of BCL6 does not restore the proliferation of GC-derived lymphoma cells depleted of OBF1 unless IRF4 is also depleted, indicating that OBF1 controls an essential regulatory node in GC differentiation.

Conflict of interest: No COI declared

COI notes: B.D.F., F.W. and M.F. are employees and shareholders of Novartis.

Preprint server: No;

Author contributions and disclosures: S.S. and P.M. designed the project; S.S. performed all experiments, bioinformatics analysis and interpreted the data for manuscript under the supervision of M.A.C., M.S. and P.M.; S.S. and C.C. maintained all the mouse strains involved in this work; F.T. and G.C. contributed to the design of knockdown experiments; S.S., C.C. and H.K. performed all FACS analyses; S.S. performed all CUT&RUN experiments and sequencing library preparation under the supervision of F.W., B.F. and M.F.; S.D. collected human tonsil after tonsillectomy; S.N.W., J.J. and S.L.N. provided cultured Spi1KO/KO B cells; S.S. and P.M. wrote the manuscript and all authors contributed to the final version.

Non-author contributions and disclosures: No;

Agreement to Share Publication-Related Data and Data Sharing Statement: Further information and requests for resources and reagents should be directed to Patrick Matthias (patrick.matthias@fmi.ch). ChIP-seq and RNA-seq data are available at GEO under accession number GSE142040.

Clinical trial registration information (if any):

OBF1 and Oct factors control the germinal center transcriptional program

Shuang Song[1, 2], Chun Cao[1], Mohamed-Amin Choukrallah[1], Fengyuan Tang[3], Gerhard Christofori[3], Hubertus Kohler[1], Fabian Wu[4], Barna D. Fodor[4], Mathias Frederiksen[4], Simon N. Willis[5, 6], Jacob T. Jackson[5, 6], Stephen L. Nutt[5, 6], Stefan Dirnhofer[7], Michael B. Stadler[1, 2, 8] and Patrick Matthias[1, 2, *]

Affiliations:
[1] Friedrich Miescher Institute for Biomedical Research, 4058 Basel, Switzerland
[2] Faculty of Sciences, University of Basel, 4031 Basel, Switzerland
[3] Department of Biomedicine, University of Basel, 4031 Basel, Switzerland
[4] Novartis Institutes for Biomedical Research, 4056 Basel, Switzerland
[5] The Walter & Eliza Hall Institute of Medical Research, Victoria, 3050, Australia
[6] Department of Medical Biology, University of Melbourne, Parkville, VIC 3010, Australia
[7] University Hospital Basel, 4031 Basel, Switzerland
[8] Swiss Institute of Bioinformatics, 4058 Basel, Switzerland.

***Correspondence:**
Patrick Matthias, Friedrich Miescher Institute for Biomedical Research, Maulbeerstrasse 66, 4058 Basel, Switzerland, Phone: 41-61-697-6661, FAX: 41-61-697-3976, E-mail: patrick.matthias@fmi.ch

Running title: OBF1 controls germinal center identity

Text word count: 4331
Abstract word count: 175
Number of figures: 7 main figures with 6 supplemental figures
Number of references: 85

Key points:

1. In primary germinal center (GC) B cells, OBF1 bind to genes important for GCs, including the key regulators BCL6 and FOXO1.
2. OBF1 maintains proliferation of GC-derived lymphoma cells and is a master regulator controlling the GC program.

Abstract

OBF1 is a specific coactivator of the POU family transcription factors OCT1 and OCT2. OBF1 and OCT2 are B cell-specific and indispensable for germinal center (GC) formation, but their mechanism of action is unclear. Here, we show by ChIP-seq that OBF1 extensively colocalizes with OCT1 and OCT2. We found that these factors also often colocalize with transcription factors of the ETS family. Furthermore, we showed that OBF1, OCT2 and OCT1 bind widely to the promoters or enhancers of genes involved in GC formation in mouse and human GC B cells. shRNA knockdown experiments demonstrated that OCT1, OCT2 and OBF1 regulate each other and are essential for proliferation of GC-derived lymphoma cell lines. OBF1 downregulation disrupts the GC transcriptional program: genes involved in GC maintenance -such as *BCL6*- are downregulated, while genes related to exit from the GC program -such as *IRF4*- are upregulated. Ectopic expression of BCL6 does not restore the proliferation of GC-derived lymphoma cells depleted of OBF1 unless IRF4 is also depleted, indicating that OBF1 controls an essential regulatory node in GC differentiation.

Introduction

OCT2 and OCT1 are transcription factors (TFs) founding members of the POU family[1]. They bind to the octamer motif 5'-ATGCAAAT-3'[2], first identified as a conserved element in the regulatory regions of immunoglobulin (Ig) genes[3-6]. OBF1 (aka OCA-B/Bob.1) is a B cell-specific coactivator that interacts with OCT1 and/or OCT2 on cognate sites[7,8]. It enhances transcriptional activity of OCT1 and OCT2 on target genes[7,9-12]. Based on their B cell specificity, OCT2 and OBF1 have been thought to be critical factors for Ig expression[9,13,14]. However, mice lacking OCT2 or OBF1 showed largely normal Ig transcription and B cell development but demonstrated the importance of these factors for humoral immune responses and transcription of isotype-switched Igs[15]. OBF1 is essential for the formation of germinal centers (GCs)[15-19], dynamic structures that form in peripheral lymphoid organs in response to T cell-dependent antigens[20-22]. The role of OCT2 in GC formation is unclear: OCT2-deficient mice lacked GCs following NP-OVA immunization[15], but showed normal GCs formation after influenza virus or NP-KLH challenge[23,24]. However, in the latter cases GC B cells function was impaired[23,24].

During the GC reaction, B cells undergo rapid proliferation and Ig gene hypermutation, leading to the selection of B cells with increased affinity for specific antigens and differentiation into high-affinity antibody-secreting plasma cells (PCs) or memory B cells[20-22,25-27]. BCL6 is a master regulator of GCs, maintaining the GC transcriptional program and suppressing the expression of genes essential for further differentiation to PCs[27-31]. The majority of non-Hodgkin lymphomas (NHLs), including Burkitt's lymphoma (BL) and most diffuse large B-cell lymphoma (DLBCL)[32,33], originate from GCs. BLs and GCB-subtype DLBCLs closely resemble GC cells[34,35], therefore they serve as a model to study the transcriptional regulation of the GC reaction.

OCT2 and OBF1 are expressed throughout B cell development, but particularly highly in GC B cells and BL cells[36,37]. DLBCLs are sensitive to the loss of OBF1, which is controlled by a super-enhancer[38]; a screen identified OBF1 as an essential gene for a GC-derived BL cell line[39]. While the link between OCT/OBF1 and GCs is established, the underlying mechanism remains unclear.

Here, we generated genomic binding maps of OCT1, OCT2 and OBF1 in primary B cells, and investigated the functional relevance of these factors for GCs. We found that they extensively colocalize with each other and with ETS factors, and we showed that OBF1 stabilizes the binding of OCT factors on chromatin. Moreover, they bind to the regulatory elements of multiple genes essential for the GC reaction in mouse and human GC B cells. In GC-derived lymphoma cells, OBF1 is the master regulator that controls cellular proliferation by modulating IRF4 expression. Upon OBF1 depletion, the GC transcriptional program was dramatically disrupted, as manifested by the downregulation of GC reaction master regulators, including *BCL6* and *FOXO1*, and the de-repression of genes promoting post-GC differentiation, such as *IRF4* and *BCL2L1*. Thus, OBF1 controls the GC

transcriptional program integrity and represents a promising therapeutic target for GC-derived lymphoma cells.

Methods

Mice

All strains were maintained on the C57BL/6J background. Animal experiments were carried out according to valid project licences, approved and regularly controlled by the Swiss Cantonal Veterinary Office of Basel-Stadt. *Pou2af1*$^{KO/KO}$ mice have been described[16]. *Rosa26*$^{BirA/BirA}$ mouse line[40] was obtained from M. Busslinger (IMP, Vienna).

Antibodies

The antibodies used in this study are provided in Supplemental Table 1.

Cell Culture

293T cells were cultured in DMEM medium, 10% FBS (Sigma). Lymphoma cell lines were cultured in RPMI1640 medium (10% FBS, 50 µM 2-mercaptoethanol) and passaged every two to three days. Splenic B cells were stimulated with Lipopolysaccharide (LPS) or anti-CD40/Interleukin-4 (IL4). See Supplemental Methods for detailed stimulation methods.

Generation of transgenic mice

See Supplemental Methods for details.

MACS sorting of CD19^{+} mature B cells

Splenic mature B cells were purified by MACS using anti-mouse CD19 Microbeads (Miltenyi Biotec). See Supplemental Methods for details.

Gene knockdown experiment by shRNA

pLKO.1 lentiviral vectors with specific shRNAs for human *Pou2f1*, human *Pou2f2* and human *Pou2af1* were purchased from Sigma. shRNA sequences used in this paper are listed in Supplemental Table 5. See Supplemental Methods for details.

Immunoblotting

Cell lysates for immunoblotting were generated with RIPA buffer (50 mM Tris-HCl, pH7.4, 150 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% SDS). Lysates were separated on 4-12% SDS-polyacrylamide gels and transferred to PVDF membranes. Antibodies are listed in Supplemental Table 1.

RNA Extraction and qPCR

Total cellular RNA from 5×10^5 cells was extracted using RNeasy Micro kit (QIAGEN), complementary DNA was generated using ImProm-II Reverse Transcription System (Promega), and quantitative PCR performed using FastStart Universal SYBR Green Master (Roche) with gene specific primers; human *Actb* was used as a reference gene. The StepOne Real-Time PCR System (Applied Biosystems) was used to read quantitative PCR signals. Oligonucleotide sequences are specified in Supplemental Table 6.

Chromatin immunoprecipitation (ChIP), Bio-ChIP-seq

5×10^7 of in vitro cultured CD19^{+} splenic B cells stimulated with LPS or anti-CD40/IL4 were used for chromatin immunoprecipitation as described[41] with modifications, or for Bio-ChIP-seq. See Supplemental Methods for detailed experimental procedures.

ChIP-seq data processing

See Supplemental Methods for details.

RNA-seq and data processing

See Supplemental Methods for details.

Human GC B cell isolation

Human GC B cell populations were purified by FACS from human tonsillectomy specimens with chronic tonsillitis. Tonsils were collected and anonymized during standard surgical pathology work-up. All patients had given informed consent, in accordance with the Swiss Federal Act on Research involving Human Beings, art. 38 and in accordance with the declaration of Helsinki. GC B cells were sorted by flow cytometry based on surface expression of CD19, CD38 and IgD.

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to Patrick Matthias (patrick.matthias@fmi.ch). ChIP-seq and RNA-seq data are available at GEO under accession number GSE142040.

Results

OCT1, OCT2 and OBF1 show genome-wide colocalization

To elucidate the chromatin occupancy of OCT1 (*Pou2f1*), OCT2 (*Pou2f2*) and OBF1 (*Pou2af1*) in B cells, we generated transgenic mice in which these TFs are endogenously tagged with an AviTag-FLAG[42,43] (*Pou2f1*^{AviF} and *Pou2af1*^{AviF}) or an AviTag (*Pou2f2*^{Avi}; Figure S1A shows the targeting strategy). Intercrossing these mice with mice ubiquitously expressing the *Escherichia coli* biotin ligase BirA from the *Rosa26* locus[40] led to robust biotinylation of the tagged factors in B cells, which allowed their efficient streptavidin-mediated precipitation.

To determine the genome-wide binding patterns of OCT1, OCT2 and OBF1 (hereafter OCT1/OCT2/OBF1), we collected mouse CD19^{+} splenic B cells from the different strains, cultured them in presence of bacterial lipopolysaccharides (LPS), and then performed Bio-ChIP-seq[44]. We identified 7,822, 13,482 and 32,596 Bio-ChIP-seq peaks for OBF1/OCT1/OCT2, respectively, with strong enrichment of the signal over input (Figure S1B). Notably, the expression level of OCT2 was much higher than that of OCT1 (Figure S1F), which likely explains why more OCT2 than OCT1 peaks were identified. For OBF1, conventional ChIP with antibodies identified much fewer peaks with lower overall enrichment; however, all antibody-derived peaks were contained within the Bio-ChIP-seq peaks (Figure S1C and S1D).

We observed considerable overlap between the three factors: 95.7% of OCT1 peaks overlapped with OCT2 peaks, 89.4% of OBF1 peaks overlapped with OCT1 peaks, and 95.6% of OBF1 peaks overlapped with OCT2 peaks (Figure 1A-1C). ChIP-seq read counts were highly correlated between all three factors (Figure 1D-1F). Furthermore, the ChIP signals of all three factors were strongly concentrated in the center of a 5kb window around OCT1 peak summits (Figure 1G and 1H).

OCT1/OCT2/OBF1 occupy the promoters of known target genes, such as *Cd36*[45,46], *Gadd45a*[47], *Id3*[48], *Spib*[12] and *Syk*[49] (Figure 1I; Figure S1E). RNA-seq analysis with LPS-stimulated splenic B cells showed that genes bound by at least one factor had significant higher expression than those not bound (Figure 1J). Thus, OCT1/OCT2/OBF1 colocalize at genome-wide level and their binding generally correlates with higher gene expression levels.

OCT1/OCT2/OBF1 occupy active functional genomic elements

Previous studies have shown the association of OCT1/OCT2/OBF1 with individual gene promoters and enhancers[11,47,50,51]. Here, we found that most OBF1 and OCT1 peaks were located at promoter regions (±1kb of the transcription start site, TSS; Figure 2A and 2B). By contrast, most OCT2 peaks were located at regions distal to TSSs (> 1 kb away), and only 30.4% of OCT2 peaks mapped to promoters (Figure 2A and 2B).

To evaluate whether the proximal and distal binding sites are indeed active promoters or enhancers, we examined the level of histone H3 lysine 27 acetylation (H3K27ac), a histone modification representative of active regulatory elements[52,53]. H3K27ac ChIP-seq identified 46,732 H3K27ac-enriched regions both proximal and distal to TSSs (Figure 2B). H3K27ac-enriched loci highly overlapped with OCT1/OCT2/OBF1 binding sites, the vast majority of which were enriched for H3K27ac marks (Figure 2C). The colocalization of OCT1/OCT2/OBF1 on active promoters (marked by high H3K4me3, low H3K4me1 and high H3K27ac signal) and active enhancers (marked by high H3K4me1 and H3K27ac signal[54]) is illustrated in Figure 2D. Furthermore, the H3K27ac level is significantly higher on the promoters of genes bound by OCT1, OCT2 or OBF1 (Figure S1G) and these factors are mainly bound to active genes (Figure S1H). In summary, OCT1/OCT2/OBF1 binding mainly occurs within active regulatory elements, indicating a functional role in transcriptional activation.

OCT1/OCT2/OBF1 colocalize with TFs of the ETS family

OBF1 lacks a DNA-binding domain, and its only established recruiting partners are OCT1 and OCT2[7,55-59]. *De novo* motif discovery of OCT1/OCT2/OBF1 peak regions identified among the top enriched motifs the sequence 5'-ATGCAAAT-3'[60,61] that matches the octamer consensus motif (Figure 3A). The second most frequent motif identified corresponded to DNA sequences bound by ETS factors, with the core sequence 5'-GGAA-3' (Figure 3A). PU.1, encoded by the *Spi1* gene, is a key ETS factor involved in cell fate decisions between myeloid and lymphoid lineages[62-65]. To understand the relative distribution of these factors, we determined the location of PU.1 and octamer motifs in the OCT1/OCT2/OBF1 peak regions: both types of motifs were most frequently found around peak summits (Figure 3B and 3C). The occurrences of PU.1 motifs at different distances around octamer sites within peak regions were determined (Figure 3D) and showed that PU.1 binding sites were preferentially found close to the octamer motifs, with the frequency peaking at a distance of about 50bp (Figure 3D, black lines); shuffled PU.1 motifs showed significantly lower enrichments (Figure 3D, grey lines, P=0.001996).

We next performed PU.1 ChIP-seq and found that OCT1/OCT2/OBF1 extensively colocalized with PU.1 (Figure 3E and S2A); ETS1 also colocalized with OCT1/OCT2/OBF1 in public ETS1 ChIP-seq data[66] (Figure S2B and S2C). We then systematically analyzed co-occurrences of binding motifs for all ETS factors which, based on our RNA-seq data, are expressed in B cells. Several of these ETS factors showed highly significant co-occurrence with octamer binding sites present in OBF1 peaks (Figure S2D), indicating that, besides PU.1 and ETS1, additional ETS factors may colocalize with OCT1/OCT2/OBF1. To test whether OBF1 and PU.1 regulate genes in common, we performed RNA-seq analysis with splenic B cells isolated from *Spi1*$^{KO/KO}$ or *Pou2af1*$^{KO/KO}$ knockout mice. However, only

minor overlap was identified among the differentially expressed genes (DEGs) (Figure 3F-3H). Together, our results indicate that OCT1/OCT2/OBF1 generally tend to colocalize with ETS factors in B cells, but the functional consequence is unclear.

OBF1 stabilizes the binding of OCT1 to chromatin

OBF1 has no independent DNA-binding capability and requires recruitment by Oct factors to associate with target sites[11,13,67]. Thereby, it might help to stabilize Oct factors on chromatin. We observed a strong positive correlation between OCT1/2 and OBF1 binding (Figure 4A; Figure S3A): OCT1 or OCT2 peaks were significantly more enriched (signal/input) when colocalized with OBF1 (Figure 4B and 4C). By contrast, the enrichment of Oct factor peaks was not dependent on colocalization with PU.1 (Figure 4D; Figure S3B). To confirm that OBF1 stabilizes the binding of Oct factors on chromatin, we generated *Pou2af1*$^{KO/KO}$ mice expressing AviTag-tagged OCT1 (*Pou2f1*$^{AviF/AviF}$*Rosa26*$^{BirA/BirA}$*Pou2af1*$^{KO/KO}$, Figure 4E) and isolated CD19^{+} splenic B cells. We examined by immunoblotting the expression level of OCT1 and OCT2 in these samples; while OCT1 level was unchanged, the OCT2 level was strongly dependent on OBF1 (Figure S3C and see below). 72 hours post LPS stimulation, we performed Bio-ChIP-seq for OCT1. We observed a general decrease in OCT1 peak enrichment among the regions bound by both OCT1 and OBF1 in WT B cells (Figure 4F), and 739 peaks (ca. 10.3% of total overlapping peaks) had significantly lower enrichment (WT/OBF1KO>2.5-fold; FDR<0.05) in *Pou2af1*$^{KO/KO}$ than in WT cells (Figure 4F and 4G). Only 73 peaks showed stronger binding signal in *Pou2af1*$^{KO/KO}$ cells (Figure 4F). We assigned the significantly differentially enriched peaks to the nearest genes and performed gene ontology (GO) analysis: the genes identified are involved in leukocyte activation/differentiation, immune system development and immune response (Figure 4G). The H3K27ac level was significantly reduced at the 739 peak regions when OBF1 was ablated (Figure 4H and 4I), suggesting that reduced OCT1 binding leads to a reduction in H3K27ac level. Our results demonstrate that OBF1 stabilizes OCT1 binding, which helps to regulate the H3K27ac levels of target loci.

OBF1 regulates the GC transcriptional program

To understand the roles of OCT1/OCT2/OBF1 in GC formation, we stimulated B cells with anti-CD40/IL4, a T cell-dependent stimulation mimicking the GC reaction, and performed RNA-seq and Bio-ChIP-seq. LPS, a T cell-independent signal, was used as a comparison. As expected, GC reaction-related genes were upregulated after anti-CD40/IL4 stimulation (Figure 5A). OBF1 showed enhanced binding to the promoters of multiple GC-associated genes (Figure 5B), suggesting that it regulates GCs by activating the expression of genes associated with this pathway.

De novo motif analysis of regions differentially bound by OBF1 under anti-CD40/IL4 showed that motifs for GC-related TFs were highly enriched (Figure S4A). We observed similar patterns for OCT1 and OCT2 (Figure S4B and S4C). OBF1, OCT1 and OCT2 bound to multiple genes encoding TFs important for the initiation phase of GCs (Figure 5C; Figure S4D). After anti-CD40/IL4 stimulation, OBF1 showed enhanced binding to several key GC-related genes, e.g., *Bcl6*[68], *Myc, Foxo1, Mef2b and Spi1*[27,69] (Figure 5D; Figure S4E).

RNA-seq analysis of anti-CD40/IL4-stimulated wild type (WT) and *Pou2af1*$^{ko/ko}$ B cells identified 746 DEGs (Figure 5E) which are mainly involved in immune responses (Figure 5F). Gene set enrichment analysis (GSEA) showed that GC signatures are dramatically disrupted when OBF1 is deleted, while a number of genes associated with PC differentiation (e.g. *Slamf7*, *Tlr7*, *Cd38*, *Stat1* and *Ccr2*) are upregulated (Figure 5G and 5H). Seven days after sheep red blood cells (SRBCs) immunization of mice, GCs formed robustly in WT spleens, and GC B cells (B220^{+}CD43^{-}CD38^{-}CD95^{+}) could be detected by FACS. In contrast, GC B cells were absent from *Pou2af1*$^{ko/ko}$ spleens (Figure S4F and 5I). Bio-ChIP-seq of OBF1 in primary murine GC B cells identified OBF1 binding to the regulatory regions of hallmark genes for GCs, such as *Bcl6*, *Foxo1*, *Aicda*, *Mef2b*, *Bach2* and *Egr3* (Figure 5J). These results indicate that OBF1 directly regulates the GC transcriptional program and is a master regulator of GCs.

OBF1 controls the proliferation of GC-derived B lymphoma cell by modulating the level of IRF4

To test the roles of OBF1/OCT1/OCT2 in the context of GCs, these factors were knocked down in Raji, a GC-derived human BL cell line, using shRNAs. The proliferation of Raji cells was strongly reduced upon knockdown of any of these factors, but OBF1 depletion elicited the strongest effect (Figure 6A-6C). Impaired proliferation upon OBF1 knockdown was also observed in other BL and DLBCL lines, such as Ramos, Daudi, HT and SUDHL4 (Figure S5A-S5C and S5I). Interestingly, OBF1 expression was downregulated when OCT1 or OCT2 were depleted (Figure S5D-S5F). Ectopic expression of the nuclear OBF1 isoform in these cells fully rescued proliferation, demonstrating that the defect observed upon OCT1 or OCT2 depletion is due to simultaneous downregulation of OBF1 (Figure 6D and 6E).

RNA-seq analysis of OBF1 knockdown Raji cells (discussed in detail below) showed that BCL6, which is essential to organize the transcriptome of GC B cells and the GC reaction[31], was strongly downregulated (Figure 6F; Figure 7B). Ectopically expressed BCL6 in OBF1-depleted Raji cells failed to restore proliferation (Figure S5G), indicating that alternative targets of OBF1, other than BCL6, are important. We wondered whether genes governing the post-GC differentiation (i.e., exit from GC program) are deregulated when OBF1 is ablated. IRF4 and BLIMP1 (encoded by the *PRDM1* gene) are master regulators controlling GC B to PC differentiation[70-72]. In absence of OBF1, *PRDM1* fails to

be upregulated, thus impeding the final stages of antibody-secreting cell development[73]. Consistently, *PRDM1* was downregulated upon OBF1 knockdown in our RNA-seq data (Figure 7B). *IRF4* expression was barely detectable in untreated Raji cells; however, it was elevated over ten times upon OBF1 downregulation (Figure 6G; Figure S5H; Figure 7B). Increased IRF4 impairs proliferation of BL cells and induces their differentiation towards PCs[74]. Indeed, Raji proliferation was restored when knocking down these two factors simultaneously (Figure 6G-6H). We observed the same outcome with SUDHL4 cells (Figure S5H and S5I), a DLBCL. This result therefore shows that OBF1 regulates lymphoma proliferation by controlling IRF4 expression. Moreover, ectopic IRF4 expression leads to impaired Raji cells proliferation (Figure 6I), a result which phenocopies the OBF1 knockdown. In conclusion, OCT1/OCT2/OBF1 are all required to maintain normal proliferation of lymphoma cells: OCT1 and OCT2 control cell proliferation by regulating the expression of OBF1, which itself is essential for controlling the expression of BCL6 and particularly IRF4.

OBF1 is essential for maintenance of the GC transcriptional program

To understand the role of OBF1 in the GC reaction, we performed RNA-seq with WT and OBF1 knockdown Raji cells, identifying 1175 DEGs (Figure 7A). Multiple genes related to GC formation or maintenance were downregulated, including *BCOR*, *BCL6*, *MEF2B* and *FOXO1*. *CXCR4*, which encodes a chemokine receptor important for GCs, was also downregulated. Genes important for somatic hypermutation or DNA repair were downregulated as well, including *AICDA* (encoding activation induced deaminase, AID), *GADD45A*, *BACH1*, *CEBPG*, *BOD1l1*, *HMGB2* and *RRM2B*. By contrast, genes involved in post-GC differentiation were upregulated, including *IRF4*, *BCL2L1*, *NFKB1*, *NFKB2* and *ZBTB32* (Figure 7B).

GO analysis showed that genes associated with protein secretion were enriched upon OBF1 depletion (Figure 7A), indicating that some aspects of PC differentiation are triggered, in agreement with what was observed in anti-CD40/IL4-stimulated *Pou2af1*$^{KO/KO}$ mouse B cells (Figure 5G). Furthermore, signatures associated with PC differentiation were upregulated in OBF1-depleted Raji cells (Figure 7C), and genes activated by IRF4 or repressed by BCL6 were upregulated (Figure S6A). These results demonstrate that, in the absence of OBF1, the GC transcriptional program collapses and the PC differentiation program partially initiates. We also found that the transcriptomes of Raji cells with depleted OBF1 and enforced IRF4 expression are highly similar; overall, genes significantly up- or downregulated upon OBF1 depletion also show stronger or weaker expression in Raji cells overexpressing IRF4, respectively (Figure 7D and 7E).

Next, we assessed the genomic binding of OBF1 in human lymphomas using CUT&RUN assays. OBF1 directly binds to the *IRF4* and *BCL6* gene loci in both BL (Raji) and DLBCL cells (SUDHL4), thus

explaining its aforementioned effects on the transcription of these genes (Figure 7F; Figure S6B). To examine OBF1 binding under normal physiological conditions, we performed CUT&RUN for OBF1 in primary human GC B cells, obtained from tonsillectomy patients. As shown in Figure 7G and S6C, OBF1 binds at the enhancers or promoters of genes that are critical for the GC reaction, such as *BCL6* and *IRF4* among others. Moreover, gene expression signatures associated with favorable prognosis outcomes for lymphoma patients[76] are upregulated in OBF1-depleted Raji cells (Figure 7H). Together, these findings further confirm our conclusion that OBF1 maintains the GC transcriptome by activating BCL6 and repressing IRF4. Once OBF1 expression is abrogated, the BCL6 pathway is attenuated, IRF4 expression is elevated, and the GC program collapses while PC differentiation initiates.

Discussion

OBF1 stabilizes the binding of OCT factors

Bio-ChIP-seq revealed the genomic location of OCT1/OCT2/OBF1 to a much greater depth than possible with antibodies (Figure S1C and S1D). Previous antibody-based ChIP-seq studies examined the binding of OCT2 only[75] or OCT2 and OBF1[24] in B cells, without appreciating the role of OCT1. We found extensive overlap between OCT1 and OCT2 peaks, and more than 90% of OBF1 peaks overlapped with either OCT1 or OCT2 peaks, consistent with the notion that this cofactor is recruited by OCT factors. Motif analysis of the different fractions (e.g. bound by OBF1 and OCT1, or by OCT1 alone) did not reveal significant differences and the consensus motif for POU domains 5'-ATGCAAAT-3' or variants thereof were identified at the top of the list. Previous studies performed on selected binding sites had concluded that OBF1 does not or only weakly stabilizes the binding of OCT factors on the DNA[11,13,76]. Comparison of OCT1 Bio-ChIP-seq peaks in WT vs OBF1 KO cells conclusively established the stabilization of OCT1 by OBF1. OCT2 could not be examined in a similar manner, as its expression is OBF1-dependent; however, given the high similarity between OCT1 and OCT2, stabilization of OCT2 appears likely. By contrast, PU.1, which often co-localized with OCT1 and OCT2, does not stabilize their binding. Cells lacking OBF1 exhibited reduced H3K27ac levels over the OCT1 peaks compared to WT cells, indicating that OBF1-depletion also affects the structure of chromatin.

Extensive colocalization between Oct factors and ETS factors

De novo motif analysis of the OCT1/OCT2/OBF1 peaks identified, in addition to octamer-related motifs, a high enrichment of ETS motifs. In particular, a PU.1 ChIP-seq study in mature B cells found octamer-like motifs near the PU.1 motifs[75]. We found by ChIP-seq that PU.1 colocalizes with a large

fraction of the OCT1/OCT2/OBF1 peaks. Prompted by these observations, we compared the transcriptome of B cells from *Spi1*$^{KO/KO}$ [77] and *Pou2af1*$^{KO/KO}$ mice. However, we found only little overlap in the deregulated genes, indicating that these two factors, although often co-localizing, are engaged mostly in separate regulatory pathways. Mapping the motifs of all ETS factors expressed in B cells showed an enrichment for several ETS motifs within the OBF1 peak regions; in addition to PU.1 and ETS1, other ETS factors such as ELK4 or ETV6 were also highly significant statistically (Figure S2D). Using publicly available ETS1 ChIP-seq data we found that colocalization is even greater than with PU.1 (Figure S2B). Thus, OCT1/OCT2/OBF1 colocalize with several ETS factors in B cells and the extent of overlap between these TF networks had not been appreciated. While the functional relevance of these observations will require further systematic analysis, given the number of different ETS factors involved in co-localization with OCT1/OCT2/OBF1, it seems likely that in some cases these factors will be co-regulating their targets.

OBF1 is essential for maintenance of the GC program and post-GC differentiation

BCL6 is a master regulator of the GC reaction[31], and *Bcl6*$^{KO/KO}$ mice lack GCs[78]. Similarly, GCs are missing in *Pou2af1*$^{KO/KO}$ mice[15-19]. These highly similar phenotypes suggested a possible regulatory relationship between these two factors. OBF1 and BCL6 are highly co-expressed in dark zone GC B cells and show attenuated expression in the light zone[36,79]. IRF4 expression is induced by CD40 signaling in light zone B cells and promotes PC differentiation[27,79,80]. The expression of IRF4 is essential to induce OBF1 and BCL6 at pre-GC stage, and repression of IRF4 by BCL6 allows the GC reaction to start[81]. Together with our data, this indicates that OBF1 and IRF4 expression are mutually exclusive, and that OBF1 is the key factor regulating the switch between GC reaction and further differentiation. More recently, OBF1 has been reported to regulate BCL6 in CD4$^+$ T_{FH} cells[82] and we showed here that it directly regulates the *BCL6* promoter in primary human and mouse GC B cells. Depletion of OBF1 in GC-derived lymphoma cells leads to reduced BCL6 expression and impaired proliferation. In these OBF1-depleted cells, GSEA analysis showed that GC signatures are lost, while PC differentiation signatures are upregulated. Consistently, IRF4 is strongly upregulated when OBF1 is depleted.

The functional role of OBF1 on BCL6 and IRF4 expression seems to be GC-restricted. Indeed, OBF1 and IRF4 are co-expressed in PCs, where both factors are critical for PC differentiation[36,70,73,81]. In PCs, IRF4 and BLIMP1 activate XBP1, which in turn induces high OBF1 expression[27,83]. Therefore, the coexistence of OBF1 and IRF4 in PCs indicates that OBF1 regulates IRF4 specifically in the GC context: once GC program integrity is lost and the cells have adopted a PC signature, OBF1 is no longer able to regulate IRF4. Likewise, BCL6 is regulated by OBF1 but only in GC B cells, because BCL6 and IRF4

expression was found to be normal in *Pou2af1*$^{KO/KO}$ splenic cells stimulated in vitro [84] (and our data not shown).

OBF1 as a therapeutic target for GC-derived B cell lymphoma

A CRISPR/Cas9 screen identified OBF1 as essential for BL cell lines[39] and it is required for maintaining proliferation of GC-derived lymphoma cells[24,38]. However, the functional roles of OCT1 and OCT2, and the functional hierarchy of OCT1/OCT2/OBF1 in B cell lymphomas remained unclear. We found that the proliferation of B cell lymphoma cells is also strongly impaired when OCT1 or OCT2 is downregulated (Figure 6B-6C); however, this leads to a reduction in OBF1 and proliferation is fully rescued when OBF1 expression is restored in these cells. Therefore, OBF1 is the factor critical for the fast proliferation of B cell lymphomas. Recent studies provide possible functional mechanisms for the importance of OBF1 for lymphoma proliferation. JQ1 blocks BRD4 binding to acetylated chromatin, which results in reduced activity of enhancers or promoters[85]. DLBCL proliferation is impaired following JQ1 administration, as it abolishes the activity of a super-enhancer controling OBF1 expression[38]. Moreover, functional importance has been ascribed to the OCT2-OBF1 interaction interface[24], yet the specific downstream target(s) had not been identified. We demonstrated here that *BCL6* and *IRF4* are critical targets of OBF1 in BL and DLBCL. Strikingly, ectopic expression of BCL6 failed to rescue the proliferation of OBF1-depleted B lymphoma cells (Figure S5G). IRF4, a negative regulator of BL proliferation[74], was upregulated when OBF1 is abrogated. Simultaneous knockdown of IRF4 restored the proliferation of OBF1-depleted GC-derived B lymphoma cells. Moreover, OBF1 binds to the majority of GC-related genes, particularly the *IRF4* and *BCL6* loci (Figure 7F-G). In sum, we demonstrate that OBF1 facilitates the fast proliferation of GC-derived lymphoma cells by repressing IRF4. In addition, lymphoma cells with reduced OBF1 level adopt signatures associated with favourable prognosis, which suggests that OBF1 could serve as a valuable clinical indicator for lymphoma severity classification and prognosis. Therefore, we propose OBF1 to be a novel and potent therapeutic target for lymphoma treatment.

Acknowledgements

We are grateful to M. Busslinger (IMP, Vienna) for *Rosa26*$^{BirA/BirA}$ mice, S. Junker (Aarhus University) for Ramos cell line, and J. Chao (FMI, Basel) for lentiviral expression plasmids. We thank S. Smallwood and the genomics facility team for help with library preparation and next generation sequencing of RNA-seq and ChIP-seq samples, J.-F. Spetz for zygote microinjection, H.-R. Hotz for advice with RNA-seq analysis, J. Wilbertz for advice on lentivirus production, Y. Zhan for comments on the manuscript, D. Finke and G. Aurey (DBM, Basel) for help with FACS sorting of human GC B cells. We also thank L. Wang, G. Matthias, D. Mayer, M. Flemr, P. Knuckles, T. Kitazawa, P. Bammer, G. Fanourgakis and all the Matthias laboratory members for their helpful discussions and advices, and Life Science Editors for editing services. This work was supported by the Novartis Research Foundation.

Author Contributions

S.S. and P.M. designed the project; S.S. performed all experiments, bioinformatics analysis and interpreted the data for manuscript under the supervision of M.A.C., M.S. and P.M.; S.S. and C.C. maintained all the mouse strains involved in this work; F.T. and G.C. contributed to the design of knockdown experiments; S.S., C.C. and H.K. performed all FACS analyses; S.S. performed all CUT&RUN experiments and sequencing library preparation under the supervision of F.W., B.F. and M.F.; S.D. collected human tonsil after tonsillectomy; S.N.W., J.J. and S.L.N. provided cultured *Spi1*$^{KO/KO}$ B cells; S.S. and P.M. wrote the manuscript and all authors contributed to the final version.

Disclosure of Conflicts of Interest

B.D.F., F.W. and M.F. are employees and shareholders of Novartis.

References:

1. Herr W, Cleary MA. The POU domain: versatility in transcriptional regulation by a flexible two-in-one DNA-binding domain. *Genes Dev*. 1995;9(14):1679-1693.
2. Staudt LM, Clerc RG, Singh H, LeBowitz JH, Sharp PA, Baltimore D. Cloning of a lymphoid-specific cDNA encoding a protein binding the regulatory octamer DNA motif. *Science*. 1988;241(4865):577-580.
3. Staudt LM, Lenardo MJ. Immunoglobulin gene transcription. *Annu Rev Immunol*. 1991;9:373-398.
4. Parslow TG, Blair DL, Murphy WJ, Granner DK. Structure of the 5' ends of immunoglobulin genes: a novel conserved sequence. *Proc Natl Acad Sci U S A*. 1984;81(9):2650-2654.
5. Falkner FG, Zachau HG. Correct transcription of an immunoglobulin kappa gene requires an upstream fragment containing conserved sequence elements. *Nature*. 1984;310(5972):71-74.
6. Bergman Y, Rice D, Grosschedl R, Baltimore D. Two regulatory elements for immunoglobulin kappa light chain gene expression. *Proc Natl Acad Sci U S A*. 1984;81(22):7041-7045.
7. Tomilin A, Remenyi A, Lins K, et al. Synergism with the coactivator OBF-1 (OCA-B, BOB-1) is mediated by a specific POU dimer configuration. *Cell*. 2000;103(6):853-864.
8. Luo Y, Fujii H, Gerster T, Roeder RG. A novel B cell-derived coactivator potentiates the activation of immunoglobulin promoters by octamer-binding transcription factors. *Cell*. 1992;71(2):231-241.
9. Luo Y, Roeder RG. Cloning, functional characterization, and mechanism of action of the B-cell-specific transcriptional coactivator OCA-B. *Mol Cell Biol*. 1995;15(8):4115-4124.
10. Ren X, Siegel R, Kim U, Roeder RG. Direct interactions of OCA-B and TFII-I regulate immunoglobulin heavy-chain gene transcription by facilitating enhancer-promoter communication. *Mol Cell*. 2011;42(3):342-355.
11. Lins K, Remenyi A, Tomilin A, et al. OBF1 enhances transcriptional potential of Oct1. *EMBO J*. 2003;22(9):2188-2198.
12. Bartholdy B, Du Roure C, Bordon A, Emslie D, Corcoran LM, Matthias P. The Ets factor Spi-B is a direct critical target of the coactivator OBF-1. *Proc Natl Acad Sci U S A*. 2006;103(31):11665-11670.
13. Strubin M, Newell JW, Matthias P. OBF-1, a novel B cell-specific coactivator that stimulates immunoglobulin promoter activity through association with octamer-binding proteins. *Cell*. 1995;80(3):497-506.
14. Muller MM, Ruppert S, Schaffner W, Matthias P. A cloned octamer transcription factor stimulates transcription from lymphoid-specific promoters in non-B cells. *Nature*. 1988;336(6199):544-551.
15. Schubart K, Massa S, Schubart D, Corcoran LM, Rolink AG, Matthias P. B cell development and immunoglobulin gene transcription in the absence of Oct-2 and OBF-1. *Nat Immunol*. 2001;2(1):69-74.
16. Schubart DB, Rolink A, Kosco-Vilbois MH, Botteri F, Matthias P. B-cell-specific coactivator OBF-1/OCA-B/Bob1 required for immune response and germinal centre formation. *Nature*. 1996;383(6600):538-542.
17. Kim U, Qin XF, Gong S, et al. The B-cell-specific transcription coactivator OCA-B/OBF-1/Bob-1 is essential for normal production of immunoglobulin isotypes. *Nature*. 1996;383(6600):542-547.

18. Nielsen PJ, Georgiev O, Lorenz B, Schaffner W. B lymphocytes are impaired in mice lacking the transcriptional co-activator Bob1/OCA-B/OBF1. *Eur J Immunol*. 1996;26(12):3214-3218.
19. Fehr T, Lopez-Macias C, Odermatt B, et al. Correlation of anti-viral B cell responses and splenic morphology with expression of B cell-specific molecules. *Int Immunol*. 2000;12(9):1275-1284.
20. Allen CD, Okada T, Cyster JG. Germinal-center organization and cellular dynamics. *Immunity*. 2007;27(2):190-202.
21. De Silva NS, Klein U. Dynamics of B cells in germinal centres. *Nat Rev Immunol*. 2015;15(3):137-148.
22. Klein U, Dalla-Favera R. Germinal centres: role in B-cell physiology and malignancy. *Nat Rev Immunol*. 2008;8(1):22-33.
23. Karnowski A, Chevrier S, Belz GT, et al. B and T cells collaborate in antiviral responses via IL-6, IL-21, and transcriptional activator and coactivator, Oct2 and OBF-1. *J Exp Med*. 2012;209(11):2049-2064.
24. Hodson DJ, Shaffer AL, Xiao W, et al. Regulation of normal B-cell differentiation and malignant B-cell survival by OCT2. *Proc Natl Acad Sci U S A*. 2016;113(14):E2039-2046.
25. Basso K, Dalla-Favera R. Germinal centres and B cell lymphomagenesis. *Nat Rev Immunol*. 2015;15(3):172-184.
26. Victora GD, Nussenzweig MC. Germinal centers. *Annu Rev Immunol*. 2012;30:429-457.
27. Song S, Matthias PD. The Transcriptional Regulation of Germinal Center Formation. *Front Immunol*. 2018;9:2026.
28. Shaffer AL, Yu X, He Y, Boldrick J, Chan EP, Staudt LM. BCL-6 represses genes that function in lymphocyte differentiation, inflammation, and cell cycle control. *Immunity*. 2000;13(2):199-212.
29. Tunyaplin C, Shaffer AL, Angelin-Duclos CD, Yu X, Staudt LM, Calame KL. Direct repression of prdm1 by Bcl-6 inhibits plasmacytic differentiation. *J Immunol*. 2004;173(2):1158-1165.
30. Shaffer AL, Lin KI, Kuo TC, et al. Blimp-1 orchestrates plasma cell differentiation by extinguishing the mature B cell gene expression program. *Immunity*. 2002;17(1):51-62.
31. Hatzi K, Melnick A. Breaking bad in the germinal center: how deregulation of BCL6 contributes to lymphomagenesis. *Trends Mol Med*. 2014;20(6):343-352.
32. Shaffer AL, 3rd, Young RM, Staudt LM. Pathogenesis of human B cell lymphomas. *Annu Rev Immunol*. 2012;30:565-610.
33. Mlynarczyk C, Fontan L, Melnick A. Germinal center-derived lymphomas: The darkest side of humoral immunity. *Immunol Rev*. 2019;288(1):214-239.
34. Wright G, Tan B, Rosenwald A, Hurt EH, Wiestner A, Staudt LM. A gene expression-based method to diagnose clinically distinct subgroups of diffuse large B cell lymphoma. *Proc Natl Acad Sci U S A*. 2003;100(17):9991-9996.
35. Piccaluga PP, De Falco G, Kustagi M, et al. Gene expression analysis uncovers similarity and differences among Burkitt lymphoma subtypes. *Blood*. 2011;117(13):3596-3608.
36. Greiner A, Muller KB, Hess J, Pfeffer K, Muller-Hermelink HK, Wirth T. Up-regulation of BOB.1/OBF.1 expression in normal germinal center B cells and germinal center-derived lymphomas. *Am J Pathol*. 2000;156(2):501-507.
37. Saez AI, Artiga MJ, Sanchez-Beato M, et al. Analysis of octamer-binding transcription factors Oct2 and Oct1 and their coactivator BOB.1/OBF.1 in lymphomas. *Mod Pathol*. 2002;15(3):211-220.

38. Chapuy B, McKeown MR, Lin CY, et al. Discovery and characterization of super-enhancer-associated dependencies in diffuse large B cell lymphoma. *Cancer Cell*. 2013;24(6):777-790.
39. Wang T, Birsoy K, Hughes NW, et al. Identification and characterization of essential genes in the human genome. *Science*. 2015;350(6264):1096-1101.
40. Driegen S, Ferreira R, van Zon A, et al. A generic tool for biotinylation of tagged proteins in transgenic mice. *Transgenic Res*. 2005;14(4):477-482.
41. Choukrallah MA, Song S, Rolink AG, Burger L, Matthias P. Enhancer repertoires are reshaped independently of early priming and heterochromatin dynamics during B cell differentiation. *Nat Commun*. 2015;6:8324.
42. Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F. Genome engineering using the CRISPR-Cas9 system. *Nat Protoc*. 2013;8(11):2281-2308.
43. Yang H, Wang H, Jaenisch R. Generating genetically modified mice using CRISPR/Cas-mediated genome engineering. *Nat Protoc*. 2014;9(8):1956-1968.
44. Revilla IDR, Bilic I, Vilagos B, et al. The B-cell identity factor Pax5 regulates distinct transcriptional programmes in early and late B lymphopoiesis. *EMBO J*. 2012;31(14):3130-3146.
45. Konig H, Pfisterer P, Corcoran LM, Wirth T. Identification of CD36 as the first gene dependent on the B-cell differentiation factor Oct-2. *Genes Dev*. 1995;9(13):1598-1607.
46. Shore P, Dietrich W, Corcoran LM. Oct-2 regulates CD36 gene expression via a consensus octamer, which excludes the co-activator OBF-1. *Nucleic Acids Res*. 2002;30(8):1767-1773.
47. Takahashi S, Saito S, Ohtani N, Sakai T. Involvement of the Oct-1 regulatory element of the gadd45 promoter in the p53-independent response to ultraviolet irradiation. *Cancer Res*. 2001;61(3):1187-1195.
48. Bordon A, Bosco N, Du Roure C, et al. Enforced expression of the transcriptional coactivator OBF1 impairs B cell differentiation at the earliest stage of development. *PLoS One*. 2008;3(12):e4007.
49. Corcoran L, Emslie D, Kratina T, et al. Oct2 and Obf1 as Facilitators of B:T Cell Collaboration during a Humoral Immune Response. *Front Immunol*. 2014;5:108.
50. Reece-Hoyes JS, Keenan ID, Pownall ME, Isaacs HV. A consensus Oct1 binding site is required for the activity of the Xenopus Cdx4 promoter. *Dev Biol*. 2005;282(2):509-523.
51. Annweiler A, Muller-Immergluck M, Wirth T. Oct2 transactivation from a remote enhancer position requires a B-cell-restricted activity. *Mol Cell Biol*. 1992;12(7):3107-3116.
52. Creyghton MP, Cheng AW, Welstead GG, et al. Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc Natl Acad Sci U S A*. 2010;107(50):21931-21936.
53. Rada-Iglesias A, Bajpai R, Swigut T, Brugmann SA, Flynn RA, Wysocka J. A unique chromatin signature uncovers early developmental enhancers in humans. *Nature*. 2011;470(7333):279-283.
54. Heintzman ND, Stuart RK, Hon G, et al. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat Genet*. 2007;39(3):311-318.
55. Sauter P, Matthias P. Coactivator OBF-1 makes selective contacts with both the POU-specific domain and the POU homeodomain and acts as a molecular clamp on DNA. *Mol Cell Biol*. 1998;18(12):7397-7409.
56. Remenyi A, Tomilin A, Pohl E, et al. Differential dimer activities of the transcription factor Oct-1 by DNA-induced interface swapping. *Mol Cell*. 2001;8(3):569-580.
57. Jerabek S, Ng CK, Wu G, et al. Changing POU dimerization preferences converts Oct6 into a pluripotency inducer. *EMBO Rep*. 2017;18(2):319-333.

58. Pfisterer P, Zwilling S, Hess J, Wirth T. Functional characterization of the murine homolog of the B cell-specific coactivator BOB.1/OBF.1. *J Biol Chem*. 1995;270(50):29870-29880.
59. Luo Y, Ge H, Stevens S, Xiao H, Roeder RG. Coactivation by OCA-B: definition of critical regions and synergism with general cofactors. *Mol Cell Biol*. 1998;18(7):3803-3810.
60. Staudt LM, Singh H, Sen R, Wirth T, Sharp PA, Baltimore D. A lymphoid-specific protein binding to the octamer motif of immunoglobulin genes. *Nature*. 1986;323(6089):640-643.
61. Scholer HR, Balling R, Hatzopoulos AK, Suzuki N, Gruss P. Octamer binding proteins confer transcriptional activity in early mouse embryogenesis. *EMBO J*. 1989;8(9):2551-2557.
62. Pang SHM, de Graaf CA, Hilton DJ, et al. PU.1 Is Required for the Developmental Progression of Multipotent Progenitors to Common Lymphoid Progenitors. *Front Immunol*. 2018;9:1264.
63. Dakic A, Wu L, Nutt SL. Is PU.1 a dosage-sensitive regulator of haemopoietic lineage commitment and leukaemogenesis? *Trends Immunol*. 2007;28(3):108-114.
64. DeKoter RP, Lee HJ, Singh H. PU.1 regulates expression of the interleukin-7 receptor in lymphoid progenitors. *Immunity*. 2002;16(2):297-309.
65. DeKoter RP, Singh H. Regulation of B lymphocyte and macrophage development by graded expression of PU.1. *Science*. 2000;288(5470):1439-1441.
66. Saelee P, Kearly A, Nutt SL, Garrett-Sinha LA. Genome-Wide Identification of Target Genes for the Key B Cell Transcription Factor Ets1. *Front Immunol*. 2017;8:383.
67. Chasman D, Cepek K, Sharp PA, Pabo CO. Crystal structure of an OCA-B peptide bound to an Oct-1 POU domain/octamer DNA complex: specific recognition of a protein-DNA interface. *Genes Dev*. 1999;13(20):2650-2657.
68. Kitano M, Moriyama S, Ando Y, et al. Bcl6 protein expression shapes pre-germinal center B cell dynamics and follicular helper T cell heterogeneity. *Immunity*. 2011;34(6):961-972.
69. Dominguez-Sola D, Victora GD, Ying CY, et al. The proto-oncogene MYC is required for selection in the germinal center and cyclic reentry. *Nat Immunol*. 2012;13(11):1083-1091.
70. Klein U, Casola S, Cattoretti G, et al. Transcription factor IRF4 controls plasma cell differentiation and class-switch recombination. *Nat Immunol*. 2006;7(7):773-782.
71. Shapiro-Shelef M, Lin KI, McHeyzer-Williams LJ, Liao J, McHeyzer-Williams MG, Calame K. Blimp-1 is required for the formation of immunoglobulin secreting plasma cells and pre-plasma memory B cells. *Immunity*. 2003;19(4):607-620.
72. Kallies A, Hasbold J, Tarlinton DM, et al. Plasma cell ontogeny defined by quantitative changes in blimp-1 expression. *J Exp Med*. 2004;200(8):967-977.
73. Corcoran LM, Hasbold J, Dietrich W, et al. Differential requirement for OBF-1 during antibody-secreting cell differentiation. *J Exp Med*. 2005;201(9):1385-1396.
74. Teng Y, Takahashi Y, Yamada M, et al. IRF4 negatively regulates proliferation of germinal center B cell-derived Burkitt's lymphoma cell lines and induces differentiation toward plasma cells. *Eur J Cell Biol*. 2007;86(10):581-589.
75. Heinz S, Benner C, Spann N, et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell*. 2010;38(4):576-589.
76. Babb R, Cleary MA, Herr W. OCA-B is a functional analog of VP16 but targets a separate surface of the Oct-1 POU domain. *Mol Cell Biol*. 1997;17(12):7295-7305.

77. Champhekar A, Damle SS, Freedman G, Carotta S, Nutt SL, Rothenberg EV. Regulation of early T-lineage gene expression and developmental progression by the progenitor cell transcription factor PU.1. *Genes Dev*. 2015;29(8):832-848.
78. Ye BH, Cattoretti G, Shen Q, et al. The BCL-6 proto-oncogene controls germinal-centre formation and Th2-type inflammation. *Nat Genet*. 1997;16(2):161-170.
79. Basso K, Klein U, Niu H, et al. Tracking CD40 signaling during germinal center development. *Blood*. 2004;104(13):4088-4096.
80. Saito M, Gao J, Basso K, et al. A signaling pathway mediating downregulation of BCL6 in germinal center B cells is blocked by BCL6 gene alterations in B cell lymphoma. *Cancer Cell*. 2007;12(3):280-292.
81. Ochiai K, Maienschein-Cline M, Simonetti G, et al. Transcriptional regulation of germinal center B and plasma cell fates by dynamical control of IRF4. *Immunity*. 2013;38(5):918-929.
82. Stauss D, Brunner C, Berberich-Siebelt F, Hopken UE, Lipp M, Muller G. The transcriptional coactivator Bob1 promotes the development of follicular T helper cells via Bcl6. *EMBO J*. 2016;35(8):881-898.
83. Shen Y, Hendershot LM. Identification of ERdj3 and OBF-1/BOB-1/OCA-B as direct targets of XBP-1 during plasma cell differentiation. *J Immunol*. 2007;179(5):2969-2978.
84. Qin XF, Reichlin A, Luo Y, Roeder RG, Nussenzweig MC. OCA-B integrates B cell antigen receptor-, CD40L- and IL 4-mediated signals for the germinal center pathway of B cell development. *EMBO J*. 1998;17(17):5066-5075.
85. Filippakopoulos P, Qi J, Picaud S, et al. Selective inhibition of BET bromodomains. *Nature*. 2010;468(7327):1067-1073.

Figure legends

Figure 1. OCT1, OCT2 and OBF1 show genome-wide colocalization.

A-C. Venn diagrams showing the overlaps between OCT1, OCT2 and OBF1 peaks in LPS-stimulated splenic B cells.

D-F. Correlation of enrichment between OCT1 and OCT2 (D), OBF1 and OCT1 (E), OBF1 and OCT2 (F) ChIP-seq samples under LPS stimulation.

G. Heatmap showing OCT1, OCT2 and OBF1 ChIP enrichment under LPS stimulation in 5-kb windows centered on OCT1 peak summits.

H. Mean of alignments of OCT1, OCT2 and OBF1 ChIP-seq signals centered on OCT1 peak summits within 5-kb genomic window under LPS stimulation.

I. OBF1/OCT1/OCT2 ChIP-seq read densities in LPS-stimulated mouse splenic B cells at two known OBF1 target gene loci.

J. Violin plots showing distributions of expression levels of genes grouped by their association with OBF1 (left), OCT1 (middle) or OCT2 (right) in splenic B cells stimulated with LPS. (mean ± s.d; *p < 0.05; **p < 0.01; ***p < 0.001; ****p < 0.0001; Mann-Whitney U test)

Figure 2. OCT1, OCT2 and OBF1 occupy active functional genomic elements.

A. Genomic distribution of OBF1/OCT1/OCT2 peaks under LPS stimulation. “Promoters” are defined as TSS ± 1kb. “Downstream” is defined as within 3kb downstream of 3’UTR.

B. Distributions of distances between OBF1, OCT1, OCT2, and H3K27ac peaks and the nearest TSS (log_{10} scale).

C. OBF1, OCT1, OCT2 and H3K27ac ChIP-seq enrichments in 10kb genomic windows centered on the middle of overlapping peaks between OBF1, OCT1 and OCT2. ChIP-seq was performed with $CD19^+$ splenic B cells treated with LPS for 72 hours.

D. OCT1, OBF1, OCT2, H3K27ac, H3K4me1 and H3K4me3 ChIP-seq read densities at two individual gene loci, illustrating the colocalization of measured signals.

Figure 3. OCT1, OCT2 and OBF1 colocalize with TFs of the ETS family.

A. *De novo* motif analysis for Bio-ChIP-seq peak regions of OCT2 (top), OBF1 (middle) and OCT1 (bottom) in LPS-stimulated $CD19^+$ mouse splenic B cells. Sequence logos and *P* values are shown for the most highly enriched sequence motifs.

B. Enrichment of Octamer motifs in OCT1, OBF1 and OCT2 in 500-bp windows centered on peak summits. Controls are randomly shuffled peak sequences that retain dinucleotide frequency.

C. Enrichment of PU.1 binding motifs (MA0080.3 in JASPAR database) in 500-bp windows centered on OCT1, OBF1 and OCT2 peak summits. Controls are randomly shuffled peak sequences that retain dinucleotide frequency.

D. Occurrence of PU.1 binding motif in a 500-bp window surrounding Octamer motifs in OBF1 overlapping peak regions. The black line shows the occurrences of original motifs within 500bp around octamer motifs found in OBF1 peaks. Grey lines show the occurrences of randomly shuffled ETS motifs as backgrounds. Each of the original ETS motifs were randomized 500 times. After randomization, *P*-value was calculated by comparing the 95^{th} percentile of the occurrences of original motifs with the 95^{th} percentiles of all randomly shuffled ETS motifs (500 times for each motif).

E. Venn diagrams showing overlaps between OCT1/OCT2/OBF1 peaks and PU.1 ChIP-seq peaks.

F. Average expression (log_2 counts per million, x-axis) and change of expression (log_2 Fold Change, y-axis) for genes in wild type and OBF1 knockout $CD19^+$ splenic B cells stimulated with LPS. Differentially expressed genes are color-coded (DEGs fold change >= 1, FDR < 0.01, logCPM > 2): red data points represent genes with higher expression level in OBF1 knockout cells, green data points represent genes with higher expression in wild type cells.

G. Average expression (log_2 counts per million, x-axis) and change of expression (log_2 Fold Change, y-axis) for genes in wild type and PU.1 knockout $CD19^+$ splenic B cells stimulated with LPS. Differentially expressed genes are color-coded (DEGs fold change >= 1, FDR < 0.01, logCPM > 2): red data points represent genes with higher expression level in PU.1 knockout cells, green data points represent genes with higher expression in wild type cells.

H. Venn diagrams showing the overlaps between significantly upregulated (upper panel) and downregulated (bottom panel) genes identified in OBF1 and PU.1 knockout B cells under LPS treatment. All samples in this Figure were LPS-stimulated $CD19^+$ mouse splenic B cells.

Figure 4. OBF1 stabilizes the binding of OCT1 to chromatin.

A. Normalized number of OBF1 and OCT1 Bio-ChIP-seq reads in overlapping peaks from LPS-stimulated $CD19^+$ mouse splenic B cells.

B. Violin plots showing enrichment of OCT1 Bio-ChIP-seq samples according to their overlap with OBF1.

C. Violin plots showing enrichment of OCT2 Bio-ChIP-seq samples according to their overlap with OBF1.

D. Violin plots showing enrichment of OCT1 Bio-ChIP-seq samples according to their overlap with OBF1, PU.1 or both.

Panel B-D: mean ± s.d; *p < 0.05; **p < 0.01; ***p < 0.001; ****p < 0.0001; two-tailed Student's t test.

E. Immunoblot showing OBF1 levels in LPS-treated splenic B cells in wild type and *Pou2af1* knockout mice. The membrane was probed with an antibody against OBF1; actin is shown as loading control.

F. Differential OCT1 binding regions in WT and *Pou2af1*$^{KO/KO}$ CD19^{+} splenic B cells stimulated by LPS (red dots, OCT1 binding regions showing significantly weaker binding in *Pou2af1* knockout samples than in WT; green dots, OCT1 binding regions showing significantly stronger binding in *Pou2af1* knockout samples than in WT).

G. Heatmap showing the OCT1 Bio-ChIP-seq signal of WT and *Pou2af1*$^{KO/KO}$ CD19^{+} splenic B cells on differentially bound regions. 739 peaks regions show significantly decreased OCT1 Bio-ChIP-seq signal in *Pou2af1* knockout samples. Gene Ontology analysis was performed using genes nearest to the peaks regions.

H. Mean of alignments of H3K27ac ChIP-seq signals centered on OCT1 peak summits within 6-kb genomic window in WT and *Pou2af1*$^{KO/KO}$ CD19^{+} splenic B cells stimulated by LPS. (*P* value was calculated by Mann-Whitney U test)

I. OCT1 ChIP-seq read densities between *Pou2af1*$^{KO/KO}$ and wild type CD19^{+} splenic B treated with LPS for 72 hours.

Figure 5. OBF1 regulates the GC transcriptional program.

A. Differentially expressed genes between anti-CD40/IL4 and LPS-stimulated CD19^{+} mouse splenic B cells. GC-related genes with significantly higher expression under anti-CD40/IL4 stimulation are indicated (yellow dots).

B. Heatmap showing OBF1 ChIP enrichment at promoter regions (TSS ± 1kb) of selected GC-related genes under LPS or anti-CD40/IL4 stimulations (red dots, OCT1 binding regions showing significantly higher binding in B cell treated with anti-CD40/IL4; green dots, OCT1 binding regions showing significantly higher binding in B cell treated with LPS).

C. OCT2, OBF1 and OCT1 reads densities at three individual gene loci, as indicated.

D. Anti-CD40/IL4-induced OBF1 binding to the *Bcl6* and *Myc* gene promoters.

E. Average expression ($\log_2$ counts per million, x-axis) and change of expression ($\log_2$ Fold Change, y-axis) for genes in wild type and OBF1 knockout CD19^{+} splenic B cells stimulated with anti-CD40/IL4. Differentially expressed genes are color-coded (DEGs fold change >= 1, FDR < 0.01, logCPM > 2): red data points represent genes with higher expression level in OBF1 knockout cells, green data points represent genes with higher expression in wild type cells.

F. Gene ontology analysis was performed with the genes up- and downregulated in OBF1 knockout versus wild type splenic B cells stimulated with anti-CD40/IL4.

G. GSEA of relative gene expression in OBF1 knockout versus wild type anti-CD40/IL4 stimulated splenic B cells against the genes downregulated in GC B cells vs plasma cells. FDR, false discovery rate; NES, normalized enrichment score.

H. Heatmaps showing differentially expressed genes between OBF1 knockout versus wild type B cells. Differentially expressed genes involved in GC reaction are selected.

I. Boxplot showing the percentages of germinal center B cells in wild type and *Pou2af1*$^{KO/KO}$ mice by FACS at day 7 post SRBCs challenge. (n=5; mean ± s.d; *p < 0.05; **p < 0.01; ***p < 0.001; ****p < 0.0001; two-sided Student's t test)

J. Upper panel, workflow of germinal center induction by intraperitoneal injection of 10% SRBC solution and the isolation of GC B cells for Bio-ChIP-seq. Lower panel, OBF1 and H3K27ac reads densities at six individual gene loci, as indicated, in purified mouse germinal center B cells induced by SRBCs.

Figure 6. OBF1 is required for the proliferation of GC-derived lymphoma cell lines.

A. Left, immunoblot showing OBF1 knockdown efficiency using shRNA specific for OBF1 (OBF1-sh1) or control (SHC002); histone 3 (H3) is shown as loading control. Right, assessment of proliferation of Raji cell line following shRNA-mediated depletion of OBF1. Cells were seeded at day 1 and counted every second day until day 7 (n = 3).

B. Left, immunoblot of OCT1 demonstrating knockdown efficiency using shRNA specific for OCT1; actin is shown as loading control. Right, assessment of proliferation of Raji cell line following shRNA-mediated depletion of OCT1 (n = 3).

C. Left, immunoblot of OCT2 demonstrating knockdown efficiency using shRNA specific for OCT2; actin is shown as loading control. Right, assessment of proliferation of Raji cell line following shRNA-mediated depletion of OCT2 (n = 3).

D. Left, immunoblot demonstrating knockdown efficiency using shRNA specific for OCT1 and ectopic expression of OBF1 tagged with AviTag and FLAG; actin is shown as loading control. Right, assessment of proliferation of Raji cell line ectopically expressing AviTag-FLAG tagged OBF1 following shRNA-mediated depletion of OCT1 (n = 3).

E. Left, immunoblot demonstrating knockdown efficiency using shRNA specific for OCT2 and ectopic expression of OBF1 tagged with AviTag and FLAG; actin is shown as loading control. Right, assessment of proliferation of Raji cell line ectopically expressing AviTag-FLAG tagged OBF1 following shRNA-mediated depletion of OCT1 (n = 3).

F. Left, immunoblot showing the downregulation of BCL6 upon shRNA-mediated knockdown of OBF1; MCM7 is shown as loading control. Right, quantitative RT-PCR (qPCR) measurements of BCL6 expression upon OBF1 knockdown (same data as shown in Fig. S5E).

G. Left, confirmation of downregulation of OBF1 by shRNA specific for OBF1 by qPCR (n=4); right, confirmation of downregulation of IRF4 by shRNA specific for IRF4 by qPCR (n=4).

H. Assessment of Raji cells proliferation following shRNA-mediated depletion of OBF1 and depletion of both OBF1 and IRF4 (n=3).

I. Assessment of proliferation of Raji cell line following ectopic expression of IRF4. Cells were seeded at day 1 and counted every second day until day 7 (n = 3).

mean ± s.d; *p < 0.05; **p < 0.01; ***p < 0.001; ****p < 0.0001; two-tailed Student's t test.

Figure 7. OBF1 is required for the maintenance of the GC transcriptional program.

A. Left, average expression (log_2 counts per million, x-axis) and change of expression (log_2 Fold Change, y-axis) for genes in control (SHC002) and OBF1 knockdown (Obf1sh1) conditions. Differentially expressed genes are color-coded (DEGs fold change > 1.5, FDR < 0.001, logCPM > 2): red data points represent genes with higher expression level in Raji cells treated with OBF1-specific shRNA, green data points represent genes with higher expression in control-treated (scrambled shRNA sequence) Raji cells. Right, gene ontology analysis was performed with the genes up- and downregulated in OBF1 knockout versus wild type splenic B cells stimulated with anti-CD40/IL4.

B. Heatmaps showing differentially expressed genes involved in GC reaction between Raji cells infected with lentivirus expressing control shRNA (SHC002) and OBF1-specific shRNA.

C. GSEA of relative gene expression in OBF1 depleted versus wild type Raji cells against the gene set identified as genes upregulated in plasma cells. FDR, false discovery rate; NES, normalized enrichment score.

D. Average expression (log_2 counts per million, x-axis) and change of expression (log_2 Fold Change, y-axis) for genes in wild type and ectopically IRF4 expressing Raji cells. Differentially expressed genes are color-coded (DEGs fold change >= 1, FDR < 0.01, logCPM > 2): red data points represent genes with higher expression level in ectopically IRF4-expressing Raji cells, green data points represent genes with higher expression in wild type Raji cells.

E. GSEA of relative gene expression in ectopically IRF4-expressing versus wild type Raji cells against the differentially expressed genes in OBF1-depleted Raji cells. FDR, false discovery rate; NES, normalized enrichment score.

F. OBF1 CUT&RUN read densities at two individual loci in Raji and SUDHL4 cells.

G. H3K27ac and OBF1 CUT&RUN read densities at two individual loci, as indicated, in purified germinal center B cells from human tonsils.

H. GSEA of relative gene expression in OBF1-depleted versus wild type Raji cells against the gene set identified as genes associated with favourable prognosis for DLBCL patients. FDR, false discovery rate; NES, normalized enrichment score.

Figure 1

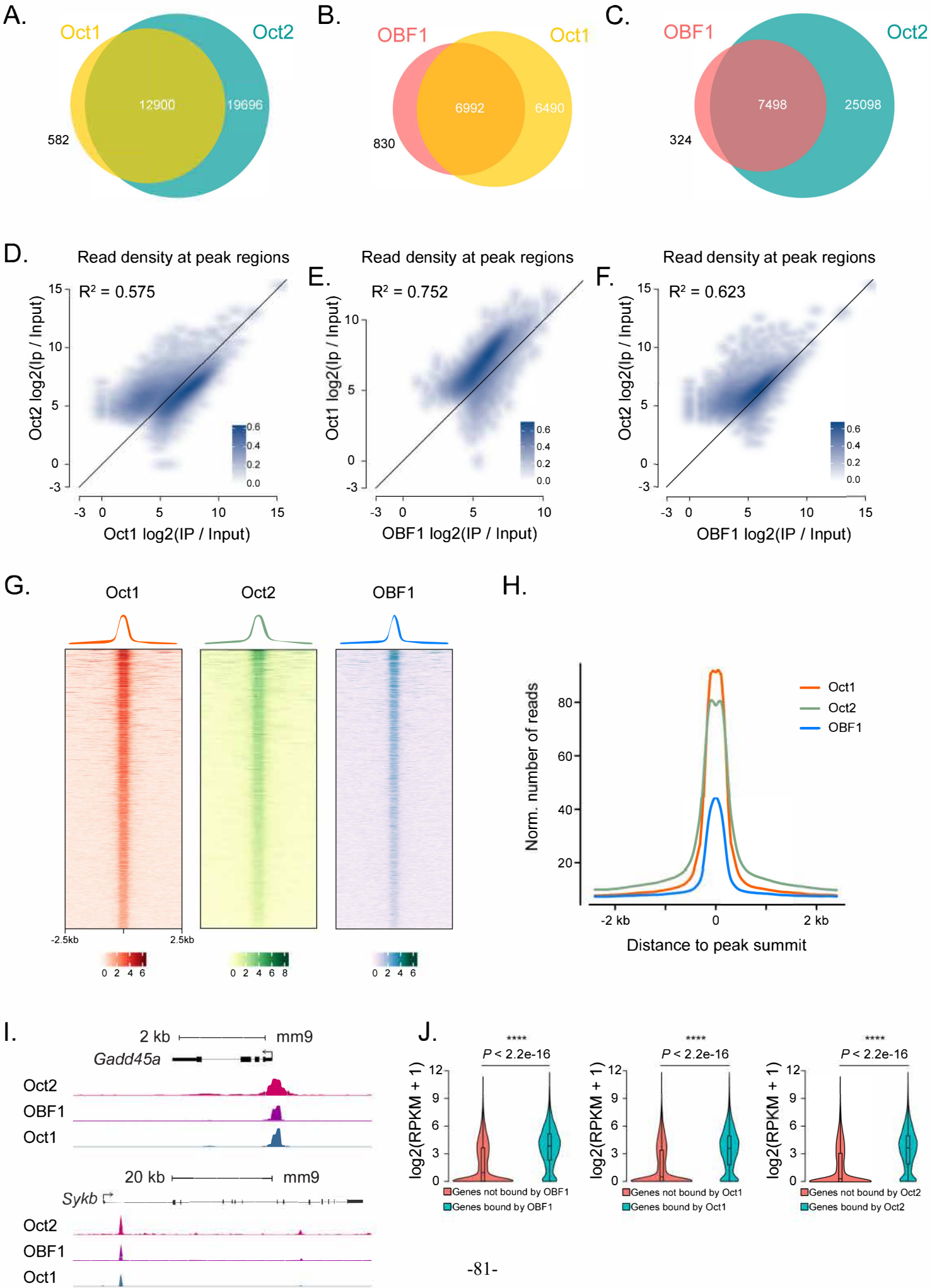
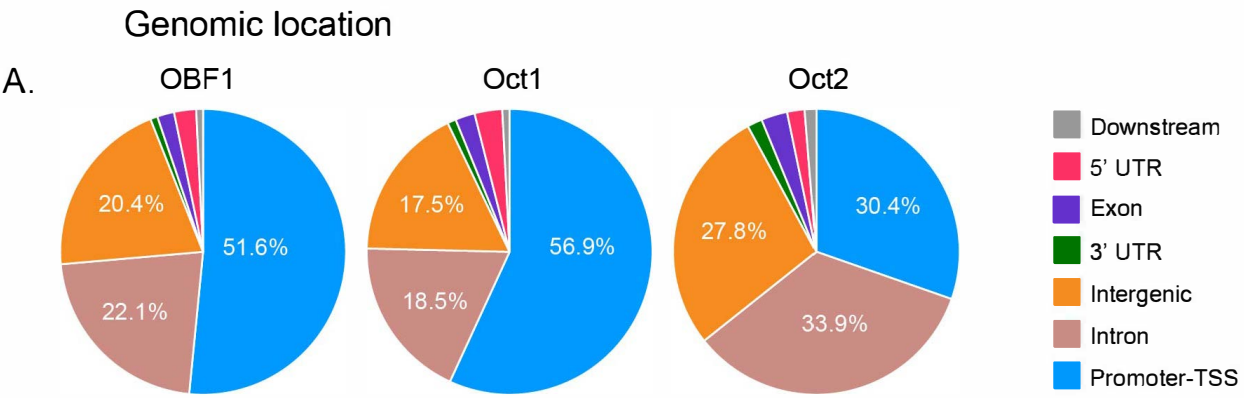
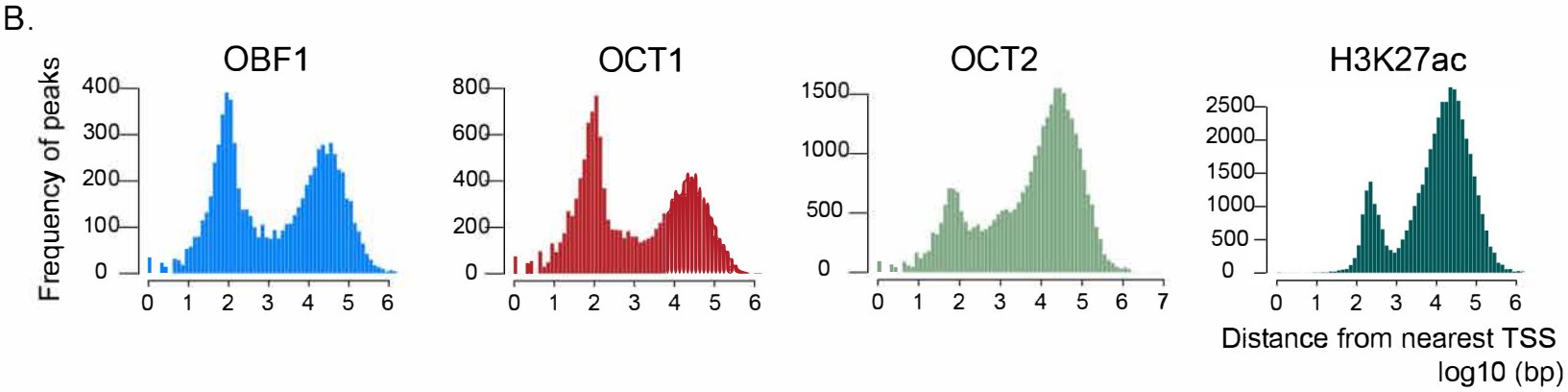
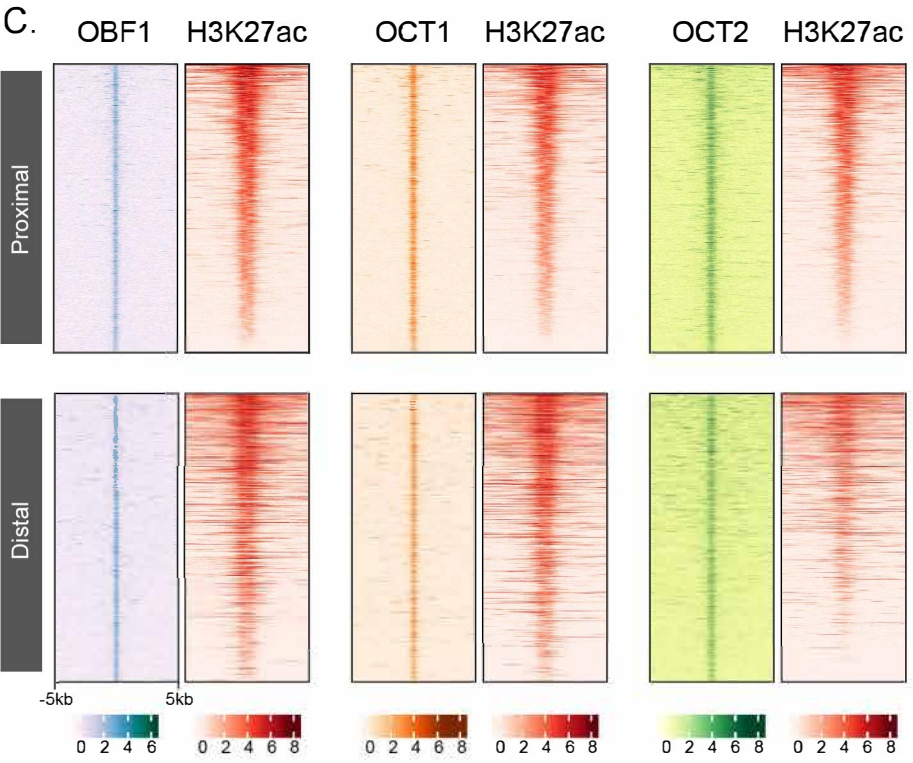

A.
Oct1
Oct2
12900
19696
582
B.
OBF1
Oct1
6992
6490
830
C.
OBF1
Oct2
7498
25098
324
D.
Read density at peak regions
R² = 0.575
Oct2 log2(Ip / Input)
Oct1 log2(IP / Input)
E.
Read density at peak regions
R² = 0.752
Oct1 log2(Ip / Input)
OBF1 log2(IP / Input)
F.
Read density at peak regions
R² = 0.623
Oct2 log2(Ip / Input)
OBF1 log2(IP / Input)
G.
Oct1
Oct2
OBF1
-2.5kb
2.5kb
H.
Norm. number of reads
Oct1
Oct2
OBF1
-2 kb
0
2 kb
Distance to peak summit
I.
2 kb
mm9
Gadd45a
Oct2
OBF1
Oct1
20 kb
mm9
Sykb
Oct2
OBF1
Oct1
J.
****
P < 2.2e-16
log2(RPKM + 1)
Genes not bound by OBF1
Genes bound by OBF1
Genes not bound by Oct1
Genes bound by Oct1
Genes not bound by Oct2
Genes bound by Oct2

Figure 2


A.
Genomic location
OBF1
Oct1
Oct2
20.4%
51.6%
22.1%
17.5%
56.9%
18.5%
27.8%
30.4%
33.9%
Downstream
5' UTR
Exon
3' UTR
Intergenic
Intron
Promoter-TSS


B.
OBF1
OCT1
OCT2
H3K27ac
Frequency of peaks
Distance from nearest TSS
log10 (bp)


C.
OBF1
H3K27ac
OCT1
H3K27ac
OCT2
H3K27ac
Proximal
Distal
-5kb
5kb

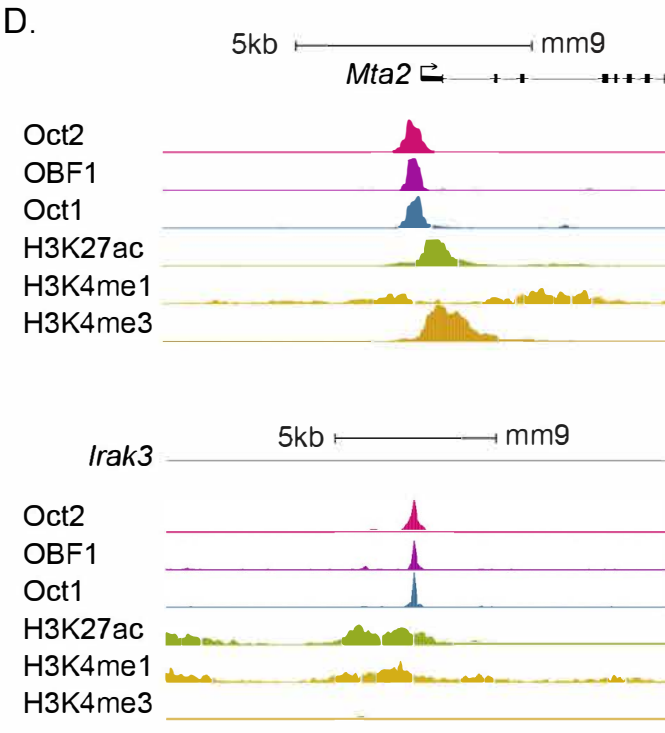

D.
5kb
mm9
Mta2
Oct2
OBF1
Oct1
H3K27ac
H3K4me1
H3K4me3
Irak3

Figure 3

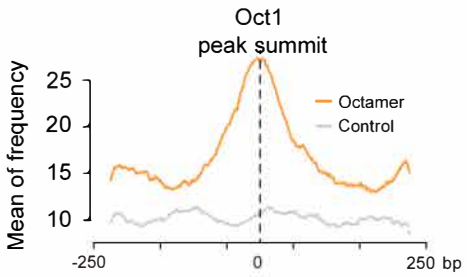
A.

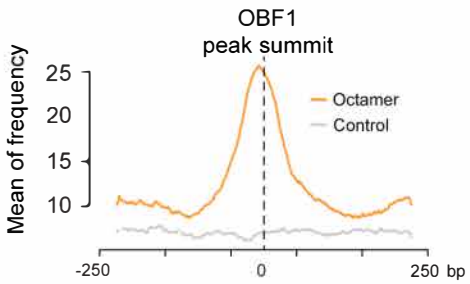
Oct2 peak	Motif *P* value	% peak with motif	Name
	1e-521	40.89%	POU homeobox
	1e-472	49.33%	POU homeobox
	1e-416	39.57%	POU homeobox
	1e-276	40.75%	POU homeobox
	1e-101	56.23%	ETS

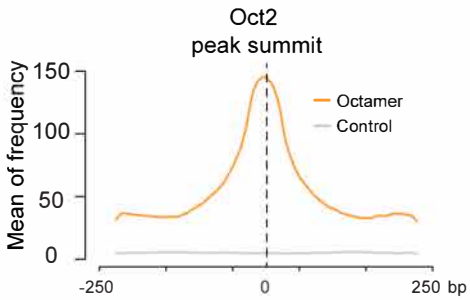
OBF1 peak	Motif *P* value	% peak with motif	Name
	1e-1710	24.57%	POU homeobox
	1e-845	39.84%	Zf
	1e-706	32.83%	ETS
	1e-432	15.55%	NRF

Oct1 peak	Motif *P* value	% peak with motif	Name
	1e-1536	41.73%	Zf
	1e-1154	13.36%	POU homeobox
	1e-706	31.13%	ETS
	1e-432	15.72%	CCAAT box

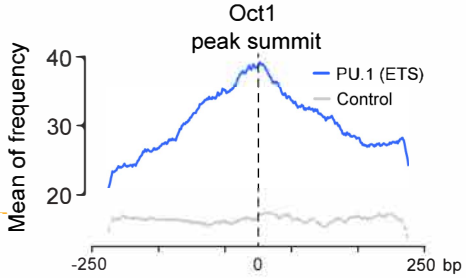
B.

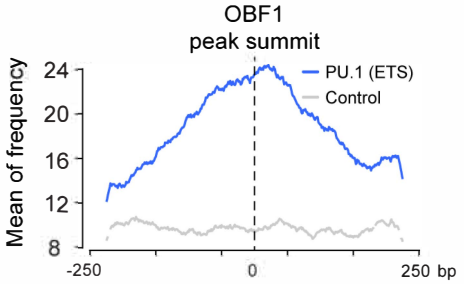


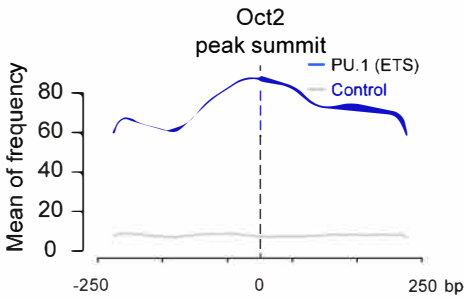




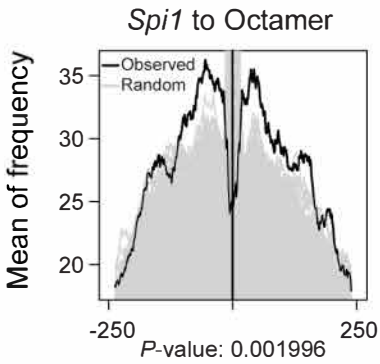
C.



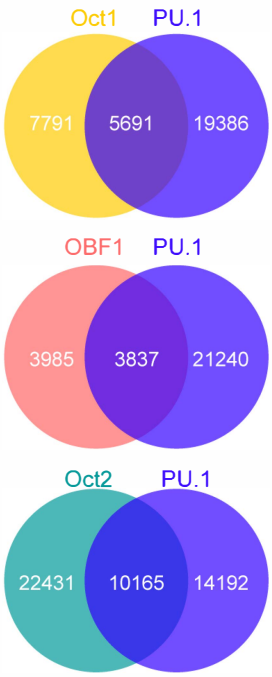




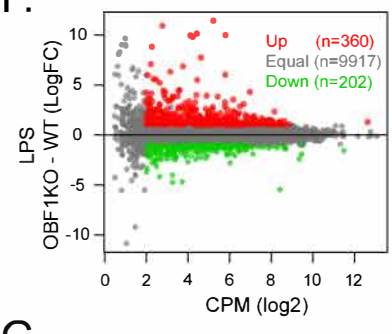
D.



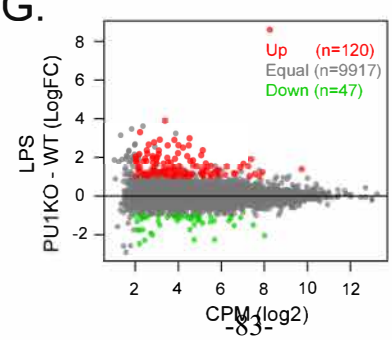
E.



F.



G.



H

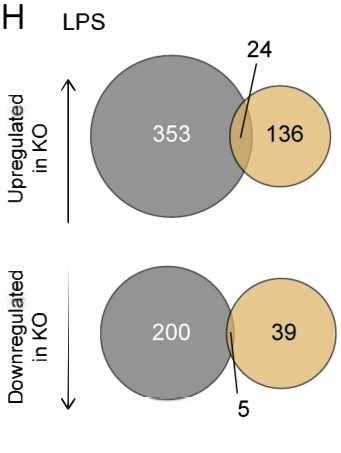


Figure 4

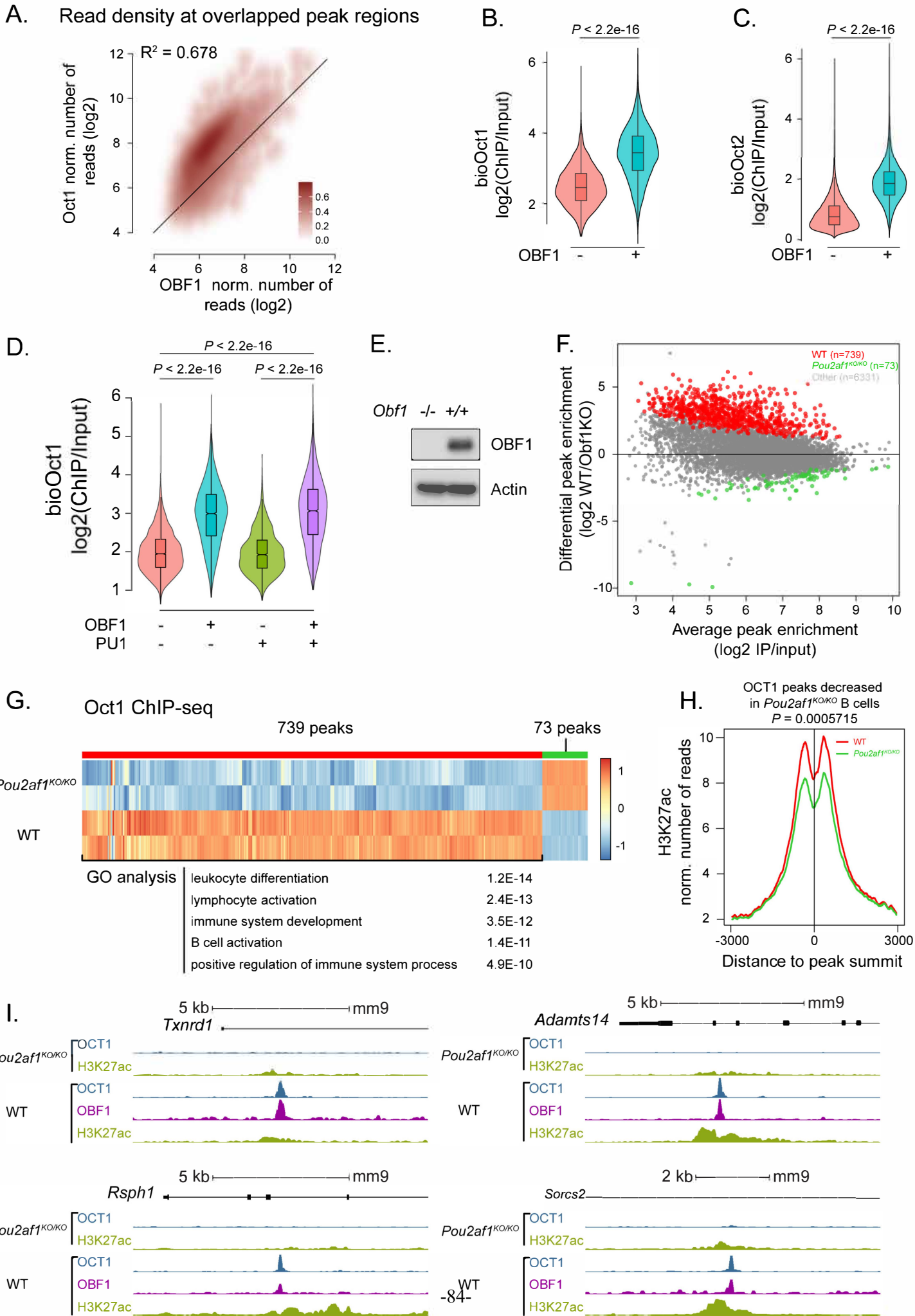

A.
Read density at overlapped peak regions
R² = 0.678
Oct1 norm. number of reads (log2)
OBF1 norm. number of reads (log2)
B.
P < 2.2e-16
bioOct1 log2(ChIP/Input)
OBF1 - +
C.
P < 2.2e-16
bioOct2 log2(ChIP/Input)
OBF1 - +
D.
P < 2.2e-16
P < 2.2e-16
P < 2.2e-16
bioOct1 log2(ChIP/Input)
OBF1 - + - +
PU1 - - + +
E.
Obf1 -/- +/+
OBF1
Actin
F.
WT (n=739)
Pou2af1KO/KO (n=73)
Other (n=6331)
Differential peak enrichment (log2 WT/Obf1KO)
Average peak enrichment (log2 IP/input)
G.
Oct1 ChIP-seq
739 peaks
73 peaks
Pou2af1KO/KO
WT
GO analysis
leukocyte differentiation 1.2E-14
lymphocyte activation 2.4E-13
immune system development 3.5E-12
B cell activation 1.4E-11
positive regulation of immune system process 4.9E-10
H.
OCT1 peaks decreased in Pou2af1KO/KO B cells
P = 0.0005715
WT
Pou2af1KO/KO
H3K27ac norm. number of reads
Distance to peak summit
I.
5 kb mm9
Txnrd1
Adamts14
Rsph1
Sorcs2
2 kb mm9
Pou2af1KO/KO
WT
OCT1
OBF1
H3K27ac

Figure 5

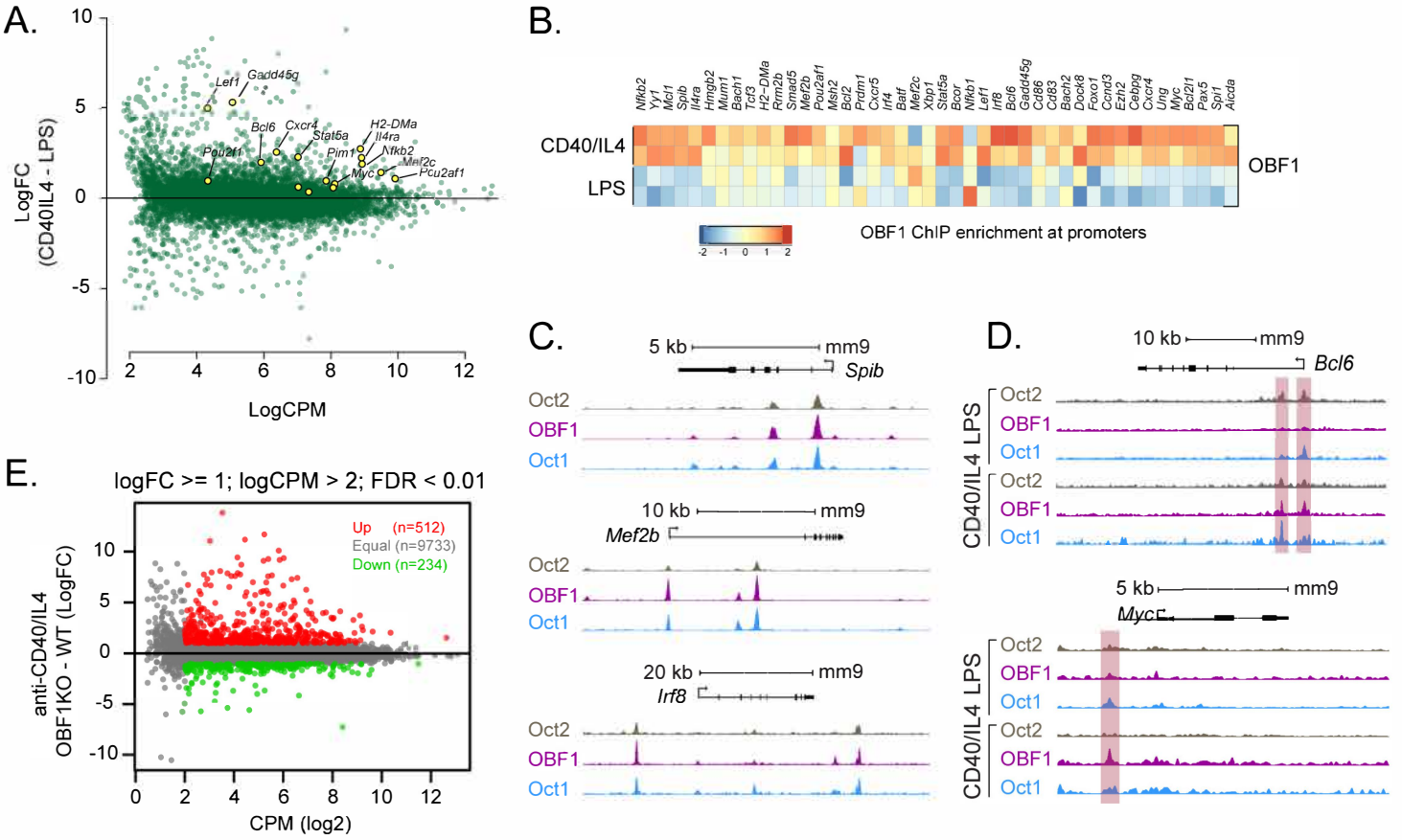

A.
LogFC (CD40IL4 - LPS)
LogCPM
Lef1
Gadd45g
Bcl6
Cxcr4
Stat5a
H2-DMa
Il4ra
Pou2f1
Pim1
Nfkb2
Myc
Mef2c
Pcu2af1
B.
CD40/IL4
LPS
OBF1
OBF1 ChIP enrichment at promoters
C.
5 kb mm9
Spib
Oct2
OBF1
Oct1
10 kb mm9
Mef2b
20 kb mm9
Irf8
D.
10 kb mm9
Bcl6
CD40/IL4 LPS
5 kb mm9
Myc
E.
logFC >= 1; logCPM > 2; FDR < 0.01
Up (n=512)
Equal (n=9733)
Down (n=234)
anti-CD40/IL4 OBF1KO - WT (LogFC)
CPM (log2)
F.
Significantly enriched GO terms
immune system process
immune response
innate immune response
interspecies interaction between organisms
symbiosis, encompassing mutualism through parasitism
defense response
multi-organism cellular process
viral process
regulation of multicellular organismal process
immune system response
positive regulation of biological process
positive regulation of cellular process
single-multicellular organism process
anantomical structure development
multicellular organism develpoment
regulation of immune system process
-log10(P-value)
G.
Enrichment plot:
GSE12366
Genes downregulated in
GC B cell vs plasma cell
Enrichment score (ES)
NES: 1.54
FDR: 0.046
OBF1KO splenic B cells
WT splenic B cells
Ranked list metric (Signal2Noise)
Rank in Ordered Dataset
H.
WT
Pou2af1koko
I.
Percentage of GC population post SRBC immunization (%)
P = 0.0002525
P = 0.0002143
PBS
SRBC
OBF1KO
WT
J.
SRBC; i.p.
Day 0
GC indcued
Day 7
MACS
PNA+ beads
Bio-ChIP-seq
100 kb mm9
Foxo1
10 kb mm9
Mef2b
200 kb mm9
Bach2
5 kb mm9
Egr3
5 kb mm9
Aicda
20 kb mm9
Bcl6
H3K27ac
OBF1

Figure 6

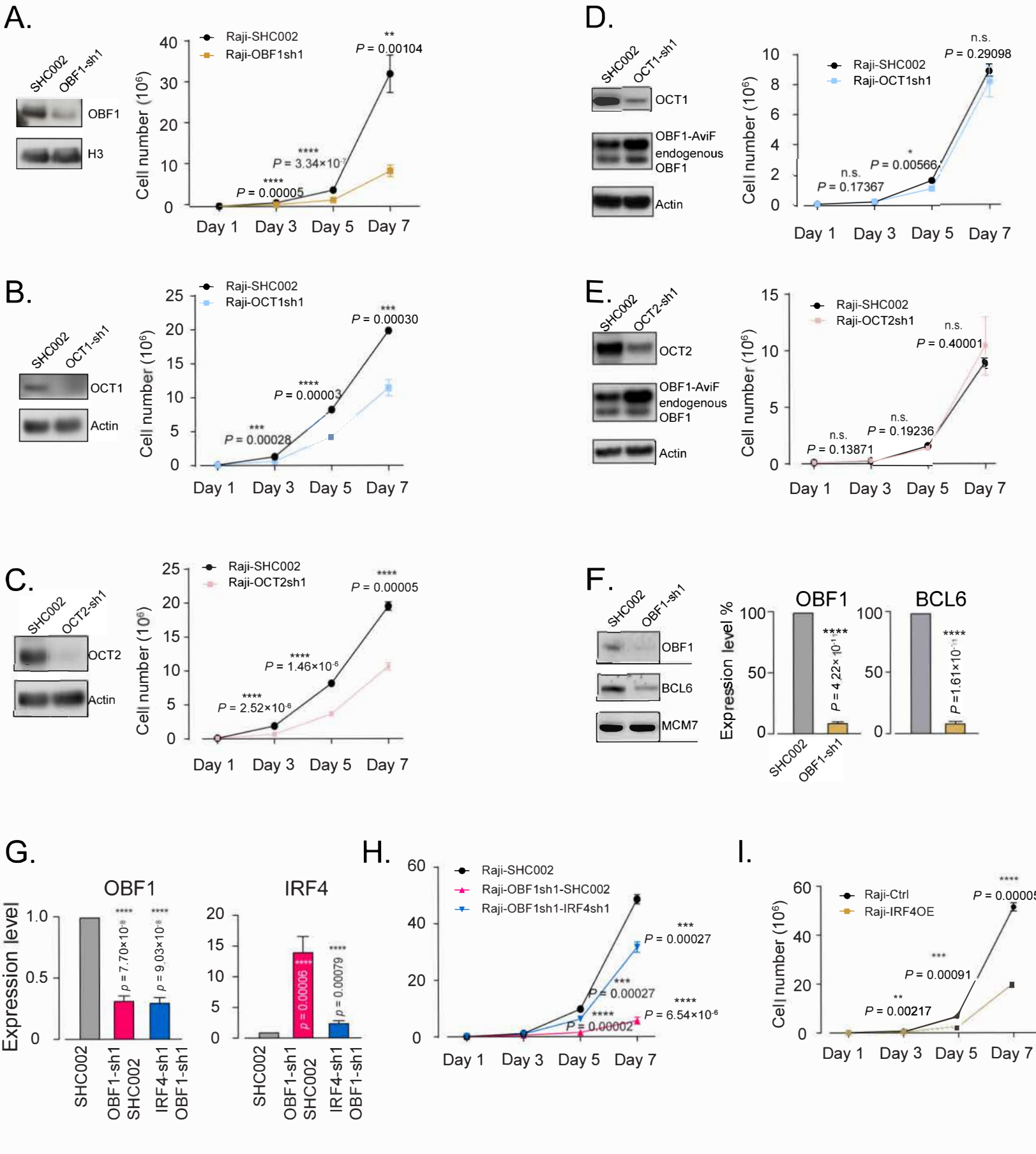
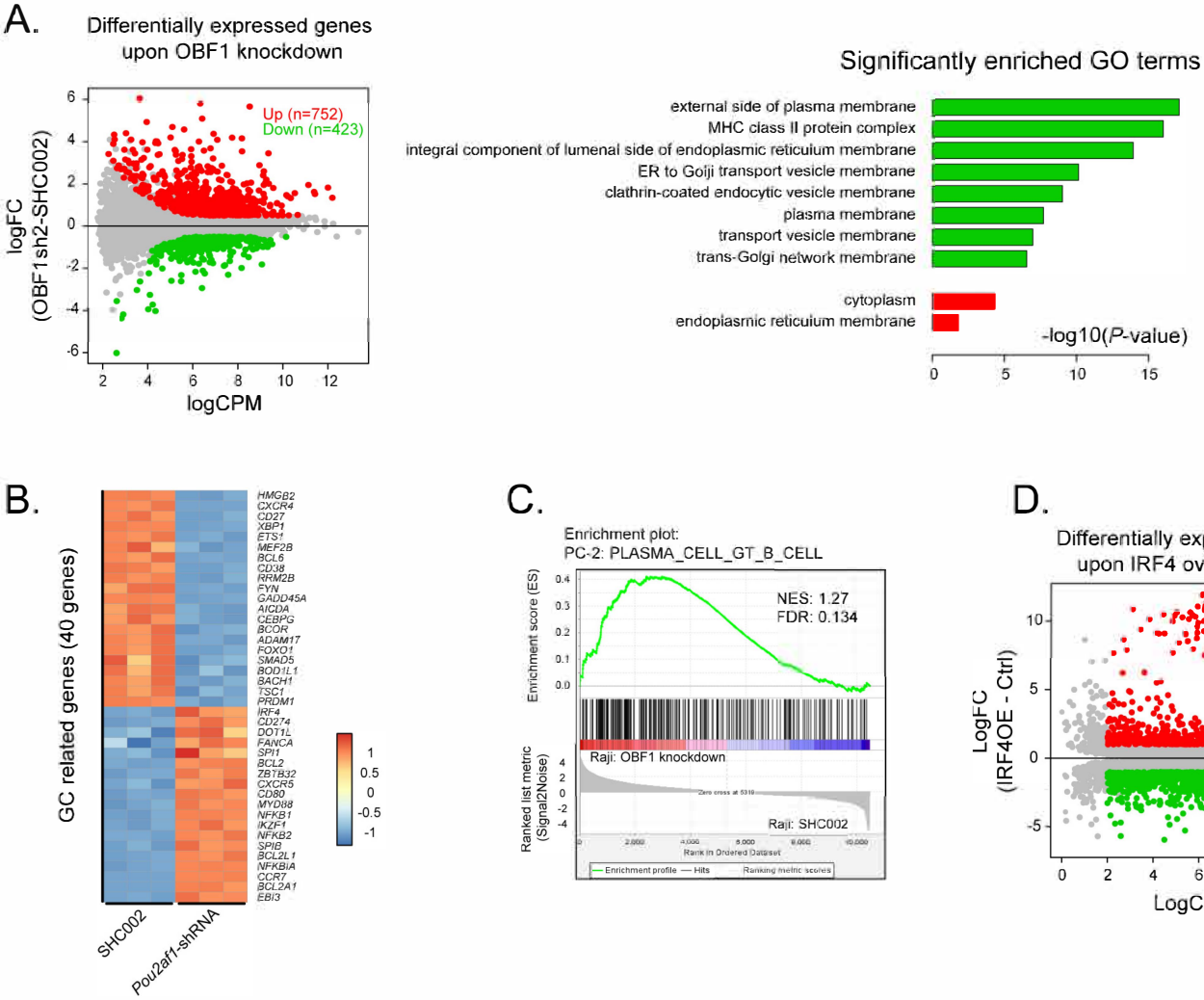
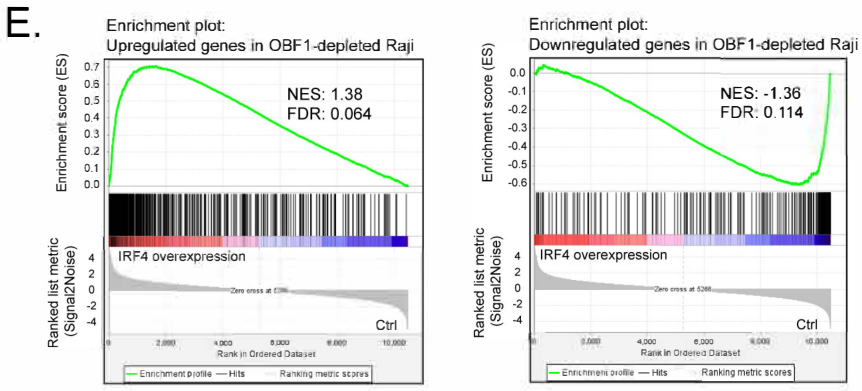
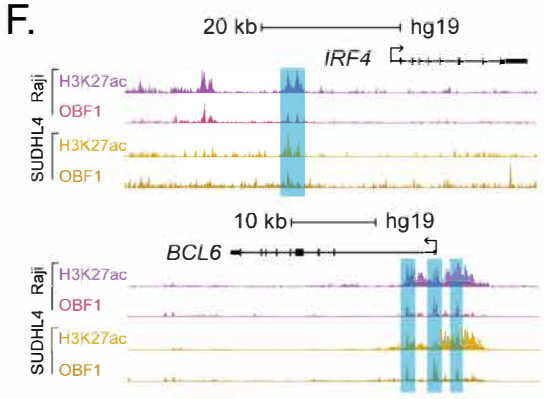
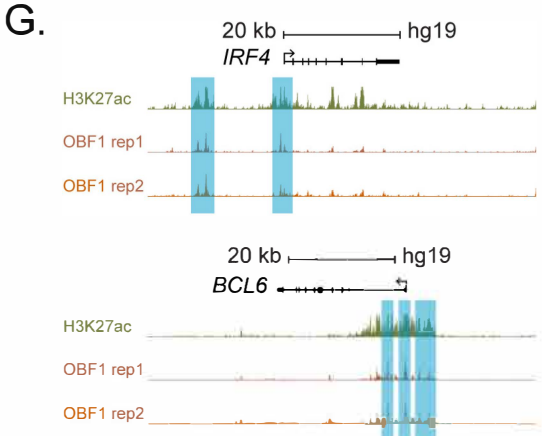

A.
SHC002
OBF1-sh1
OBF1
H3
Raji-SHC002
Raji-OBF1sh1
Cell number (10⁶)
P = 0.00005
P = 3.34×10⁻⁷
P = 0.00104
Day 1
Day 3
Day 5
Day 7
B.
SHC002
OCT1-sh1
OCT1
Actin
Raji-SHC002
Raji-OCT1sh1
P = 0.00028
P = 0.00003
P = 0.00030
C.
SHC002
OCT2-sh1
OCT2
Actin
Raji-SHC002
Raji-OCT2sh1
P = 2.52×10⁻⁶
P = 1.46×10⁻⁶
P = 0.00005
D.
SHC002
OCT1-sh1
OCT1
OBF1-AviF
endogenous
OBF1
Actin
Raji-SHC002
Raji-OCT1sh1
P = 0.17367
P = 0.00566
P = 0.29098
E.
SHC002
OCT2-sh1
OCT2
OBF1-AviF
endogenous
OBF1
Actin
Raji-SHC002
Raji-OCT2sh1
P = 0.13871
P = 0.19236
P = 0.40001
F.
SHC002
OBF1-sh1
OBF1
BCL6
MCM7
Expression level %
OBF1
BCL6
P = 4.22×10⁻¹
P = 1.61×10⁻⁵
G.
OBF1
IRF4
Expression level
p = 7.70×10⁻⁸
p = 9.03×10⁻⁸
p = 0.00006
p = 0.00079
SHC002
OBF1-sh1 SHC002
IRF4-sh1 OBF1-sh1
H.
Raji-SHC002
Raji-OBF1sh1-SHC002
Raji-OBF1sh1-IRF4sh1
P = 0.00027
P = 0.00027
P = 0.00002
P = 6.54×10⁻⁶
I.
Raji-Ctrl
Raji-IRF4OE
P = 0.00217
P = 0.00091
P = 0.00005

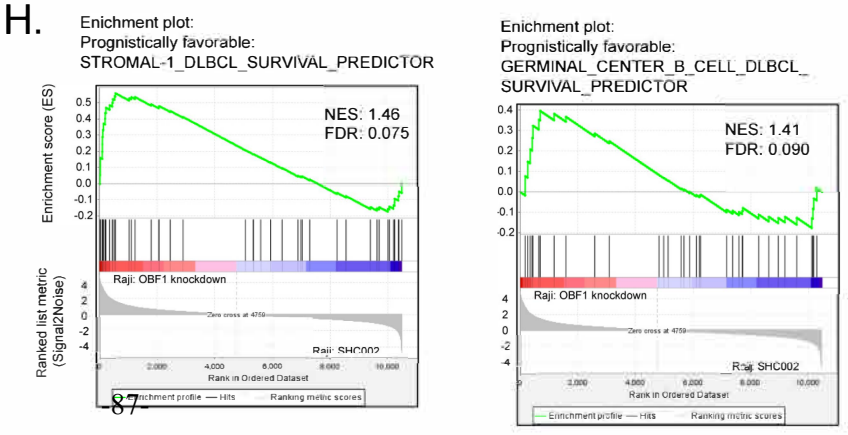
Figure 7


A.
Differentially expressed genes upon OBF1 knockdown
Up (n=752)
Down (n=423)
logFC (OBF1sh2-SHC002)
logCPM
Significantly enriched GO terms
external side of plasma membrane
MHC class II protein complex
integral component of lumenal side of endoplasmic reticulum membrane
ER to Golgi transport vesicle membrane
clathrin-coated endocytic vesicle membrane
plasma membrane
transport vesicle membrane
trans-Golgi network membrane
cytoplasm
endoplasmic reticulum membrane
-log10(P-value)
B.
GC related genes (40 genes)
SHC002
Pou2af1-shRNA
C.
Enrichment plot:
PC-2: PLASMA_CELL_GT_B_CELL
NES: 1.27
FDR: 0.134
Raji: OBF1 knockdown
Raji: SHC002
D.
Differentially expressed genes upon IRF4 overexpression
Up (n=528)
Down (n=988)
LogFC (IRF4OE - Ctrl)
LogCPM


E.
Enrichment plot:
Upregulated genes in OBF1-depleted Raji
NES: 1.38
FDR: 0.064
IRF4 overexpression
Ctrl
Enrichment plot:
Downregulated genes in OBF1-depleted Raji
NES: -1.36
FDR: 0.114
IRF4 overexpression
Ctrl


F.
20 kb
hg19
IRF4
Raji
H3K27ac
OBF1
SUDHL4
H3K27ac
OBF1
10 kb
hg19
BCL6
Raji
H3K27ac
OBF1
SUDHL4
H3K27ac
OBF1


G.
20 kb
hg19
IRF4
H3K27ac
OBF1 rep1
OBF1 rep2
20 kb
hg19
BCL6
H3K27ac
OBF1 rep1
OBF1 rep2


H.
Enichment plot:
Prognistically favorable:
STROMAL-1_DLBCL_SURVIVAL_PREDICTOR
NES: 1.46
FDR: 0.075
Raji: OBF1 knockdown
Raji: SHC002
Enichment plot:
Prognistically favorable:
GERMINAL_CENTER_B_CELL_DLBCL_SURVIVAL_PREDICTOR
NES: 1.41
FDR: 0.090
Raji: OBF1 knockdown
Raji: SHC002

Supplemental Methods

Cell Culture

For T-independent stimulation, primary spleen mature B cells were cultured in RPMI1640 medium (10% FBS, 50 μm 2-mercaptoethanol (AppliChem), 50 μg/mL LPS (Sigma), 2 mM Glutamine, 1% ITS (Sigma), 1% MEM (Sigma), 1mM Sodium Pyruvate (ThermoFisher Scientific)), medium was changed every day. For T-dependent stimulation, primary spleen mature B cells were cultured in RPMI1640 medium (10% FBS, 50 μm 2-mercaptoethanol (AppliChem), 1 ug/mL anti-murine-CD40 + 20 ng/mL IL4 (R&D Systems), 2 mM Glutamine, 1% ITS (Sigma), 1% MEM (Sigma), 1 mM Sodium Pyruvate (ThermoFisher Scientific)), medium was changed every day. For each stimulation, B cells were harvested after 72 hours.

Mouse immunization with SRBCs

The immunization method has been described [1]. 1mL sheep blood was washed twice by adding 10 volume PBS into the 15mL falcon tube. The washed SRBC pellet was then resuspended in 10 mL PBS to make a 10% (v/v) SRBC suspension. 0.1 mL of 10% SRBC suspension was injected intraperitoneally into each mouse (8-12 weeks old). Seven days after injection, germinal center cells were harvested for purification or analysis.

MACS sorting of $CD19^+$ mature B cells

Single cell suspension was prepared by smashing the spleen, and then washing in freshly prepared MACS buffer (0.5% BSA and 2mM EDTA in PBS) (20mL for two spleens) three times. After washing, cell concentration was adjusted to 10^8 cells/mL. To label the B cells for MACS LS column separation, 10 μL CD19 beads was added to every 10^7 cells (100 μL), and the mixture was incubated at 4 ℃ for 15min. The anti-CD19-labeled spleen cells were further washed with MACS buffer and then applied to MACS LS column. $CD19^+$ mature B cells were purified according to the standard protocol of Miltenyi Biotec LS column (Cat# 130-042-401). Purified $CD19^+$ mature B cells were subjected to in vitro culture with LPS or anti-CD40/IL4 stimulants at 2×10^6 cells/mL.

MACS sorting of murine GC B cells

Seven days after SRBC injection, mice were sacrificed, and single cell suspension was prepared by smashing the spleen, and then washing in freshly prepared MACS buffer (0.5% BSA and 2mM EDTA in

PBS) (20mL for two spleens) three times. The germinal center were then prepared following procedures from PNA MicroBead Kit (Germinal Center B Cell (PNA) MicroBead Kit, 130-110-479, Miltenyi Biotec).

Generation of transgenic mice

Gene-specific sgRNAs were designed using the tool provided by the Zhang lab (https://zlab.bio/guide-design-resources). The generation of sgRNA constructs was described in [2]. The pX458 plasmid was used as the vector for sgRNA cloning. Commercially synthesized top and bottom strands of sgRNA templates specific for each gene were in vitro annealed, and subsequently ligated into the BbsI site of pX458 plasmid. The T7 promoter was added to the 5' end of sgRNA coding sequences by PCR, and sgRNAs were transcribed by T7 RNA polymerase (MEGAscript T7 Transcription kit). Then, in vitro transcribed sgRNAs were purified by MEGAclear Transcription Clean-Up kit (Thermo Fisher Scientific). The quality of sgRNAs were analyzed by electrophoresis on TBE-Urea gel (Novex TBE-Urea gel, EC6885BOX, Thermo Fisher Scientific). Synthesized gRNA strands for generating sgRNA-pX458 plasmids are listed in Supplemental Table 2.

Single-stranded oligo DNA nucleotides (ssODN) for each gene (*Pou2f1*, *Pou2f2* and *Pou2af1*) were synthesized by Integrated DNA Technologies, and were reconstituted with ultrapure water at a concentration of 1 μg/μL. Each ssODN consists of 40-nt upstream and downstream homologous arms at the 5'- and 3'-ends. Coding sequence for AviTag and FLAG was located between the upstream and downstream homologous arms. Sequences of ssODNs are specified in Supplemental Table 5.

We performed microinjection with the described protocol [2] with some modifications. sgRNAs, Cas9 protein (Toolgen), Cas9 mRNA (Sigma) and ssODNs were mixed and co-injected into one-cell stage mouse zygotes. For each zygote, a solution containing 25 ng/μL sgRNA, 25 ng/μL Cas9 protein, 50 ng/μL Cas9 mRNA and 7 ng/μL ssODN was micro-injected into the pronucleus. ~200 zygotes were injected for each edited gene. Injected zygotes were immediately transplanted back to the pseudopregnant foster mothers. In 3 weeks, CRISPR/Cas9 treated pups were delivered.

To genotype the resulting transgenic pups, genomic DNA was extracted from CRISPR/Cas9 treated pups using DNA extraction buffer (10% Chelex 100 in ultrapure water). The targeted regions of the genome were amplified by PCR. The resulting PCR products were TA-cloned using pGEM-T Easy vector systems (A1360, Promega), and subjected to Sanger sequencing. Primers used in genotyping for genome-edited mice are listed in Supplemental Table 3.

Gene knockdown experiment by shRNA

To prepare lentivirus for infection, 1.3×10^6 293T cells were seeded on T25 flask with 5mL DMEM medium 24 hours before transfection and, the following day, 7 µg pLKO.1-shRNA vector, together with plasmids coding for Rev, Vsv-g, Tat and Gag/Pol, was transfected into 293T cells with FuGENE HD (Promega). Supernatant was collected at 24, 48 and 72 hours post transfection, and filtered through 0.4 µm EMD Millipore Steriflip Sterile Disposable Vacuum Filter units. Lentivirus particles in this 15 mL suspension were then concentrated by adding 5 mL of LentiX concentrator (Takara), and incubated for 24 hours at 4 ℃. Post incubation, lentiviral particles were precipitated and centrifuged at 1500 g for 45 min at 4 ℃. Pelleted lentiviruses were then concentrated 150 times by resuspending with 100 µL culture medium (RPMI1640, 10% FCS, 50 µM 2-mercaptoethanol). 10 µL lentivirus suspension was added to 2×10^5 cells (Raji and Ramos) in 6-well plate containing 8 µg/mL Polybrene (Merck) in RPMI1640 medium. B lymphoma cells with lentiviruses were infected for 5 days at 37℃ and 5% CO_2. Post infection, stable integrations were selected for 7 days in the presence of 3 µg/mL Puromycin (GIBCO) or Blasticidin (GIBCO). Non-target shRNA (Cat# SHC002, Sigma) was used as a control in shRNA-mediated gene knockdown experiments. After selection, stably integrated B lymphoma cells were used for downstream assays. For the proliferation measurements, cells treated with shRNAs were quantified by flow cytometry using Vi-CELL XR Cell Viability Analyzer (Beckman Coulter) at the indicated time points.

Chromatin immunoprecipitation (ChIP)

Immediately after centrifugation, supernatant was discarded, and cells were resuspended and fixed with 10 mL 1% formaldehyde in PBS at room temperature for 10min, followed by quenching with 530µL 2.5 M Glycine at room temperature for 5 min. Cells were lysed for 10 min in 10 mL lysis buffer A (10 mM HEPES, pH8.0, 10 mM EDTA, 0.5 mM EGTA, 0.25% Triton X-100, PI). The supernatant was discarded, and pellets were incubated for 10 min in lysis buffer B (10 mM HEPES, pH8.0, 200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 0.01% Triton X-100, PI). Then, the nuclei were resuspended in chromatin lysis buffer (50 mM Tris-HCl, pH 8.0, 10 mM EDTA, 0.5% SDS) for 30 min at 4 ℃. Subsequently, sonication was performed using Bioruptor Next Gen (Diagenode) at high output intensity for 25 cycles (30s on/30s off). Finally, the fragmented chromatin suspension was diluted 5 times with chromatin dilution buffer (50 mM Tris-HCl, pH8.0, 10 mM EDTA, 167 mM NaCl, 1.67% Triton X-100). 30-50 µg chromatin was subjected to immunoprecipitation with 5-10 µg ChIP-grade antibodies and incubated overnight. Subsequently, the samples were incubated with Dynabeads Protein G (Thermo Fisher Scientific) for 2 hours. The beads were washed with the following buffers: twice with low salt buffer (20 mM Tris pH 8.0, 0.1% SDS, 1% Triton X-100, 2 mM EDTA, 150 mM NaCl), twice with high salt buffer (20 mM Tris pH 8.0, 0.1% SDS, 1%

Triton X-100, 2 mM EDTA, 500 mM NaCl), twice with LiCl buffer (10 mM Tris pH 8.0, 250 mM LiCl, 1 mM EDTA, 1% NP-40, 1% Na Deoxychholate) and once with TE (20 mM Tris-HCl, pH8.9, 2 mM EDTA). The protein-DNA complexes were eluted from the beads with 350 µL elution buffer (10 mM Tris-HCl, pH7.5, 1 mM EDTA, 1% SDS, 100 mM NaHCO3) and reverse crosslinked overnight at 65 ℃ with 60 µg Proteinase K and 60 µg RNase A. Immunoprecipitated DNA was recovered using AMPure XP beads. PU.1, H3K27ac, H3K4me3 and H3K4me1 ChIPs from spleen B cells were subjected to high-throughput sequencing on Illumina HiSeq 2500 sequencer using standard NEB library preparation kits and protocols.

Bio-ChIP-seq

Cells were fixed with 10 mL 1% formaldehyde in PBS at room temperature for 10min, followed by quenching with 530 µL 2.5 M Glycine at room temperature for 5 min. Cells were lysed for 10 min in 10 mL lysis buffer A (10 mM HEPES, pH8.0, 10 mM EDTA, 0.5 mM EGTA, 0.25% Triton X-100, PI). The supernatant was discarded, and pellets were incubated for 10 min in lysis buffer B (10 mM HEPES, pH8.0, 200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 0.01% Triton X-100, PI). Then, the nuclei were resuspended in chromatin lysis buffer (50 mM Tris-HCl, pH 8.0, 10 mM EDTA, 0.5% SDS) for 30 min at 4 ℃. Subsequently, sonication was performed using Bioruptor Next Gen (Diagenode) at high output intensity for 25 cycles (30 s on/30 s off). Finally, the fragmented chromatin suspension was diluted 5 times with chromatin dilution buffer (50 mM Tris-HCl, pH8.0, 10 mM EDTA, 167 mM NaCl, 1.67% Triton X-100). 30-50 µg chromatin was subjected to immunoprecipitation with Dynabeads M-280 Streptavidin and incubated overnight. Subsequently, the samples were incubated with Dynabeads M-280 Streptavidin (Thermo Fisher Scientific) for 2 hours. The beads were washed with the following buffers: twice with 2% SDS/TE buffer (20 mM Tris-HCl, pH8.9, 2 mM EDTA, 2% SDS), twice with high salt buffer (50 mM HEPES, pH7.5, 1 mM EDTA, 1% Triton X-100, 0.1% Na Deoxychholate, 0.1% SDS, 500 mM NaCl, 1 mM DTT, PI), twice with DOC buffer (10 mM Tris-HCl, pH8.0, 500 mM LiCl, 1% NP-40, 1% Na Deoxychholate, 1 mM EDTA) and twice with TE (20 mM Tris-HCl, pH8.9, 2 mM EDTA). To elute ChIP DNA, the bead-bound protein-DNA complexes were incubated with 300 µL TE (10 mM Tris-HCl, pH7.5, 1 mM EDTA) and 60 µg RNase for 30 min at 37℃. Then 30 µL 10% SDS, 6 µL 5 M NaCl and 60 µg Proteinase K was added and the samples incubated overnight at 65 ℃. Immunoprecipitated DNA was recovered using AMPure XP beads. Bio-ChIPed DNA for the three factors was subjected to high-throughput sequencing on Illumina HiSeq 2500 sequencer using standard NEB library preparation kits and protocols.

Gene Set Enrichment Analysis

To perform gene set enrichment analysis (GSEA) [3,4], RNA-seq data count table were normalized by TMM method using R package edgeR[5]. For RNA-seq of mouse B cells, Mouse gene symbols were converted to human counterparts, and those without corresponding human symbols were omitted for GSEA analysis. Gene sets with a FDR < 0.25 were defined as significantly enriched. Molecular Signatures Database v7.2 (c7: immunologic signature gene sets) and lymphoma signature database (https://lymphochip.nih.gov/signaturedb/) of the Staudt laboratory [6] were used in GSEA analysis.

CUT&RUN

CUT&RUN experiments were performed following the protocol as described with modifications [7,8]. Cells were fixed with 10 mL 1% formaldehyde in PBS at room temperature for 10min, and resuspended with wash buffer (20 mM HEPES-KOH pH 7.9, 10 mM NaCl, 1 mM spermidine, 0.02% Digitonin, and 1× protease inhibitor cocktails). The cells were captured with Concanavalin A beads (BioMagPlus) and incubated with 10 uL antibodies in 200 µL antibody buffer (20 mM HEPES-KOH pH 7.9, 10 mM NaCl, 1 mM spermidine, 0.02% Digitonin, 2 mM EDTA, and 1× protease inhibitor cocktails) overnight at 4 ℃. Then, cells were washed with 1mL antibody buffer twice. Then pA/G-MN (in house) was added with final concentration at 700 ng/uL to 200 µL antibody buffer and incubated at 4 ℃ for 1 hour. Cells were washed again and resuspended in 100 µL wash buffer. $CaCl_2$ was added at a final concentration of 2 mM to activate the enzyme. The reaction was carried out on ice for 30min and stopped by 100 µL of 2X STOP buffer (340 mM NaCl, 20 mM EDTA, 4 mM EGTA, 0.02% Digitonin, 100 µg/mL RNase A, and 50 µg/mL glycogen). The released protein-DNA complex was digested by proteinase K at 65 ℃ overnight and purified by Phenol-Chloroform-Isoamyl method. Antibody targeting OBF1 and H3K27ac were 33483 (Cell Signaling Technology) and 39133 (Active Motif), respectively.

Library Preparation and Sequencing for CUT&RUN

Sequencing library for CUT&RUN DNA using NEBNext Ultra II DNA Library Prep Kit aimed at preserve short DNA fragments (30-80 bp) was described [9]. Detailed stepwise protocol can be found at protocol.io (https://doi.org/10.17504/protocols.io.wvgfe3w). In general, at end repair step, the CUT&RUN DNA was incubated at 50 ℃ for 1 h in order to recover short DNA. At ligation step, 2 µM NEB adapter was used to reduce the level of dimer formation. After ligation, 1.75× volume AMPure XP beads (Beckman Coulter) was applied to remove large fragments. PCR amplification was performed for 10 cycles (this step needs to be optimized). The resulting libraries were purified by a sequential 1.1× (0.8× + 0.3×) addition of

AMPure beads. Indexed libraries were pooled and paired-end sequencing was performed using Nextseq platform with NextSeq High Output (75 cycles) (2 × 38 bp, 6-bp index) protocol.

ChIP-seq data processing

ChIP-seq and BioChIP-seq sequencing reads were aligned to the mouse genome (mm9) using the qAlign function, which internally uses Bowtie [10], from Bioconductor package QuasR [11] (package version 1.16.0). Alignments were shifted by 60 bases, corresponding to an estimated fragment length of 120 bp. Peaks of OCT1, OCT2, OBF1, PU.1, H3K27ac, H3K4me3 and H3K4me1 were called in control sample (input) with MACS2 [12] (software version 2.1.1) using default parameters. Reads were counted in each peak region using the qCount function in QuasR (with shift set to 60). The read counts per sample were normalized to the mean total mapped number of reads across all samples. The fold enrichment over input was calculated using the following formula:

$$\log2(((nIP / tIP)*10^6+8) / ((nInput / tInput)*10^6+8)),$$

where nIP and nInput are the read counts in a region in ChIP and input samples, respectively, tIP and tInput are the total number of mapped reads in the ChIP and input samples, and 8 is a pseudocount to stabilize enrichments that are based on low numbers of reads.

We generated two replicates for OCT1, OCT2, OBF1, PU.1, H3K27ac, H3K4me1 and H3K4me3 in wild type $CD19^+$ mature B cells treated with LPS or anti-CD40/IL4. For comparing OCT1 ChIP-seq signals between wild type and *Pou2af1* knockout mice, we generated two replicates of OCT1 and OBF1 in wild type B cells, and two replicates of OCT1 in mutant mature B cells. For H3K27ac, we generated one replicate for both wild type and mutant mature B cells. Peaks were assigned to genes by calculating distance of peak summits to the nearest transcriptional start site (TSS) allowing for a maximum distance of 500 kb, using a set of non-redundant TSSs with a single start site randomly selected for each gene. Peaks with a distance to the nearest TSS of less than 1 kb were classified as proximal and those with a distance greater than 1 kb as distal.

Heatmaps centred on TF peak summits were generated with the qProfile function from the Bioconductor package QuasR [11] (package version 1.16.0), and visualized using ComplexHeatmap (https://bioconductor.org/packages/release/bioc/manuals/ComplexHeatmap/man/ComplexHeatmap.pdf) (package version 1.18.1).

Differentially bound regions were identified with glmQLFit/glmQLFTest method in the R package edgeR [5] (https://www.bioconductor.org/packages/release/bioc/vignettes/edgeR/inst/doc/edgeRUsersGuide.pdf) (package version 3.26.8). Raw P values were corrected for multiple testing by calculating false

discovery rates (FDRs). Significant changes in each contrast were defined as changes with a fold-change of 2 and an FDR less than 0.01.

TF motif discovery was performed using the HOMER software (version 4.8.2) using the following command (findMotifsGenome.pl input_list mm9 output_directory -size 1000). The size of the regions used to search for motifs was set to 1 kb. The lists of TF motifs identified were filtered by expression levels. We searched for motifs de novo using the MEME-suite [13] (version 5.0.5) within 31bp around peak summits. The occurrences of the enriched motifs were then counted in an area of 500bp around each peak summit using TFBSTools (https://bioconductor.org/packages/release/bioc/manuals/TFBSTools/man/TFBSTools.pdf) (package version 1.22.0) from Bioconductor. Shuffled DNA sequence with preserved dinucleotide frequency generated by the fasta-dinucleotide-shuffle command from MEME (version 4.12.0) were used as control. ETS motifs were scanned within 500bp window centered on OBF1 peak summits were generated by TFBSTools (package version 1.22.0) from Bioconductor. ETS motifs were binding motifs of ETS factors that were expressed in mature B cells filtered by our RNA-seq data. The position weight matrices of these motifs were extracted from JASPAR2018 (https://bioconductor.org/packages/release/data/annotation/manuals/JASPAR2018/man/JASPAR2018.pdf) (package version 3.9) from Bioconductor using the getMatrixByID function. Control position weight matrices were randomly generated 500 times by shuffling the columns of each original position weight matrix. P values for each ETS weight matrix over random controls were calculated by comparing 95th-percentile counts (corresponding to the strong enrichment observed at about 50bp from the octamer motif) between randomized and real motif frequencies using the following formula:

$$P = (\text{sum}(95\text{th-random-count} > 95\text{th-motif-count}) + 1) / (500 + 1)$$

RNA-seq and data processing

Cellular RNA from 5×10^5 $CD19^+$ mature B cells 72 hours post LPS and anti-CD40/IL4 treatment were purified using RNeasy Micro kit (QIAGEN), strand-specific RNA-seq libraries were generated using TruSeq mRNA preparation kit (Illumina), and libraries were sequenced on an Illumina HiSeq2500 machine (50bp single-ended reads). Three replicates were generated for each condition. Sequencing reads were aligned using qAlign from the Bioconductor package QuasR [11] (package version 1.16.0) to the mouse genome (mm9) with default parameters except for splicedAlignment = TRUE. Alignments were quantified with qCount from the Bioconductor package QuasR [11] (package version 1.16.0) for known UCSC genes

obtained from the TxDb.Mmusculus.UCSC.mm9.knownGene package (package version 3.2.2) using default parameters.

To identify genes differentially expressed in either LPS or anti-CD40/IL4 treatment, we used generalized linear models. Genes with at least 5 counts per million in at least two biological samples were considered. Statistically significantly differentially expressed genes were identified using glmFit/glmLRT method in edgeR (https://www.bioconductor.org/packages/release/bioc/vignettes/edgeR/inst/doc/edgeRUsersGuide.pdf) (package version 3.26.8). *P* values were corrected for multiple testing by FDRs. Significant effects in each contrast were defined as changes with absolute fold-change of 2 and FDR less than 0.01. Contrast group: anti-CD40/IL4 compared to LPS.

For the OBF1 knockout experiment, cellular RNA from anti-CD40/IL4 stimulated wild type and OBF1 knockout murine cells were prepared with RNeasy Micro kit (QIAGEN), and strand-specific RNA-seq libraries were generated using the TruSeq mRNA preparation kit (Illumina). The libraries were sequenced on the Illumina HiSeq2500 system (50bp single-ended reads). Three replicates were generated for each condition. Sequencing reads were aligned using qAlign from the Bioconductor package QuasR [11] (package version 1.16.0) to the mouse genome (mm9) with default parameters except for splicedAlignment = TRUE. Alignments were quantified with qCount from the Bioconductor package QuasR [11] (package version 1.16.0) for known UCSC genes obtained from the TxDb.Mmusculus.UCSC.mm9.knownGene package (package version 3.2.2) using default parameters.

To identify genes differentially expressed OBF1 knockout murine cells, we used generalized linear models. Genes with at least 5 counts per million in at least two biological samples were considered. Statistically significantly differentially expressed genes were identified using glmFit/glmLRT method in edgeR (https://www.bioconductor.org/packages/release/bioc/vignettes/edgeR/inst/doc/edgeRUsersGuide.pdf) (package version 3.26.8). *P* values were corrected for multiple testing by FDRs. Significant effects in each contrast were defined as changes with absolute fold-change of 2 and FDR less than 0.01. Contrast group: OBF1 KO compared to wild type.

For the OBF1 knockdown experiment, cellular RNA from wild type and OBF1 knockdown Raji cells were prepared with RNeasy Micro kit (QIAGEN), and strand-specific RNA-seq libraries were generated using the TruSeq mRNA preparation kit (Illumina). The libraries were sequenced on the Illumina HiSeq2500 system (50bp single-ended reads). Three replicates were generated for each condition. Sequencing reads were aligned using qAlign from the Bioconductor package QuasR [11] (package version 1.16.0) to the

human genome (hg19) with default parameters except for splicedAlignment = TRUE. Alignments were quantified with qCount from the Bioconductor package QuasR [11] (package version 1.16.0) for known UCSC genes obtained from the TxDb.Hsapiens.UCSC.hg19.knownGene package (package version 3.2.2) using default parameters.

To identify genes differentially expressed OBF1 knockdown Raji cells, we used generalized linear models. Genes with at least 5 counts per million in at least two biological samples were considered. Statistically significantly differentially expressed genes were identified using glmFit/glmLRT method in edgeR (https://www.bioconductor.org/packages/release/bioc/vignettes/edgeR/inst/doc/edgeRUsersGuide.pdf) (package version 3.26.8). *P* values were corrected for multiple testing by FDRs. Significant effects in each contrast were defined as changes with absolute fold-change of 1.5 and FDR less than 0.01. Contrast group: Raji OBF1shRNA compared to Raji SHC002.

For the IRF4 overexpression experiments, cellular RNA from wild type and IRF4 overexpressed Raji cells were prepared with RNeasy Micro kit (QIAGEN), and strand-specific RNA-seq libraries were generated using the TruSeq mRNA preparation kit (Illumina). The libraries were sequenced on the Illumina HiSeq2500 system (50bp single-ended reads). Three replicates were generated for each condition. Sequencing reads were aligned using qAlign from the Bioconductor package QuasR [11] (package version 1.16.0) to the human genome (hg19) with default parameters except for splicedAlignment = TRUE. Alignments were quantified with qCount from the Bioconductor package QuasR [11] (package version 1.16.0) for known UCSC genes obtained from the TxDb.Hsapiens.UCSC.hg19.knownGene package (package version 3.2.2) using default parameters.

To identify genes differentially expressed in IRF4 overexpressed Raji cells, we used generalized linear models. Genes with at least 5 counts per million in at least two biological samples were considered. Statistically significantly differentially expressed genes were identified using glmFit/glmLRT method in edgeR (https://www.bioconductor.org/packages/release/bioc/vignettes/edgeR/inst/doc/edgeRUsersGuide.pdf) (package version 3.26.8). *P* values were corrected for multiple testing by FDRs. Significant effects in each contrast were defined as changes with absolute fold-change of 2 and FDR less than 0.01. Contrast group: Raji IRF4OE compared to Raji ctrl.

CUT&RUN data processing

The data processing was carried out following the procedure and parameters that has been described previously [14]. For alignment, Bowtie2 version 2.2.5 [15] was used with following parameters: --local --very-

sensitive-local --no-unal --no-mixed --no-discordant --phred33 -I 10 -X 700. For MACS2 [12] peak calling, parameters: macs2 callpeak –p 1e-5 –f BEDPE –keep-dup all.

Supplemental methods references

1. Cato MH, Yau IW, Rickert RC. Magnetic-based purification of untouched mouse germinal center B cells for ex vivo manipulation and biochemical analysis. *Nat Protoc*. 2011;6(7):953-960.
2. Yang H, Wang H, Jaenisch R. Generating genetically modified mice using CRISPR/Cas-mediated genome engineering. *Nat Protoc*. 2014;9(8):1956-1968.
3. Mootha VK, Lindgren CM, Eriksson KF, et al. PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat Genet*. 2003;34(3):267-273.
4. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A*. 2005;102(43):15545-15550.
5. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*. 2010;26(1):139-140.
6. Shaffer AL, Wright G, Yang L, et al. A library of gene expression signatures to illuminate normal and pathological lymphoid biology. *Immunol Rev*. 2006;210:67-85.
7. Skene PJ, Henikoff S. A simple method for generating high-resolution maps of genome-wide protein binding. *Elife*. 2015;4:e09225.
8. Skene PJ, Henikoff JG, Henikoff S. Targeted in situ genome-wide profiling with high efficiency for low cell numbers. *Nat Protoc*. 2018;13(5):1006-1019.
9. Liu N, Hargreaves VV, Zhu Q, et al. Direct Promoter Repression by BCL11A Controls the Fetal to Adult Hemoglobin Switch. *Cell*. 2018;173(2):430-442 e417.
10. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol*. 2009;10(3):R25.
11. Gaidatzis D, Lerch A, Hahne F, Stadler MB. QuasR: quantification and annotation of short reads in R. *Bioinformatics*. 2015;31(7):1130-1132.
12. Zhang Y, Liu T, Meyer CA, et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol*. 2008;9(9):R137.
13. Bailey TL, Williams N, Misleh C, Li WW. MEME: discovering and analyzing DNA and protein sequence motifs. *Nucleic Acids Res*. 2006;34(Web Server issue):W369-373.
14. Meers MP, Bryson TD, Henikoff JG, Henikoff S. Improved CUT&RUN chromatin profiling tools. *Elife*. 2019;8.
15. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 2012;9(4):357-359.

Supplementary Figure legends

Supplemental Figure 1 (related to Fig. 1 and Fig. 2). OCT1, OCT2 and OBF1 show genome-wide colocalization.

A. Schematic strategy for targeted integration at the endogenous gene locus in order to create N- or C-termal tagged proteins. For the scenario of 3'-insertion (C-terminal targeting; C-term), a double strand break (DSB) on DNA was introduced right before the stop codon using specific sgRNA sequences and Cas9 protein. Synthesized single-stranded oligo DNA nucleotides (ssODN, less than 200nt in length) were used as donor DNA for the site-specific integration of AviTag-FLAG sequence into the desired positions. In the case of 5'-insertion (N-terminal targeting; N-term), a specific DSB was introduced right after the start codon, and the AviTag-FLAG sequence was inserted in a similar manner.

B. Scatter plots showing the enrichment of ChIP-seq signals of OCT1, OCT2 and OBF1 over Input at promoter regions under LPS (upper panel) and anti-CD40/IL4 (lower panel) stimulation.

C. Left, Venn diagrams showing the overlaps between OBF1 Bio-ChIP-seq and antibody ChIP-seq (sc955); right, correlation of enrichment between OBF1 Bio-ChIP-seq and antibody ChIP-seq.

D. Left, Venn diagrams showing the overlaps between OBF1 Bio-ChIP-seq and antibody ChIP-seq (CST33483); right, correlation of enrichment between OBF1 Bio-ChIP-seq and antibody ChIP-seq.

E. Genome browser snapshots showing the binding peaks of OBF1/OCT1/OCT2 at their known target genes ($CD19^+$ spleen B cells treated with LPS for 72 hours).

F. Barplot plots showing expression levels of *Pou2f1* and *Pou2f2* in splenic B cells stimulated with LPS, measured by RNA-seq. (mean ± s.d).

G. Violin plots showing distributions of enrichment of H3K27ac on the promoters of genes grouped by their association with OCT2 (left), OCT1 (middle) or OBF1 (right) in $CD19^+$ splenic B cells stimulated with LPS. (*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; ****$p < 0.0001$; Mann-Whitney U test)

H. Stacked barplots showing the occupancy of OBF1, OCT1 and OCT2 on the promoters of active genes (left) and inactive genes (right).

Supplemental Figure 2 (related to Fig. 3). OCT1, OCT2 and OBF1 colocalize with TFs of the ETS family.

A. OCT1, OBF1, OCT2 and PU.1 ChIP enrichments in 10-kb windows centered on peak summits, separately for OCT1 peak regions overlapping or non-overlapping with PU.1. All samples in this Figure were LPS-stimulated $CD19^+$ mouse splenic B cells.

B. Venn diagrams showing overlap of peaks between OCT1/ETS1, OCT2/ETS1 and OBF1/ETS1 (ChIP-seq were performed with spleen B cells treated with LPS; ETS1 data were downloaded from GSE83797).

C. Enrichment of ETS1 binding motifs (MA0098.2 in JASPAR database) within 500-bp windows in OCT1, OBF1 and OCT2 peak regions centred around peak summits under LPS treatment. Controls are shuffled peak sequences that retain dinucleotide frequency.

D. The enrichment of the ETS motifs of various ETS factors are visualized. ETS factors were selected based on their expression levels in our RNA-seq data. Black lines show the occurrences of original motifs within 500 bp around octamer motifs found in OBF1 peaks. Grey lines show the occurrences of randomly shuffled ETS motifs as backgrounds. Each of the original ETS motifs were randomized 500 times. After randomization, the *P*-value was calculated by comparing the 95th percentile of the occurrences of original motifs with the 95th percentiles of all randomly shuffled motifs (500 times) for each ETS motif.

Supplemental Figure 3 (related to Fig. 4). OBF1 stabilizes the binding of OCT1 to chromatin.

A. Read counts correlation between overlapping regions of OBF1 and OCT2 ChIP-seq samples from LPS-stimulated CD19^{+} mouse spleen B cells.

B. Violin plots showing enrichment of OCT2 ChIP samples from LPS-stimulated CD19^{+} mouse spleen B cells according to their overlap with OBF1 and/or PU.1.

Panel B: mean ± s.d; *p < 0.05; **p < 0.01; ***p < 0.001; ****p < 0.0001; Mann-Whitney U test.

C. Immunoblot to monitor the expression level of OCT1 and OCT2 in OBF1 KO murine CD19^{+} B cells stimulated with LPS.

Supplemental Figure 4 (related to Fig. 5). OBF1 regulates the GC transcriptional program.

A. Left, scatter plot showing differentially bound regions of OBF1 between LPS and anti-CD40/IL4 stimulations; right, *de novo* motif analysis of regions differentially bound by OBF1 under LPS and anti-CD40/IL4 stimulation using HOMER.

B. Left, scatter plot showing differentially bound regions of OCT1 between LPS and anti-CD40/IL4 stimulations. Right, de novo motif analysis of differentially bound regions of OCT1 under LPS and anti-CD40/IL4 stimulation using HOMER.

C. Left, scatter plot showing differentially bound regions of OCT2 between LPS and anti-CD40/IL4 stimulations. Right, de novo motif analysis of differentially bound regions of OCT2 under LPS and anti-CD40/IL4 stimulation using HOMER.

D. OCT1, OCT2, OBF1 ChIP-seq read densities at gene loci involved in GC reaction ($CD19^+$ spleen B cells treated with anti-CD40/IL4 for 72 hours).

E. OCT1, OCT2, OBF1 ChIP-seq read densities at *Foxo1*, *Mef2b*, *Spi1* in splenic B cell treated with anti-CD40/IL4 or LPS for 72 hours.

F. Upper panel, workflow of germinal centers induction by intraperitoneal injection of 10% SRBCs solution and the isolation of GC B cells for FACS analysis. Lower panel, FACS plots for GC B cells identification, at day 7 following immunization of wild type and *Pou2af1*$^{KO/KO}$ mice.

Supplementary Figure 5 (related to Fig. 6). OBF1 is required for the proliferation of GC-derived lymphoma cell lines

A. Left, quantitative reverse-transcription PCR (qRT-PCR) showing OBF1 knockdown following shRNA-mediated depletion of OBF1 (OBF1-shRNA1) or control (SHC002) (n=3); right, assessment of proliferation in Ramos cell line following shRNA-mediated depletion of OBF1 (n=3). Cells were seeded at day 1, and counted every second day until day 7 (n = 3).

B. Left, qRT-PCR showing OBF1 knockdown following shRNA-mediated depletion of OBF1 (OBF1-shmix) or control (SHC002) (n=3); right, assessment of proliferation of Daudi cell line following shRNA-mediated depletion of OBF1. Cells were seeded at day 1, and counted every second day until day 7 (n = 3).

C. Left, qRT-PCR showing OBF1 knockdown following shRNA-mediated depletion of OBF1 (OBF1-shRNA1) or control (SHC002) (n=3); right, assessment of proliferation of HT cell line following shRNA-mediated depletion of OBF1. Cells were seeded at day 1, and counted every second day until day 7 (n = 3).

D. qRT-PCR measurements of OCT1, OCT2, OBF1 and BCL6 expression upon shRNA-mediated knockdown of OCT1 (n=3).

E. qRT-PCR measurements of OCT1, OCT2, OBF1 and BCL6 expression upon shRNA-mediated knockdown of OBF1 (same data as shown in Fig. 6F) (n=3).

F. qRT-PCR measurements of OCT1, OCT2, OBF1 and BCL6 expression upon shRNA-mediated knockdown of OCT2 (n=3).

G. Left, immunoblot confirming the ectopic expression of BCL6 in combination with shRNA-mediated OBF1 knockdown. Right, assessment of proliferation in Raji cell line following shRNA-mediated OBF1 knockdown and ectopic expression of BCL6.

H. Left, confirmation of downregulation of OBF1 by shRNA specific for OBF1 by quantitative qRT-PCR in SUDHL4 cells; right, confirmation of downregulation of IRF4 by shRNA specific for IRF4 by qRT-PCR in SUDHL4 cells (n=3).

I. Assessment of SUDHL4 cells proliferation following shRNA-mediated depletion of OBF1 and depletion of both OBF1 and IRF4 (n=3).

mean ± s.d; *p < 0.05; **p < 0.01; ***p < 0.001; ****p < 0.0001; two-tailed Student's t test.

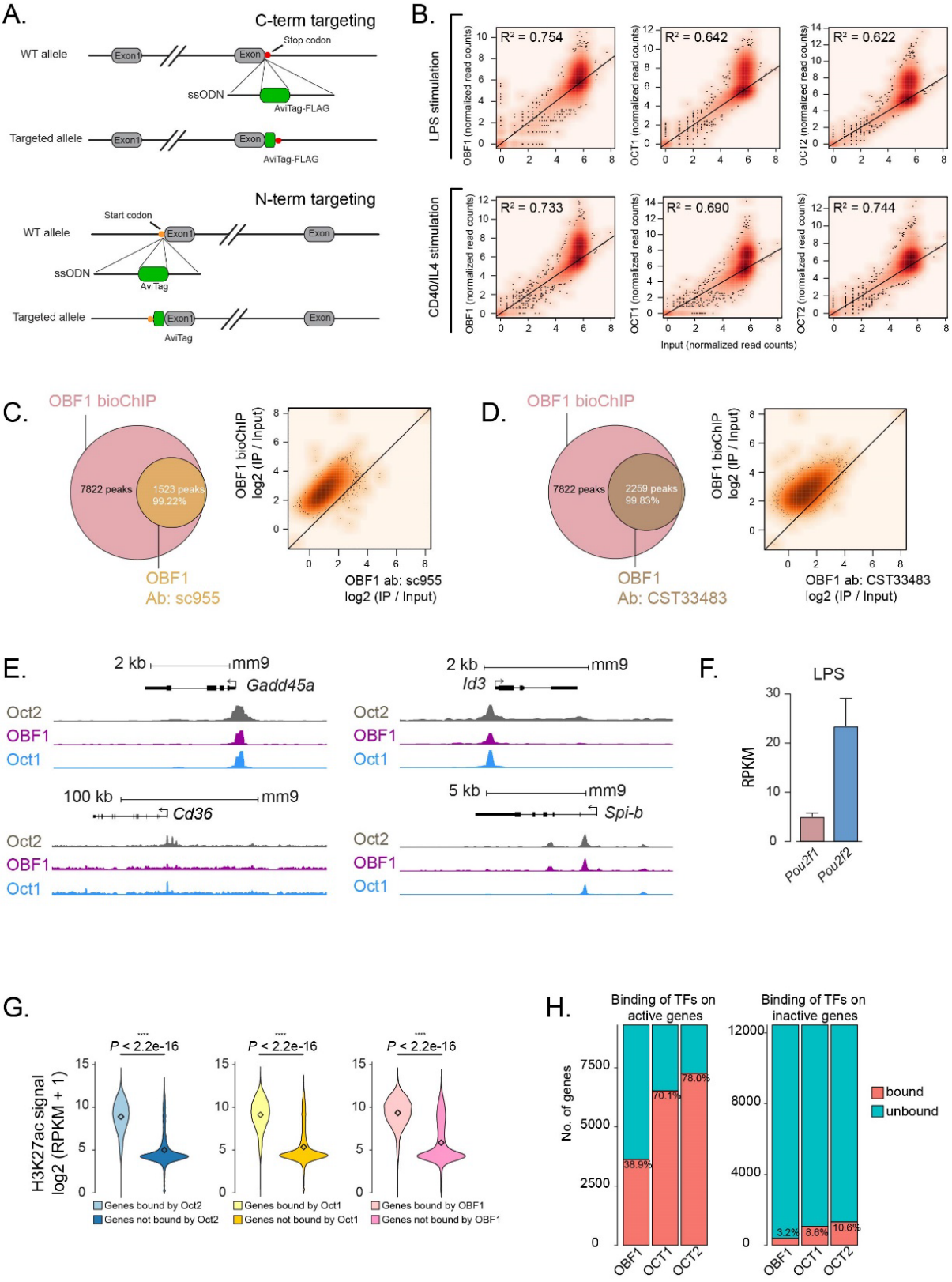
Supplementary Figure 6 (related to Fig. 7). OBF1 regulates proliferation of GC-derived lymphoma cell line via modulating the level of IRF4

A. GSEA of relative gene expression in OBF1-depleted versus wild type Raji cells against the gene set identified as genes activated by IRF4 and genes repressed by BCL6, respectively. FDR, false discovery rate; NES, normalized enrichment score.

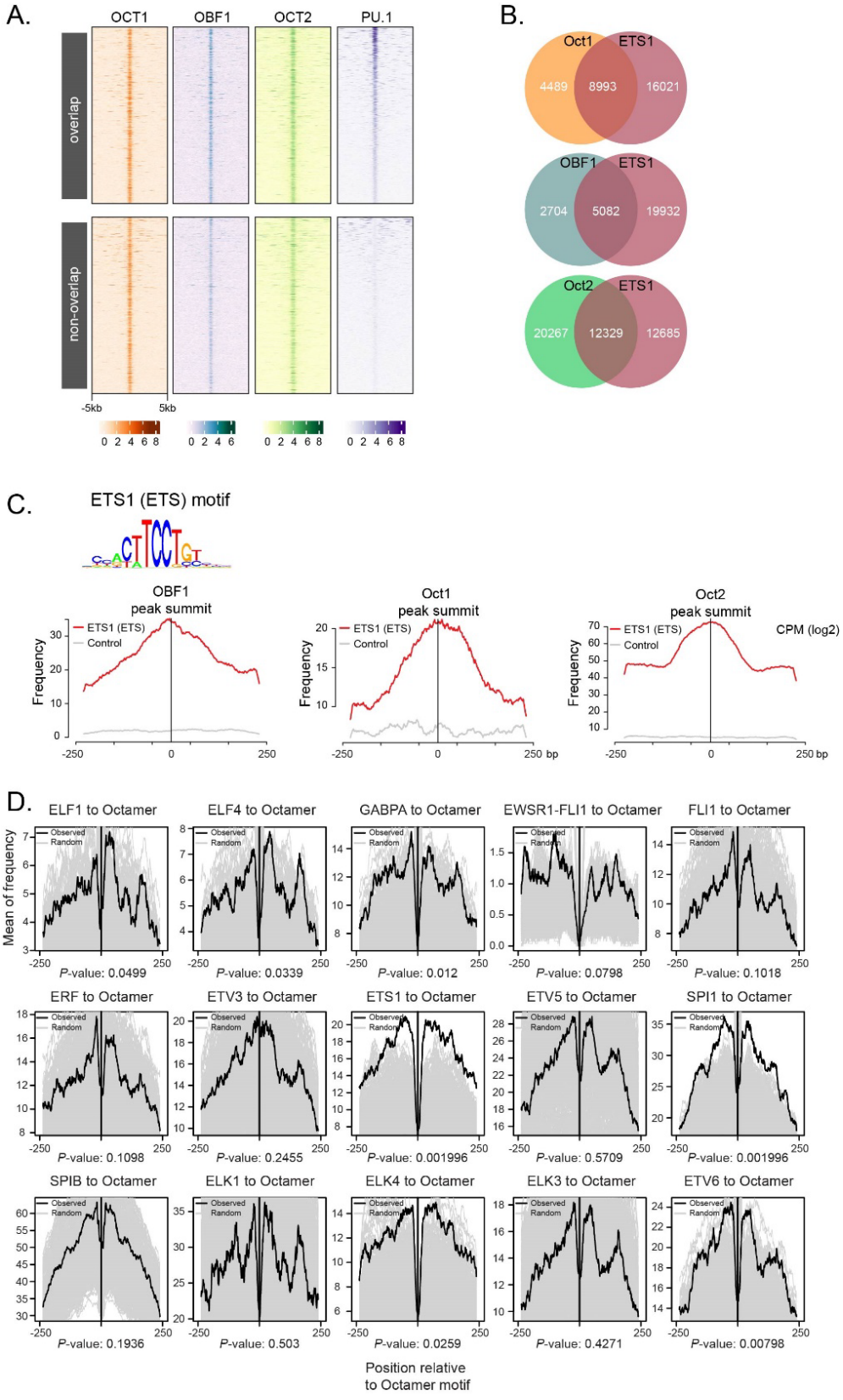
B. OBF1 CUT&RUN read densities at four individual loci in Raji and SUDHL4 cells.

C. H3K27ac and OBF1 CUT&RUN read densities at four individual loci, as indicated, in purified germinal center B cells from human tonsils.

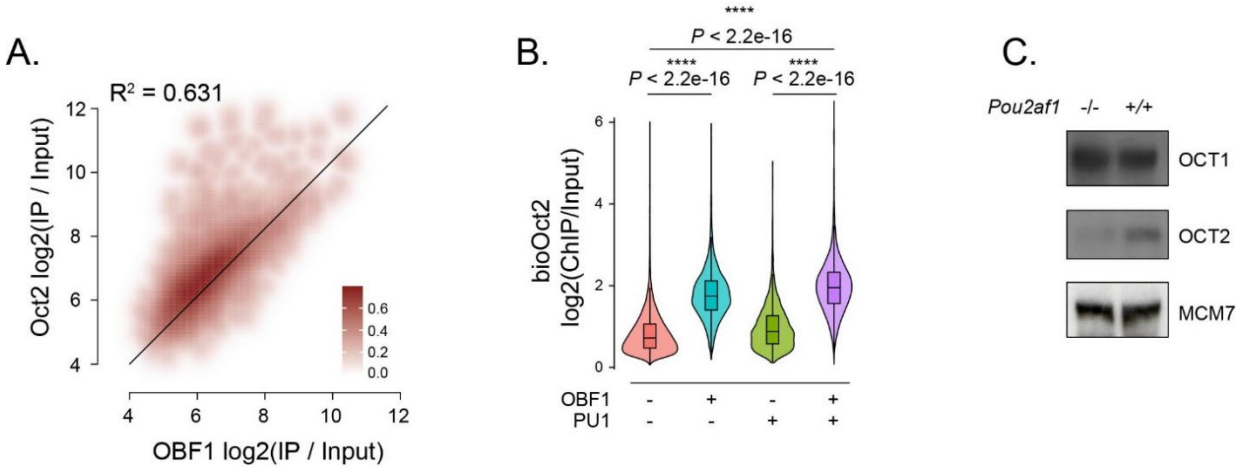
Supplemental Figure 1



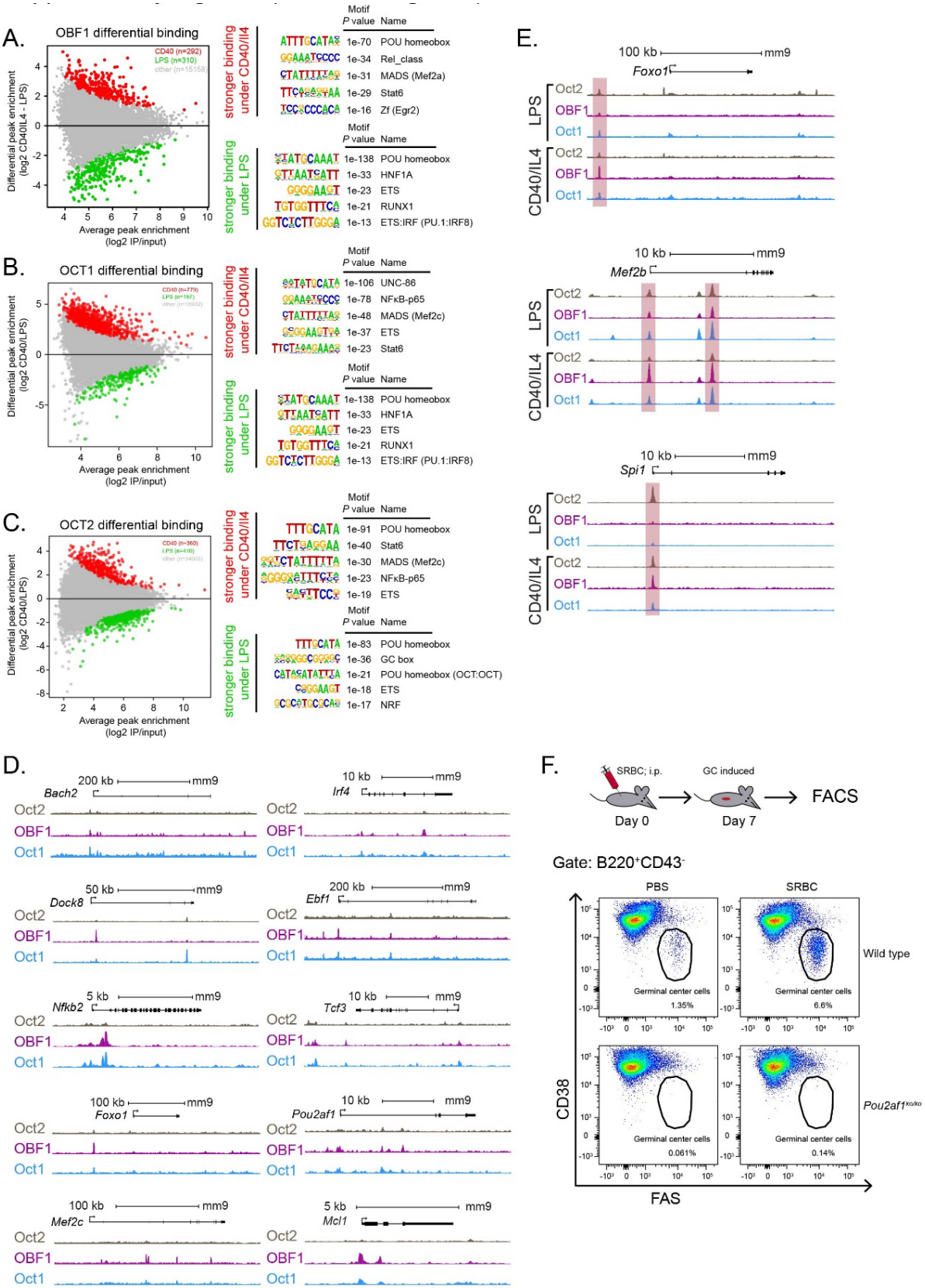
Supplemental Figure 2



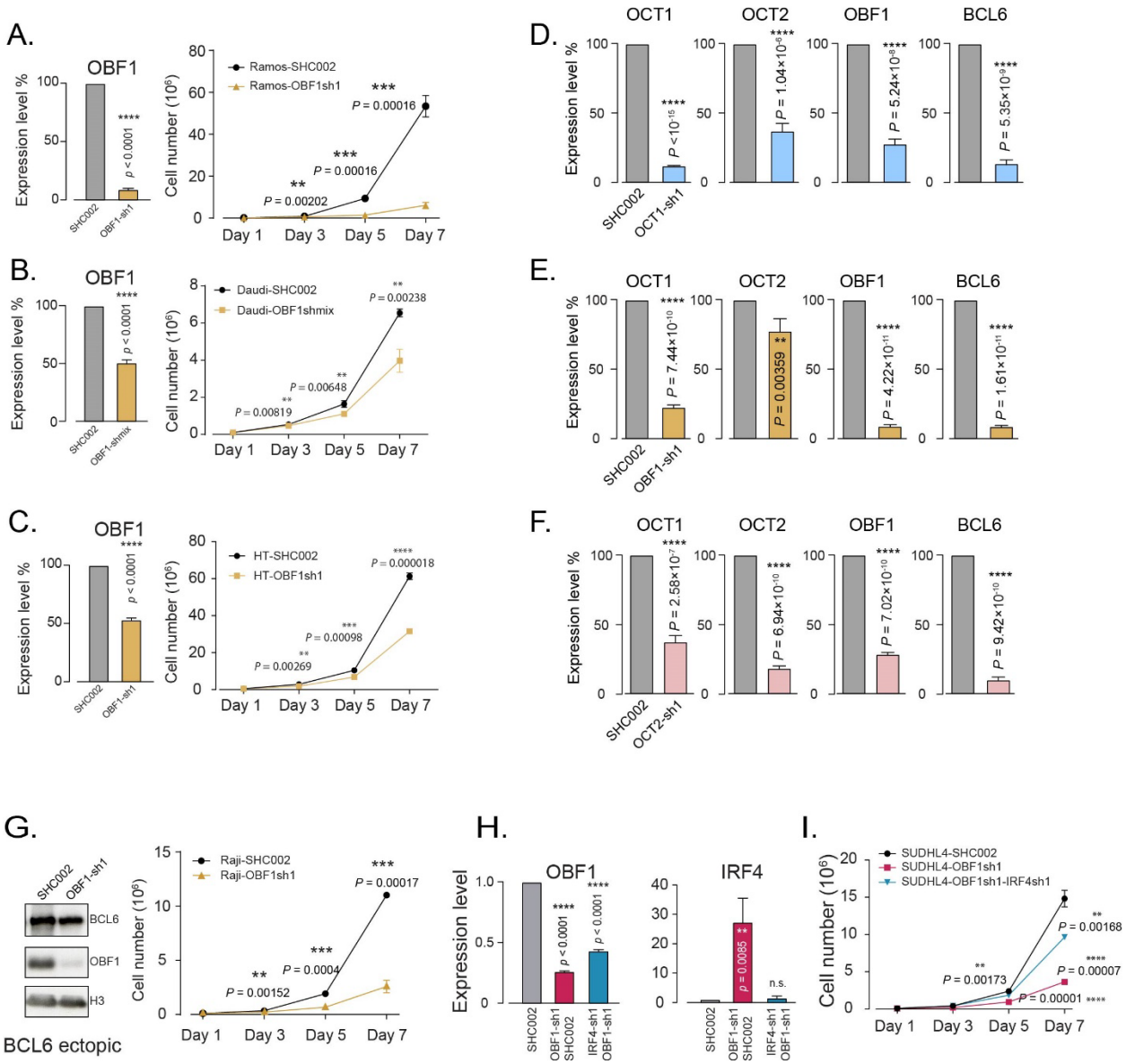
Supplemental Figure 3



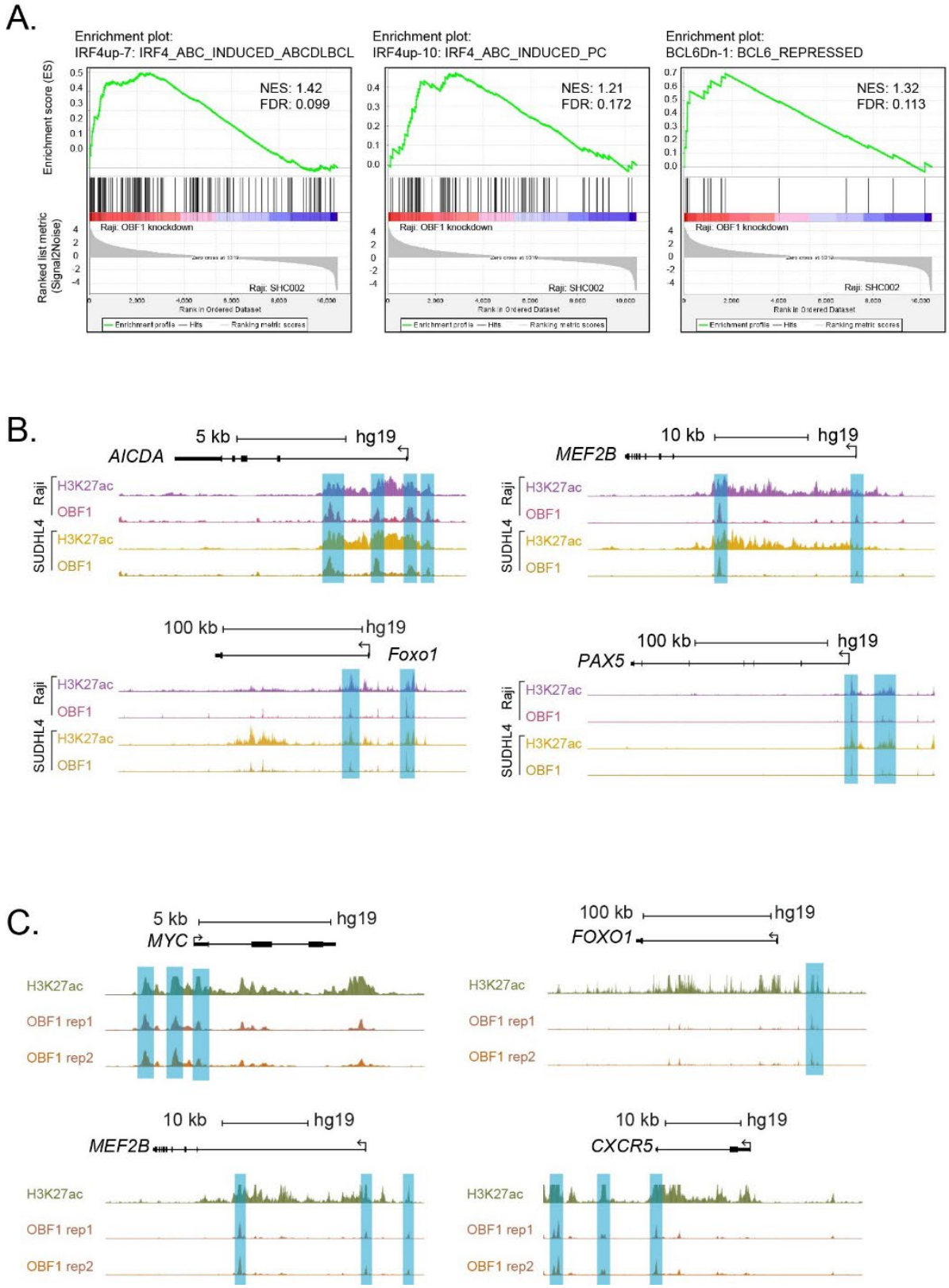
Supplemental Figure 4



Supplemental Figure 5



Supplemental Figure 6



Supplemental Table 1. Antibodies

Target	Provider	Catalog number	Clone
H3K4me3	Millipore	07-473	polyclonal
H3K4me1	Abcam	ab8895	polyclonal
H3K27ac	Active Motif	39133	polyclonal
PU.1	Cell Signaling Technology	2258	9G7
OCT2	Abcam	ab179808	EPR12482
OBF1	Cell Signaling Technology	33483	polyclonal
OBF1	Santa Cruz	sc955	C-20
OCT1	Cell Signaling Technology	8157	D7B6
BCL6	Cell Signaling Technology	14895	D4I2V
CD40	BD Biosciences	553721	HM40-3
Histone 3	Abcam	ab1791	polyclonal
MCM7	Abcam	ab2360	47DC141
CD19-PE-cy7	BD Biosciences	557835	SJ25C1
CD38-PE	BD Biosciences	555460	HIT2
IgD-BV605	BD Biosciences	563313	IA6-2
IgG	Sigma-Aldrich	I5006-10MG	

Supplemental Table 2. Synthesized gRNA strands for generating transgenic mice

Gene	Species	Model	gRNA
Pou2af1	Mus musculus	C57BL/6J	5'-CCTCTCCGTGGAGGGCTTTT-3'
Pou2f1	Mus musculus	C57BL/6J	5'-CTCAGCTCACTGTGCCTTGG-3'
Pou2f2	Mus musculus	C57BL/6J	5'-GCCCTGGCTGGCGGGCAGCA-3'

Supplemental Table 3. ssODNs for generating transgenic mice

Gene	Sequence	Tag
Pou2f1	5'-CACAGTCGCCTCTGCCAGTGGGCCTGCTTCCACCACCACAGCTGCATCCAAGGCACAGGGCGGCCTGAACGACATCTTCGAGGCTCAGAAAATCGAATGGCACGAAGGTGACTACAAAGACGATGACGATAAATGAGCTGAGTGCAGAGCTGGGTTGCCACAGGCTTTCCTCACTACA-3'	AviTag-FLAG
Pou2f2	5'-GGCGGGGAGATGACACAGTTGTTCCCCAGCCCTGGCTGGCGGGCAGCATGATGGGCCTGAACGACATCTTCGAGGCTCAGAAAATCGAATGGCACGAAGTTCATTCCAGCATGGGGGCTCCAGGTAAGAGGCTGATGCCTTCT-3'	AviTag
Pou2af1	5'-AACACGTACGAGCTCAACCACACCCTCTCCGTGGAGGGCTTTGGCGGCCTGAACGACATCTTCGAGGCTCAGAAAATCGAATGGCACGAAGGTGACTACAAAGACGATGACGATAAATGAGGCTGGCTTGCATCTAACAGATGTTTCACCCATAGCTGAG-3'	AviTag-FLAG

Supplemental Table 4. Genotyping primers for transgenic mice lines

Gene	Forward	Reverse	Species
Pou2af1	5'-TCCCTGACCATTGACAAGCT-3'	5'-TTGCTTCCCCAATGATGTGC-3'	Mus musculus
Pou2f1	5'-CTGTTCTGGGCTCTTTGTCT-3'	5'-AGGAAGCCAATGACAGGTGT-3'	Mus musculus
Pou2f2	5'-ACACTCATGATCTGGGGTCG-3'	5'-ACACACACCCCATGCCTAAT-3'	Mus musculus
Rosa26-BirA	5'-GGAGAGCTTAAATCGGGCGA-3'	5'-CCCTGTTTGTCTATTCCGCG-3'	Mus musculus
Rosa26-WT	5'-GTGTAACTGTGGACAGAGGAG-3'	5'-GGAGAGCTTAAATCGGGCGA-3'	Mus musculus

Supplemental Table 5. Gene specific shRNAs

Gene/control	Sequence	Identifier	Species
POU2AF1	5'-GGGAGGTAATTATAGGGATTT-3'	TRCN0000423580	Homo sapiens
POU2F1	5'-ACAACACAGCAACCGTGATTT-3'	TRCN0000232119	Homo sapiens
POU2F2	5'-GCTACCGACACCAAATCTATT-3'	TRCN0000245324	Homo sapiens
IRF4	5'-CCAGCAGGTTCACAACTACAT-3'	TRCN0000014766	Homo sapiens
Control	5'-CAACAAGATGAAGAGCACCAA-3'	SHC002	Homo sapiens
POU2AF1 mix	5'-GGGAGGTAATTATAGGGATTT-3' 5'-TGGACACCTTACACCGAGTAT-3'	TRCN0000423580 TRCN0000021679	Homo sapiens

Supplemental Table 6. Oligos used for quantitative PCR

Gene	Forward	Reverse	Species
POU2AF1	5'-CACCAATGTCACGACAAGAAGC-3'	5'-ACGGGAAATAGGTGAGGGGT-3'	Homo sapiens
POU2F1	5'-ATGAACAATCCGTCAGAAACCAG-3'	5'-GATGGAGATGTCCAAGGAAAGC-3'	Homo sapiens
POU2F2	5'-GAGGAGCCCAGTGATCTGGA-3'	5'-GAAGCGGGAAATGGTCGTC-3'	Homo sapiens
ACTB	5'-AGCCTCGCCTTTGCCGATC-3'	5'-AGCGGGCGATATCATCATCC-3'	Homo sapiens
BCL6	5'-CGCAACTCTGAAGAGCCACCTGCG-3'	5'-TTTGTGACGGAAATGCAGGTTA-3'	Homo sapiens
IRF4	5'-ACAGCAGTTCTTGTCAGAG-3'	5'-GAGGTTCTACGTGAGCTG-3'	Homo sapiens

Supplemental Table 7. Genes associated with differentially regions bound by OCT1 between WT and *Pou2af1*$^{-/-}$ B cells

Genes associated with stronger peak in WT B cells							Genes associated with stronger peak in OBF1KO B cells
Vxn	*Ebf1*	*Ankrd55*	*Tbc1d5*	*Kcnmb3*	*Urad*	*Mmp15*	*Lonrf2*
Sgk3	*Clint1*	*Emb*	*Fem1a*	*Foxo1*	*Slc7a1*	*Zfhx3*	*Hdac4*
Mcmdc2	*Fam71b*	*Pxk*	*Vav1*	*Cog6*	*Ubl3*	*Marveld3*	*Dbi*
Slco5a1	*Rufy1*	*Oit1*	*Fer*	*Eif2a*	*N4bp2l1*	*Vac14*	*Parp1*
Tram1	*Jade2*	*Kcnk5*	*Tgif1*	*Selenot*	*C1galt1*	*Wwox*	*Wdr26*
Tram2	*Il4*	*Zmiz1*	*Lpin2*	*E130311K13Rik*	*Umad1*	*Maf*	*1700113H08Rik*
Cnga3	*Cnot8*	*Slmap*	*Smchd1*	*Tiparp*	*Glcci1*	*Gse1*	*Jmjd4*
Mgat4a	*Pik3r5*	*Tasor*	*Fam98a*	*Lekr1*	*Bmt2*	*Irf8*	*Per1*
Rev1	*Pik3r6*	*Prkcd*	*Atl2*	*Tmem131l*	*St7*	*Trappc2l*	*Polr2a*
Cd28	*Spns3*	*Sfmbt1*	*Haao*	*Tmem154*	*Akr1b10*	*Irf2bp2*	*Fbxo47*
Ctla4	*Atp2a3*	*Sh3bp5*	*Zfp36l2*	*Kirrel*	*Trim24*	*Pard3*	*Trib2*
Ino80d	*Rap1gap2*	*Ercc6*	*Epas1*	*Etv3*	*Ndufb2*	*Sesn3*	*Frmd6*
Map2	*Cpd*	*Wdfy4*	*Gm4832*	*Mef2d*	*Mir704*	*Izumo1r*	*Cdca4*
Kansl1l	*Ssh2*	*Gch1*	*Foxn2*	*Arhgef2*	*Gimap5*	*Smco4*	*BC005537*
Mreg	*Coro6*	*Mrpl57*	*Svil*	*Pi4kb*	*Mir148a*	*S1pr2*	*Lysmd3*
Tns1	*Pipox*	*Pebp4*	*Egr1*	*Hist2h2be*	*Jazf1*	*Bbs9*	*Gpbp1*
Wnt10a	*Slfn2*	*Egr3*	*Nr3c1*	*Tent5c*	*Hpgds*	*Kcnj1*	*Mtrex*
Mrpl44	*Hnf1b*	*Rb1*	*Dtwd2*	*Man1a2*	*Tnip3*	*Ets1*	*Ctnnd2*
Dock10	*Mir21a*	*Lrch1*	*Dmxl1*	*Dennd2d*	*4930515G16Rik*	*4930581F22Rik*	*Mir30d*
Gpr55	*Dhx40*	*Gpr18*	*Zfp608*	*Prpf38b*	*Rprl1*	*Tirap*	*4930483J18Rik*
Ptma	*Ppm1e*	*Selenop*	*Gramd3*	*S1pr1*	*Hmces*	*Vsig2*	*Eef2kmt*
Atg16l1	*Cuedc1*	*Ankrd33b*	*March3*	*A930005H10Rik*	*Plxna1*	*Gramd1b*	*Abhd10*
Ackr3	*Sp6*	*Mtdh*	*Prrc1*	*Cnn3*	*Lrig1*	*Mfrp*	*Nfkbiz*
Tnfrsf11a	*Mllt6*	*Zfp706*	*Ndst1*	*Gclm*	*Suclg2*	*Ccdc84*	*Pcnp*
Bcl2	*Gsdma*	*Odf1*	*Grpel2*	*Ank2*	*Foxp1*	*Cxcr5*	*Klhdc3*
Actr3	*Krt222*	*Azin1*	*Adrb2*	*Lef1*	*Eif4e3*	*Ddx6*	*Tfeb*
Tmem163	*Atp6v0a1*	*Zhx2*	*Malt1*	*Papss1*	*Il5ra*	*BC049352*	*Oard1*
Cxcr4	*Ubtf*	*Tbc1d31*	*Fam210a*	*Bank1*	*Bhlhe40*	*Nxpe2*	*Rab5a*
Il10	*Grn*	*Tatdn1*	*Tcf4*	*Ppp3ca*	*Setd5*	*Nnmt*	*Ndufa11*
Mdm4	*Itgb3*	*Trib1*	*Nfatc1*	*Tspan5*	*Jagn1*	*Pou2af1*	*Pja2*
Tmem9	*Milr1*	*A1bg*	*Pold4*	*Sh3glb1*	*Vgll4*	*Pstpip1*	*Ska1*
Zfp281	*Pitpnc1*	*Myc*	*Pacs1*	*Mcoln3*	*Hdhd5*	*1700017B05Rik*	*Sf1*
Nek7	*Prkca*	*H2afy3*	*Ehd1*	*Tox*	*Gnb3*	*Arid3b*	*Slc22a12*
Dennd1b	*Rgs9*	*Ly6m*	*Zbtb3*	*Plekhf2*	*Ccnd2*	*Sema7a*	*Pfpl*
Zbtb41	*Kcnj16*	*Ly6e*	*Ms4a7*	*Bach2*	*Clec2d*	*Tle3*	*Tmem252*
Odr4	*Kcnj2*	*Cbx7*	*Ms4a6c*	*Smu1*	*Cd69*	*Rplp1*	*Pten*
Cacna1e	*Cd300a*	*Tnrc6b*	*Gm8369*	*Sit1*	*Etv6*	*Uchl4*	*Camsap1*
Tor3a	*Slc25a19*	*Arfgap3*	*Gna14*	*Grhpr*	*2810454H06Rik*	*Plekho2*	*Nup214*
Tex35	*Usp36*	*Arhgap8*	*1110059E24Rik*	*AI427809*	*Pbp2*	*Herc1*	*Nfe2l2*
4930523C07Rik	*Ttc32*	*5031439G07Rik*	*Plgrkt*	*Slc44a1*	*Etnk1*	*Rab8b*	*A330069E16Rik*
Rc3h1	*Trib2*	*Celsr1*	*Pten*	*D630039A03Rik*	*1700073E17Rik*	*Myo1e*	*Atp11b*
Sell	*Pqlc3*	*Gramd4*	*Cpeb3*	*Ugcg*	*Lmntd1*	*Tcf12*	*Tsc22d2*
Rcsd1	*Pik3cg*	*Ano6*	*Myof*	*Susd1*	*Bhlhe41*	*Fam214a*	*Rap2b*
Pou2f1	*Ccdc71l*	*Arid2*	*Blnk*	*Slc31a2*	*Ppfibp1*	*Hcrtr2*	*Txnip*
Slamf6	*Hdac9*	*Scaf11*	*Frat1*	*Mtap*	*Rep15*	*Ibtk*	*Pnrc1*
Zbtb18	*Ahr*	*Mir1941*	*Mirt1*	*Cdkn2a*	*Dennd5b*	*Adamts7*	*Dnaja1*
Parp1	*Brms1l*	*Zfp263*	*Tcf7l2*	*Atg4c*	*Bicd1*	*E030011O05Rik*	*Dnajc11*
1700056E22Rik	*Vcpkmt*	*Snn*	*Gm7102*	*Ssbp3*	*Leng9*	*Stag1*	*Dhrsx*
Cenpf	*Nin*	*Snx29*	*Usp6nl*	*Rab3b*	*Apoc4*	*Tmem108*	*Tmem60*
Atf3	*Fut8*	*Spidr*	*Il2ra*	*Bend5*	*Ceacam19*	*Ackr4*	*Ppm1g*
Traf5	*Gphn*	*Hira*	*St8sia6*	*C530005A16Rik*	*Kcnn4*	*Nek11*	*Sh3bp2*
Cnksr3	*Elmsan1*	*Olfr164*	*Bmi1*	*Mfsd2a*	*Xrcc1*	*Cnot10*	*Wdr19*
NA	*Tshr*	*Dgkg*	*Gfi1b*	*Zc3h12a*	*Pafah1b3*	*Cmtm6*	*Smim14*
Zc3h12d	*Ston2*	*St6gal1*	*6530402F18Rik*	*Laptm5*	*Zfp658*	*Tgfbr2*	*Uso1*
Hivep2	*Foxn3*	*Lpp*	*Ggta1*	*Snhg3*	*Abcc6*	*Dlec1*	*Ppp1cc*
Tnfaip3	*Rps6ka5*	*Uts2b*	*Ptgs1*	*Zdhhc18*	*Lrrk1*	*Csrnp1*	*Mospd3*
Map3k5	*Golga5*	*4632428C04Rik*	*Zbtb6*	*Il22ra1*	*Mef2a*	*1700048O20Rik*	*Tpra1*
1700020N01Rik	*Syne3*	*Hes1*	*Crb2*	*Rpl11*	*Chd2*	*Snrk*	*Ccdc9*
Epb41l2	*B930059L03Rik*	*Fam43a*	*Gm13498*	*Alpl*	*Slco3a1*	*Tcaim*	*Hnrnpul1*
Marcks	*Ppp1r13b*	*Pak2*	*Gpd2*	*B330016D10Rik*	*Sv2b*	*Fyco1*	*Pak4*
Armc2	*Adssl1*	*Rubcn*	*Baz2b*	*Slc25a34*	*Akap13*	*H2al1j*	*Spib*
Pdss2	*Tedc1*	*Heg1*	*Cobll1*	*Vps13d*	*Klhl25*	*Gpr34*	*Rpl18*
F930017D23Rik	*Tmem121*	*Muc13*	*Mettl8*	*Cdk6*	*Zfp710*	*Rp2*	*Rnf169*
Qrsl1	*Rapgef5*	*Slc49a4*	*Ube2e3*	*Gnat3*	*Tlnrd1*	*F9*	*Taok2*
Prdm1	*Pitrm1*	*Fbxo40*	*Prg2*	*Snord93*	*Fchsd2*	*Ldoc1*	*Ell*
Lims1	*Pfkp*	*Cd80*	*Trp53i11*	*Il6*	*Atg16l2*	*F8a*	*Klf2*
Ascc1	*Lgals8*	*Btla*	*Cd44*	*Tmem214*	*Mir139*	*5430427O19Rik*	*Exosc6*
Tbata	*Ero1lb*	*Cd47*	*Pdhx*	*Psapl1*	*Rhog*	*Sh3kbp1*	*Map10*
Arid5b	*Hist1h2ae*	*Cblb*	*Apip*	*Cpeb2*	*Dennd5a*		*Hspa8*
Tmem26	*Slc17a2*	*Alcam*	*Ehf*	*Pgm1*	*Wee1*		*Cd3d*
Dip2a	*Cmah*	*Nfkbiz*	*Rasgrp1*	*Klf3*	*Arntl*		*Stoml1*
Gadd45b	*Gm11346*	*Filip1l*	*Vps18*	*Tlr1*	*Psma1*		*Tlr9*
Gng7	*Cdkal1*	*St3gal6*	*Atp8b4*	*Chrna9*	*Nucb2*		*Dcaf1*

Txnrd1	*Dusp22*	*Gpr15*	*Bcl2l11*	*Atp8a1*	*Xylt1*
Chst11	*Irf4*	*Nrip1*	*Morrbid*	*Ociad1*	*Gga2*
Sycp3	*Pxdc1*	*Synj1*	*Cdc25b*	*Rasl11b*	*Arhgap17*
Elk3	*Eci2*	*Il10rb*	*Smox*	*Tmprss11e*	*Sbk1*
Socs2	*Ssr1*	*Tmem50b*	*Snord17*	*Jchain*	*Btbd16*
Btg1	*Gm10790*	*Atp5o*	*Xrn2*	*Thap6*	*Chst15*
Dcn	*Hivep1*	*Psmg1*	*Cd93*	*Paqr3*	*Ppfia1*
Dusp6	*Nol7*	*Ldhal6b*	*Acss1*	*Rasgef1b*	*Xkr5*
Glipr1	*Cd83*	*Snx9*	*Tm9sf4*	*Plac8*	*Ap3m2*
Kcnmb4	*Atxn1*	*Agpat4*	*Ppp1r16b*	*Aff1*	*Htra4*
Myrfl	*Card19*	*Rgmb*	*Top1*	*Rplp0*	*Zfp703*
Cpm	*Syk*	*Abca3*	*Mmp9*	*Taok3*	*Tti2*
Irak3	*Cplx2*	*Fgd2*	*Prex1*	*Med13l*	*Saraf*
Ndufa4l2	*Lpcat1*	*Pim1*	*Nfatc2*	*Dtx1*	*Dusp4*
Polm	*Ell2*	*Rsph1*	*Tshz2*	*Hvcn1*	*Ppp1r3b*
Ikzf1	*Mef2c*	*Cryaa*	*Rtf2*	*Tctn1*	*Prag1*
Plek	*Serinc5*	*Brd4*	*Rbm38*	*Mlxip*	*Pdlim3*
Spred2	*F2rl1*	*H2-Ob*	*Cdh26*	*Hip1r*	*Irf2*
Actr2	*Tnpo1*	*Clic5*	*Slco4a1*	*Rilpl2*	*Cbr4*
Vps54	*Pik3r1*	*Supt3*	*Helz2*	*Rflna*	*Tmem161a*
Zrsr1	*Elovl7*	*Vegfa*	*Fabp5*	*Slc15a4*	*Gdf15*
Eml6	*Gpbp1*	*Rftn1*	*Pde7a*	*Rabgef1*	*Inpp4b*
Slit3	*Map3k1*	*Plcl2*	*Dnajc5b*	*Lnx2*	*N4bp1*

Supplemental Table 8. Differentially expressed genes in *Pou2af1* $^{KO/KO}$ and *Spi1* $^{KO/KO}$ B cell under LPS stimulation

Genes upregulated in *Pou2af1* $^{KO/KO}$ versus WT B cells			Genes downregulated in *Pou2af1* $^{KO/KO}$ versus WT B cells		Genes upregulated in *Spi1* $^{KO/KO}$ versus WT B cells	Genes downregulated in *Spi1* $^{KO/KO}$ versus WT B cells
Wfdc17	*Mid1*	*Tcirg1*	*Btnl1*	*Snora74a*	*Hbb-bs*	*Hepacam2*
Mir22hg	*Nnt*	*Map3k15*	*Isg15*	*Tet1*	*Pik3r6*	*Cygb*
Slc15a4	*Slc11a1*	*Pla2g7*	*Snora78*	*Stk39*	*Txndc5*	*Bcl2a1b*
Plekha1	*Oit3*	*Dok3*	*Lrrc49*	*Ncapg*	*Amigo2*	*Bcl2a1c*
Inpp5f	*Pde4a*	*Dnajb9*	*Pik3r6*	*Kalrn*	*Eaf2*	*Epha2*
Myzap	*Padi2*	*Trim36*	*Ildr1*	*Btnl2*	*Pygl*	*Flt3*
Synj1	*Pfn2*	*Zfp385a*	*Ccdc122*	*Akr1e1*	*Ada*	*Gbp2b*
Cyp27a1	*Abcb4*	*Itpripl2*	*Cdca2*	*Trim17*	*Adcy6*	*Lilrb4a*
Itga9	*Pik3r3*	*Cacna2d4*	*Ada*	*Cdc42ep4*	*Apoe*	*Havcr1*
Sbk1	*Pip5k1b*	*Sidt1*	*Aire*	*Psrc1*	*Atp1b1*	*Slc12a8*
Pld4	*Pla2g4a*	*Rasgef1b*	*Alpl*	*Ehd3*	*Atp9a*	*Cfp*
Rab31	*Plaur*	*A330023F24Rik*	*Steap4*	*Pglyrp2*	*Prdm1*	*Pltp*
Unc5b	*Pltp*	*Arl4c*	*Bcl2a1c*	*Pmaip1*	*Cacna1e*	*Rab20*
Nod1	*Pou2f3*	*Scimp*	*Bmp1*	*Chst7*	*Ctla4*	*Ryr1*
Slco4a1	*Ppargc1a*	*Mcc*	*Bub1*	*C130026I21Rik*	*Cd28*	*Serpine2*
Mex3b	*Ctsa*	*Fnip2*	*Bub1b*	*Dact3*	*Cdh17*	*Tnni2*
Ston2	*Ppl*	*Slc9a9*	*Ccnb2*	*Suv39h2*	*Ackr3*	*Vill*
Actn1	*Ppt1*	*Slc41a2*	*Ccr6*	*Asb2*	*Cplx2*	*Gsn*
Pygl	*Pros1*	*Agap1*	*Cdc25c*	*Ifitm3*	*Crip1*	*Tspan33*
Nr3c2	*Lgmn*	*Peg13*	*Cdkn2a*	*Rgcc*	*Cst7*	*Kctd14*
Scn4a	*Psen2*	*Lmo7*	*Cish*	*Tmed6*	*Dapk2*	*Atp8b4*
Abcg1	*Ptger4*	*Mirt1*	*Ccr3*	*Cenpm*	*Dnase1l3*	*Parp12*
Acp5	*Ptgs1*	*Mreg*	*Ccr4*	*Zfp541*	*Emid1*	*Oas3*
Hsd17b11	*Ptpn14*	*Mir101a*	*Cryz*	*Ifit1bl1*	*Gad1*	*Cd300lf*
Acvrl1	*Wdr81*	*Mir142*	*Dapk2*	*Pxdc1*	*Galc*	*Slamf1*
Avil	*Ptpre*	*Otulin*	*Epn2*	*Rras2*	*Adgrg1*	*Cacna2d4*
Ahr	*Ptprj*	*Rgs1*	*Fscn1*	*Ndc80*	*Hsd11b1*	*Msrb3*
Ank	*Ptprs*	*Rragd*	*Fcer2a*	*Eps8l1*	*Hspg2*	*Stag3*
Apoe	*Rag1*	*Ldlrad4*	*Fes*	*Ero1lb*	*Jchain*	*Rsad2*
Atp1b1	*Rasgrf2*	*Col5a3*	*Gbp2b*	*Ccdc50*	*Il10*	*Plbd1*
Bcl6	*Rbpms*	*Slc40a1*	*Gnaz*	*Zwilch*	*Il2rb*	*1110032F04Rik*
Blk	*Reln*	*Lpar2*	*Cdca3*	*Ube2c*	*Il6st*	*Carmil1*
Ctla4	*Renbp*	*Hpgds*	*Hmmr*	*1700048O20Rik*	*Itga3*	*Fcmr*
Cd2	*Ret*	*Ptpn3*	*Lipc*	*Clybl*	*Kcnc3*	*Basp1*
Cd38	*Rgs16*	*Sertad1*	*Cxcl10*	*Iqcg*	*Hivep3*	*Nfam1*
Cd5	*Rps6ka2*	*Camkk1*	*Ifit1*	*Kif18b*	*Papln*	*Rab30*
Cd9	*Ryr2*	*Socs5*	*Ifit3*	*Taf15*	*Epcam*	*Nxpe2*
Socs3	*S100a6*	*Srpk3*	*Itm2a*	*Cenpl*	*Zfp704*	*Jdp2*
Cxcr4	*Sat1*	*Clec4n*	*Kif5c*	*Arhgap19*	*Edaradd*	*Kmo*
Plk3	*Scn3a*	*Trpm5*	*Hivep3*	*Gypc*	*Bhlha15*	
Col14a1	*Slpi*	*Slc37a2*	*Rps2*	*Osbpl3*	*Nrp1*	
Csf1	*Kdm5d*	*Smpdl3a*	*Anxa1*	*Car13*	*Pdcd1*	
Csf1r	*Plk2*	*Slc15a2*	*Lrrn3*	*Cenpn*	*Prf1*	
Pcyt1a	*Snn*	*Zfp318*	*Lta*	*Cdkn3*	*Endou*	
Dnase1l3	*Sox4*	*Pdcd1lg2*	*Cd93*	*Zfp429*	*Ptpn14*	
Dnase2a	*Serpinb6b*	*Fxyd6*	*Zfp704*	*Zfp493*	*Ell2*	
Dusp2	*Spic*	*Gm6377*	*Melk*	*Pear1*	*Rab17*	
Dyrk1b	*Sptb*	*Gm14137*	*Mns1*	*Ms4a6c*	*Rprl1*	
S1pr1	*Sdc1*	*Kcnq1ot1*	*Gbp4*	*P2ry13*	*Ryr2*	
Egr1	*Sdc4*	*Dusp10*	*Mybl1*	*Elovl7*	*Saa3*	
Egr2	*Zfp658*	*Itm2c*	*Nedd4*	*Parpbp*	*Ccl6*	
Egr3	*Lrrc25*	*Slc15a3*	*Nes*	*Nxpe2*	*Ccl9*	
Adgre1	*Inpp5a*	*Nrp*	*Nid1*	*Bcl9l*	*Spi1*	
Ephx1	*Cntn2*	*Lima1*	*Notch1*	*Ifitm2*	*Slit3*	
Eps8	*Fam83f*	*Grina*	*Nrp2*	*Slc16a3*	*Serpina3g*	
Igsf8	*Tcn2*	*Camk2n1*	*Palm*	*Sytl3*	*Spn*	
F5	*Pik3ip1*	*Susd3*	*Pgk1*	*Hemgn*	*Pde2a*	
Fcer1g	*Phf1*	*Ahnak*	*Plk1*	*Chst10*	*Creb3l2*	
Fcgrt	*Cd300a*	*Lrrk2*	*Pou2af1*		*Cpeb3*	
Fgr	*Tsc22d1*	*Plekho1*	*Pter*		*Stxbp1*	
Fos	*Tgfb3*	*Dock7*	*Rad51*		*Trib1*	
Fosb	*Tgfbi*	*Cxxc5*	*Rasgrp1*		*Tbxa2r*	
Fosl2	*Tgm2*	*Slc39a8*	*Rmrp*		*Sema6d*	
Gad1	*Serinc5*	*Dusp6*	*Ccl22*		*Ckap4*	
Gcnt1	*Lhfpl2*	*Plgrkt*	*Sema5a*		*Socs2*	
Gem	*Klf10*	*5830432E09Rik*	*Sema7a*		*Hid1*	
Gfi1	*Chdh*	*Plpp3*	*Shcbp1*		*Tsc22d1*	
Ggh	*Fam167a*	*Wls*	*Slc31a2*		*Tgfb3*	

Gnb3
Lilr4b
Lilrb4a
Gsr
Grn
Hck
Hes1
Hexa
Hexb
Hivep2
Hmox1
Nr4a1
Hpse
Hspa2
Id2
Ier2
Il10
Il13ra1
Il18
Il1r2
Il18r1
Il5ra
Il6st
Il7r
Itgax
Itgb7
F11r
Junb
Kifc3
Klf4
Lck
Ldlr
Lef1
Lfng
Lgals1
Lgals3
Lgals4
Lifr
Ltk
Glcci1
Lyz2
Mcoln3
Edaradd
Mxd4
Smad1
Havcr1
Maf
Marco

Tnf
Tnfsf10
Ttn
Tyrobp
Uty
Vcam1
Vcl
Trio
Fam43a
Snx32
Myof
Zfp36
B3gnt7
Uap1l1
Phyhd1
Gsn
Slc17a9
Megf9
Insig1
Zfyve28
Plac8
Myo1h
Hmces
Xylt1
Arhgef10
Psd3
Csgalnact1
Dmxl2
Kctd12
Tenm4
Dusp5
2900026A02Rik
Nfkbid
Shank1
Mctp2
Nlrp10
Sipa1l2
Ssh3
Rhobtb2
Srgap3
Adgre5
Slc26a11
Ddx3y
Inpp4a
Ell3
Vat1
Rab6b
Pfkfb4

B230118H07Rik
Diras2
Mcoln2
Ppfia4
Gpr155
Slc44a2
1500015A07Rik
Ptms
1810046K07Rik
Chn2
Afap1
Slc35d2
Jhy
Otud1
Fmnl2
8-Mar
Wdfy3
Malat1
Nebl
4933439C10Rik
Scpep1
Spsb1
Tbc1d30
Rhoh
Gns
Cdyl2
Mir17hg
Gm11346
Rapgef2
Abcc3
Rflnb
Chst15
Ypel2
Lbh
Gpr146
Fam20c
Tmem108
Ctns
Dpp7
Tas1r3
Gpr137b
Clec2i
Lztfl1
Brwd1
Trim7
Slc41a1
Fcrla
Adgrl2

Slc8a1
Serpina3g
Pif1
Stat4
Aurka
Cpeb3
Stxbp1
Kif15
Rnf26
Lgalsl
BC030867
Depdc1b
Dlgap5
Tpi1
Vpreb3
Wnt10a
Wnt10b
Ticam2
Pdk1
Cenpe
Gbp7
Parm1
Zfp599
Adamts14
Vash1
Serpina3f
Zfp874a
Papss2
Sh3bp5
St8sia6
Krt222
Nlrc4
Nlrx1
Prr11
Slamf1
Trpm2
Olfml2b
Msrb3
Rinl
Cyb5r2
Tmem26
Nostrin
Sfmbt2
Rap1gap2
Rsl1
Nsl1
Zfp456
Tnip3

Serinc5
Rapgef3
Wfs1
Wnt10a
Xbp1
Myof
Zap70
Farp2
Man1c1
Nyx
Adamts14
Cacna1i
Lax1
2900026A02Rik
Gprin3
Rbm47
Srgap3
Nlrc3
Pon3
Sh2d2a
Kif19a
Zfp385a
Pik3cg
Ipcef1
Dennd5b
Myo1d
Sfmbt2
Zfp395
Lmo7
Znrf3
Rragd
Tet1
Nucb2
Usp2
Ptpn3
Rec8
Tacc2
Cacna1h
Asb2
Tmem176a
Fkbp11
Rgcc
Echdc3
Mcoln2
Rere
Trp53inp2
Dab2ip
Sec24d
Mzb1
Derl3
Tbc1d9
Osbpl3
Fndc3b
Mvb12b
Pear1
Atat1
Hspa12a
Vsir
Creld2
Ypel2
Nav2
Bhlhe41
Tmem108
Cd96
B3gnt9

2.2 Enhancer repertoires are reshaped independently of early priming and heterochromatin dynamics during B cell differentiation

(Published, 2015, Nature Communications)

nature COMMUNICATIONS

ARTICLE

Received 18 Nov 2014 | Accepted 11 Aug 2015 | Published 19 Oct 2015

DOI: 10.1038/ncomms9324 OPEN

Enhancer repertoires are reshaped independently of early priming and heterochromatin dynamics during B cell differentiation

Mohamed-Amin Choukrallah[1], Shuang Song[1], Antonius G. Rolink[2], Lukas Burger[1,3] & Patrick Matthias[1,4]

A widely accepted model posits that activation of enhancers during differentiation goes through a priming step prior to lineage commitment. To investigate the chronology of enhancer repertoire establishment during hematopoiesis, we monitored epigenome dynamics during three developmental stages representing hematopoietic stem cells, B-cell progenitors and mature B-cells. We find that only a minority of enhancers primed in stem cells or progenitors become active at later stages. Furthermore, most enhancers active in differentiated cells were not primed in earlier stages. Thus, the enhancer repertoire is reshaped dynamically during B-cell differentiation and enhancer priming in early stages does not appear to be an obligate step for enhancer activation. Furthermore, our data reveal that heterochromatin and Polycomb-mediated silencing have only a minor contribution in shaping enhancer repertoires during cell differentiation. Together, our data revisit the prevalent model about epigenetic reprogramming during hematopoiesis and give insights into the formation of gene regulatory networks.

[1] Friedrich Miescher Institute for Biomedical Research, Maulbeerstrasse 66, 4058 Basel, Switzerland. [2] Department of Biomedicine, University of Basel, Mattenstrasse 28, 4058 Basel, Switzerland. [3] Swiss Institute of Bioinformatics, Maulbeerstrasse 66, 4058 Basel, Switzerland. [4] Faculty of Sciences, University of Basel, 4051 Basel, Switzerland. Correspondence and requests for materials should be addressed to M.-A.C. (email: mohamed-amin.choukrallah@fmi.ch) or to L.B. (email: lukas.burger@fmi.ch) or to P.M. (email: patrick.matthias@fmi.ch).

B cells derive from haematopoietic stem cells (HSCs) through multistep differentiation stages. HSCs have both self-renewal and multipotency capacities. They initially give rise to multipotent progenitors (MPPs) that lose self-renewal capacity but keep the ability to generate early progenitors of the lymphoid, myeloid and erythroid lineages. MPPs differentiate into lymphoid-primed MPPs that further give rise to common lymphoid progenitors (CLPs). The CLP compartment contains two distinct populations, all-lymphoid progenitors (ALPs) and B cell-biased lymphoid progenitors (BLPs)[1]. ALPs retain the full lymphoid potential, while BLPs preferentially generate B cells[1] through multiple stages that are functionally distinct: Pre-pro-B, Pro B, Pre-BI, large and small pre-B II, immature B and finally mature B cells[2,3].

B cell development is controlled by the interplay of a cohort of transcription factors (TFs) and DNA cis-regulatory elements (cis-REs)[4–6]. This interaction is crucial to establish transcriptional programs specific to each differentiation stage. Promoters and enhancers are the two major types of cis-REs in eukaryotes. Enhancers are distal cis-RES that can be located hundreds of kilobases (kb) away of their target genes and play a central role in the activation and fine-tuning of their target promoters[7]. In mammalian cells, enhancer elements have been divided into two major categories, active and primed[8], that can be distinguished functionally and by specific histone modification patterns. Active enhancers are characterized by the concomitant presence of H3K4me1 together with acetylation marks such as H3K27ac[9] and are associated with actively transcribed genes, while primed enhancers are solely marked by H3K4me1, lack acetylation marks and their target genes are weakly or not expressed. A subset of primed enhancers are also additionally marked by the Polycomb group (PcG)-related repressive mark H3K27me3; these enhancers, initially identified in human embryonic stem (ES) cells, have been termed poised enhancers[10]. Primed enhancers are thought to be bookmarked for rapid activation in response to environmental or developmental signals.

Cell differentiation from pluripotent stem cells requires not only the activation of specific sets of genes characteristic of the differentiated cell phenotype but also efficient and temporally controlled silencing of pluripotency and lineage inappropriate genes. The main chromatin-associated repressive mechanisms are the PcG-mediated repression and heterochromatin. PcG targets harbour the H3K27me3 mark, which is catalysed by EZH1 and 2 enzymes, two methyl-transferases belonging to the PRC2 complex[11,12]. Heterochromatin-enriched loci are marked by H3K9me2/3, a reaction catalysed by the H3K9 methyl-transferases G9A and G9a-like protein[13]. It has been reported that ES cells possess less expanded heterochromatin blocks than differentiated cells[14–16]. These observations suggest that the reduced prevalence of heterochromatin in stem cells plays a role in their developmental plasticity. However, this model was challenged by another study showing that the distribution of heterochromatin is largely conserved between ES cells and differentiated neurons[17]. The dynamics of heterochromatin in adult stem cells and their progeny have been less studied. Furthermore, the crosstalk between heterochromatin and the PcG machinery is a matter of debate: although some reports showed that these two mechanisms are mutually exclusive[17], other studies proposed that they can cooperate to exert their silencing function[18].

Although the epigenetic profiles at specific B cell stages are well described[19], transitions between them have been little investigated. So far it is unclear how the features of enhancers change during the transition from multipotent stem cells to committed progenitors and then to differentiated cells such as mature B cells. The prevailing model is that the enhancer landscape is largely established in early haematopoietic progenitors and that multilineage priming of enhancer elements precedes commitment to the lymphoid or myeloid lineages. This model implies that enhancers used in terminally differentiated cells are pre-marked by H3K4me1 (that is, primed) in early stages before their activation during differentiation or in response to stimuli[20–22]. This model was recently challenged by investigations in the myeloid system, which found only limited enhancer priming in early myeloid progenitors[23]. The role of early enhancer priming during B cell differentiation, before and after the lineage commitment, has not been thoroughly investigated. Furthermore, the role of repressive epigenetic mechanisms in reshaping enhancer repertoires is poorly understood.

Here, we use a genome-wide chromatin immunoprecipitation (ChIP)-sequencing approach to investigate the enhancer dynamics in three developmental stages of the murine haematopoietic system. To this end, we compare the profiles of H3K4me3, H3K4me1, H3K27ac, H3K27me3 and H3K9me2 in uncommitted progenitors/HSCs, committed pro-B and splenic mature B cells. In contrast to the prevalent model, we find that early priming has only a minor contribution in building the enhancer repertoire utilized in terminally differentiated cells. Rather, we find the enhancer landscape to be dynamically reshaped at each step of differentiation. The changes involve establishment of novel enhancers as well as closing and reopening of pre-existing ones. Globally, we find that the vast majority of enhancers are *de novo* established in the cell stages where they are required to control their target genes, without prior priming in earlier stages. Furthermore, the analysis of H3K9me2 reveals that the distribution of heterochromatin is largely invariant from HSCs to mature B cells. We also find that the heterochromatin dynamics and PcG-mediated silencing play only minor roles in reshaping enhancer repertoires during B cell differentiation. Overall, our data revisit the current model of epigenome reprogramming during the progression from adult pluripotent stem cells to terminally differentiated cells.

Results

Enhancer classification in HSCs and different B cell stages. Our primary goal was to define the enhancer landscape and its dynamics during B cell development. We focused on three major stages: pluripotent progenitors with HSC potential, committed Pro B cells and mature B cells. As pluripotent progenitors we used immortalized haematopoietic progenitors, as previously described[24,25]. Briefly, bone marrow (BM) cells were transduced with a retrovirus encoding a NUP98–HOXB4 fusion protein and cultured in SCF/IL6 containing media. These cells have stem cell properties as they are able to reconstitute all haematopoietic compartments in lethally irradiated mice and can be serially transplanted[24,25]; long-term HSCs, short-term HSCs (Supplementary Fig. 1A) and B cells (Supplementary Fig. 1B), among other lineages, can derive from these cells. Thus, the NUP98–HOXB4 transduced haematopoietic progenitors are pluripotent and have stem cell properties and therefore will hereafter be referred to as HSCs. Pro B cells were expanded on feeder cells in the presence of IL-7 and splenic mature B cells were induced to proliferate by LPS.

To identify enhancer elements, we generated genome-wide maps for the histone modifications H3K4me1 and H3K4me3. These two marks are known to distinguish enhancers from promoters: active promoters are characterized by high level of H3K4me3 and low level of H3K4me1, while enhancers exhibit high H3K4me1 and low H3K4me3 levels[26]. Thus, we identified putative enhancers as loci not overlapping with annotated promoters, lacking H3K4me3 and showing significant

enrichment in H3K4me1 signal over the corresponding input chromatin (see Material and methods). Based on these criteria, we identified ~40,000 putative enhancers in HSCs and 25,000 putative enhancers in Pro B and mature B cells (Fig. 1a,b).

We then classified putative enhancers based on their enrichment in H3K27ac, which marks active enhancers[9] and H3K27me3, which marks poised enhancers[10]; putative enhancers lacking these two marks and enriched solely in H3K4me1 were considered to be primed enhancers. Based on these criteria, we found that in HSCs roughly 42% of enhancers were active, 56.5% primed and only 1.5% poised. These proportions were moderately different in Pro B (36% active, 61% primed and 3% poised) and mature B cells (32% active, 66.5% primed and 1.5%) (Fig. 1b). We observed a reproducible reduction in the number of enhancers in Pro B and mature B cells relative to HSCs. This reduction was more pronounced at active enhancers, suggesting less enhancer usage in B cells.

Enhancers act synergistically to promote gene expression. To investigate the relationship between enhancer status and gene expression, we performed RNA-sequencing in all three differentiation stages. We assigned each identified enhancer to the closest promoter, allowing a maximal distance of 500 kb between enhancer and target promoter (see Material and methods). In all three developmental stages genes associated with active enhancers show on average the highest expression levels, followed by genes associated with primed enhancers, poised enhancers and genes not associated with any enhancer (Fig. 1c). These observations were functionally validated by testing a number of putative enhancers in a reporter assay. For this, different categories of enhancers (that is, primed in HSCs and active in Pro B, or active in both stages) were tested for their ability to drive the expression of the nano-luciferase gene (Fig. 1d, Supplementary Fig. 7D, see Material and methods). We found that the putative enhancers identified as being primed in HSCs and active in Pro B cells were

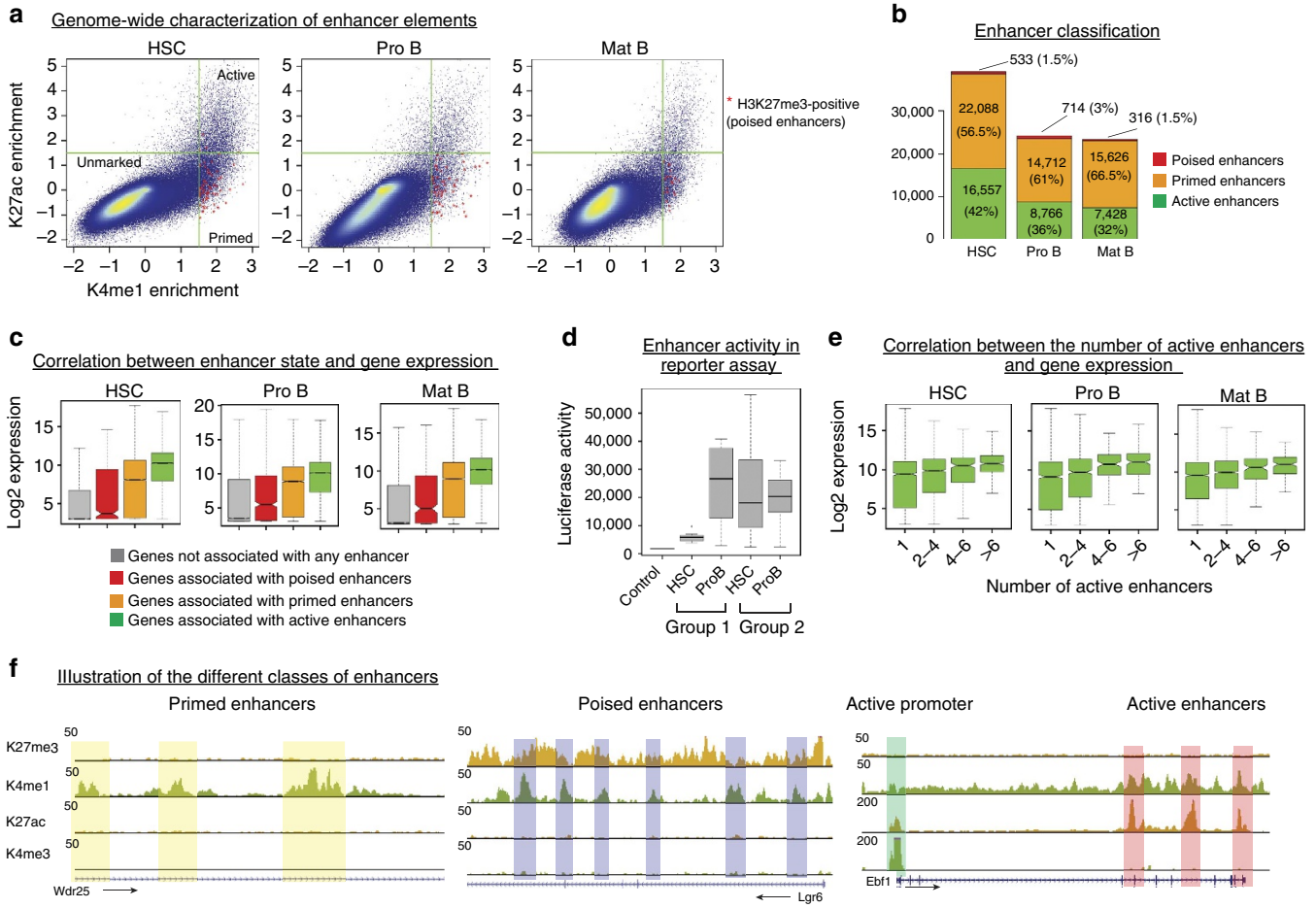


Figure 1 | Genome-wide characterization of enhancer elements in HSCs and different B cell stages. (**a**) H3K4me1 versus H3K27ac enrichments (over input) for 1 kb sliding windows across the genome. Regions overlapping with annotated promoters or enriched in H3K4me3 were excluded. Green lines indicate the cut-offs used to select enriched windows for both signals (Materials and methods). Windows additionally enriched for the H3K27me3 mark are marked by red stars (*). (**b**) Putative enhancers (corresponding to merged windows enriched in H3K4me1) were classified according to their enrichment in H3K27ac and H3K27me3 signal. H3K4me1+/H3K27ac+ loci correspond to active enhancers; H3K4me1+/H3K27ac- loci correspond to primed enhancers and H3K4me1+/H3K27me3+ loci to poised enhancers.(**c**) Box plots showing expression levels of genes according to their association with distinct classes of enhancers. (**d**) Functional validation of enhancer states by a reporter assay. Box plots show the luciferase activity for different categories of enhancers in HSCs and Pro B cells. Group 1 consists of enhancers primed in HSCs and active in Pro B cells (n=8); group 2 consists of enhancers active in both HSCs and Pro B cells (n=11). An enhancer-less construct (see Material and methods) was used as a negative control. Raw data are presented in Supplementary Fig. 7D. (**e**) Gene expression levels as a function of the number of associated active enhancers. (**f**). Genome browser snapshots illustrating primed enhancers (*Wdr25* locus), poised enhancers (*Lgr6* locus), and active enhancers and promoter (*Ebf1* locus) in Pro B cells. HSC, haematopoietic stem cells; Mat B, mature B cells; Pro B, progenitor B cells.

indeed significantly more active in Pro B cells than in HSCs, although a significant activity could be measured in HSCs when compared with the enhancers-less construct. In contrast, the enhancers that we had identified as active in both stages were roughly equally active in both differentiation stages and globally show a higher activity than primed enhancers (Fig. 1d).

Interestingly, expression levels generally increase with the number of active enhancers associated with a given gene (Fig. 1e), suggesting that enhancers act synergistically to define the expression level of their target genes.

Active enhancers specific to HSCs, Pro B or mature B cells regulate genes associated with various biological functions. Consistent with the known function of the cell type analysed, HSC enhancers are largely associated with genes involved in cytokines regulation, apoptosis and haematopoietic cell differentiation, whereas genes regulated by Pro B and mature B cell-specific enhancers are associated with lymphocyte differentiation and activation. Many stage-specific enhancers were also found to be associated with genes involved in diseases of the haematopoietic system, such as leukemia or lymphoma (Supplementary Fig. 2).

Enhancers are shaped by ubiquitous and stage-specific TFs. Although enhancer priming is well described for a number of known enhancers in the haematopoietic system[20], the balance between early priming and *de novo* establishment of new enhancers during differentiation or in response to stimuli has not been rigorously addressed. To examine this, we asked three major questions: (i) how dynamic is the enhancer repertoire from HSCs to fully differentiated cells?; (ii) what is the fraction of primed enhancers in HSCs that become active during differentiation?; and (iii) what is the fraction of active enhancers in differentiated cells that were primed in earlier stages? To answer these questions, we first compared H3K4me1 profiles in HSCs, Pro B and mature B cells. This comparison showed that the enhancer repertoire, as characterized by the H3K4me1 mark, is highly dynamic from HSCs to Pro B cells and from Pro B to mature B cells (Fig. 2a). Although the enhancer repertoires of Pro B and mature B cells are more similar to each other than to the repertoire of HSCs (Supplementary Fig. 3A), dynamic reshaping is observed at all stages of the differentiation process. These changes involve *de novo* establishment of new enhancers that are stage-specific, as well as closing and reopening of pre-existing enhancers (Fig. 2a,b).

The enhancer landscape is largely shaped by the cooperative binding of ubiquitous and cell type-specific TFs[4,27]. To address the role of these two categories in enhancer establishment during B cell development, we analysed the identified enhancer sequences for matches to known TF motifs. Enhancers specific to HSCs were largely enriched for PU.1 and many variants of ETS motifs (Fig. 2c). Enhancers specific to Pro B cells were highly enriched for motifs for the B cell-specific TF EBF1 (refs 28,29) and variants of ETS motifs (Fig. 2c); E2A and PU.1 motifs were also found to be enriched, but with less significant *P* values. Finally, mature B cell-specific enhancers were found to be enriched for binding motifs of Oct2, NFkB and IRF1/2/4, factors known to be important in mature B cells[30–32]. These results clearly support the model that cell type-specific TFs can induce the establishment of new enhancers that were absent in earlier differentiation steps in a stage and context-dependent manner.

Stage-specific enhancers are rarely primed in early stages. We next asked what fraction of primed enhancers (K4me1$^+$/K27ac$^-$) in HSCs become active (K4me1$^+$/K27ac$^+$) in differentiated cells.

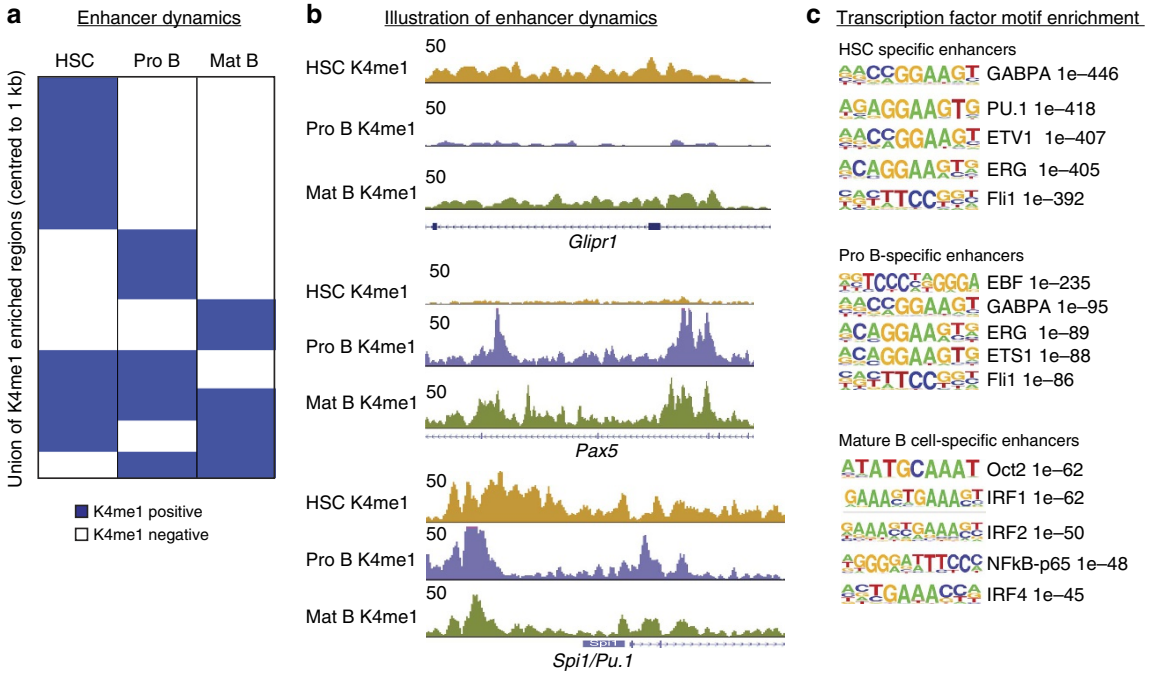


Figure 2 | Enhancer repertoires are dynamically reshaped from HSCs to mature B cells. (**a**) Heatmap showing the dynamic behaviour of the H3K4me1 mark in HSCs, Pro B and mature B cells. H3K4me1-positive regions are shown in blue, H3K4me1-negative regions in white. (**b**) Genome browser snapshot illustrating different kinds of cell type-specific enhancers: marked in HSCs and mature B but not in Pro B cells (top, *Glipr1* locus), marked in Pro B and mature B cells but not in HSCs (middle, *Pax5* locus) or marked in all stages (bottom, *Spi1/Pu1* locus). (**c**) TF motif enrichments for enhancers specific to HSCs, Pro B or mature B cells. Sequence logos and *P* values are shown for the most highly enriched sequence motifs. Only motifs corresponding to TFs that are expressed in the respective cell type are shown (only ETS1 in HSCs and Oct4 in mature B cells were excluded).

To this end, we selected all primed HSC enhancers and calculated their H3K27ac and H3K4me1 enrichments in Pro B or mature B cells, as well as in spleen and thymus using publicly available data[33]. Surprisingly, we found that only a very small fraction of primed enhancers (4%) in HSCs become active in Pro B or mature B cells (Fig. 3a,b). A small fraction of primed enhancers in HSCs remain primed in Pro B and mature B cells, whereas the majority of them (~80%) lose K4me1 and become closed in Pro and mature B cells. Similar observations were made for spleen and thymus.

Furthermore, we found that only a small fraction of primed enhancers in Pro B cells become active in mature B cells, while the majority lose both K4me1 and K27ac (Fig. 3c).

We also investigated the behaviour of poised enhancers in subsequent stages. Similar to primed enhancers, the few poised enhancers identified in HSCs and Pro B cells rarely become active in later stages (Supplementary Fig. 4A,B).

Consistent with this result, we found that only a minority of active enhancers in mature B cells were primed in HSCs or in Pro B cells, while the majority were either already active in earlier stages or completely closed (Fig. 4a). To exclude the possibility that LPS activation may affect enhancer usage in mature B cells, we also generated H3K27ac maps in splenic resting B cells. We found only a modest reduction in the overall number of enriched regions compared with LPS-induced cells (Supplementary Fig. 5A). This is consistent with a previous study showing that in resting B cells the level of histone acetylation as measured by H3K9ac mapping is lower than in other cell types, such as in proliferating *Rag2*−/− Pro B cells[19]. However, in spite of this slight reduction in the number of enriched regions, the

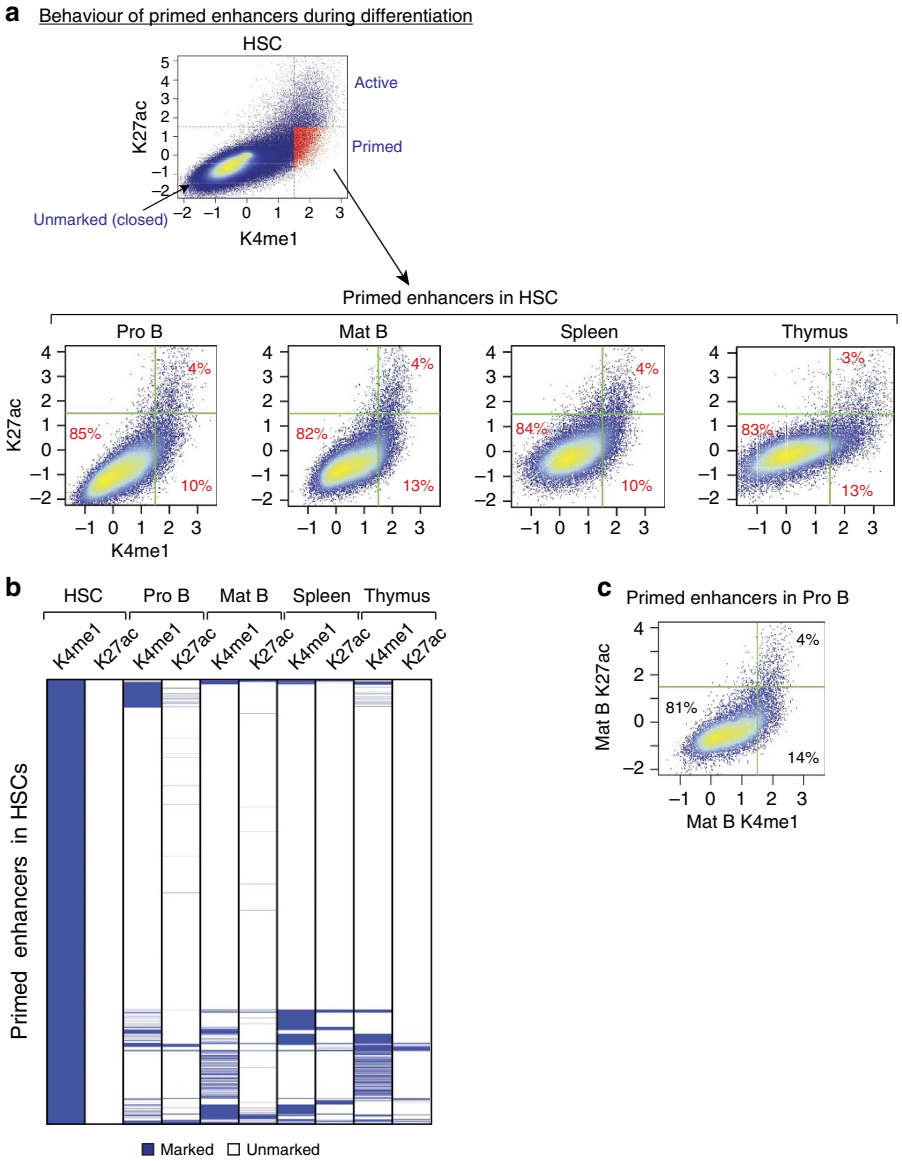


Figure 3 | Enhancers primed in HSCs rarely become active at later differentiation stages. (**a**) Chromatin state of enhancers primed in HSCs was investigated in Pro B, mature B cells, spleen and thymus. H3K4me1 and H3K27ac enrichments in the indicated cell types were calculated at genomic coordinates corresponding to primed enhancers in HSCs. Lines indicate cut-offs used to select enriched regions for each signal. The proportions of the different populations are indicated in the scatter plots. (**b**) Heatmap displaying enhancers primed in HSCs and clustered based on their H3K4me1 and H3K27ac enrichment in the indicated cell types and tissues. (**c**) Similar to A, the chromatin state of enhancers primed in Pro B cells was investigated in mature B cells.

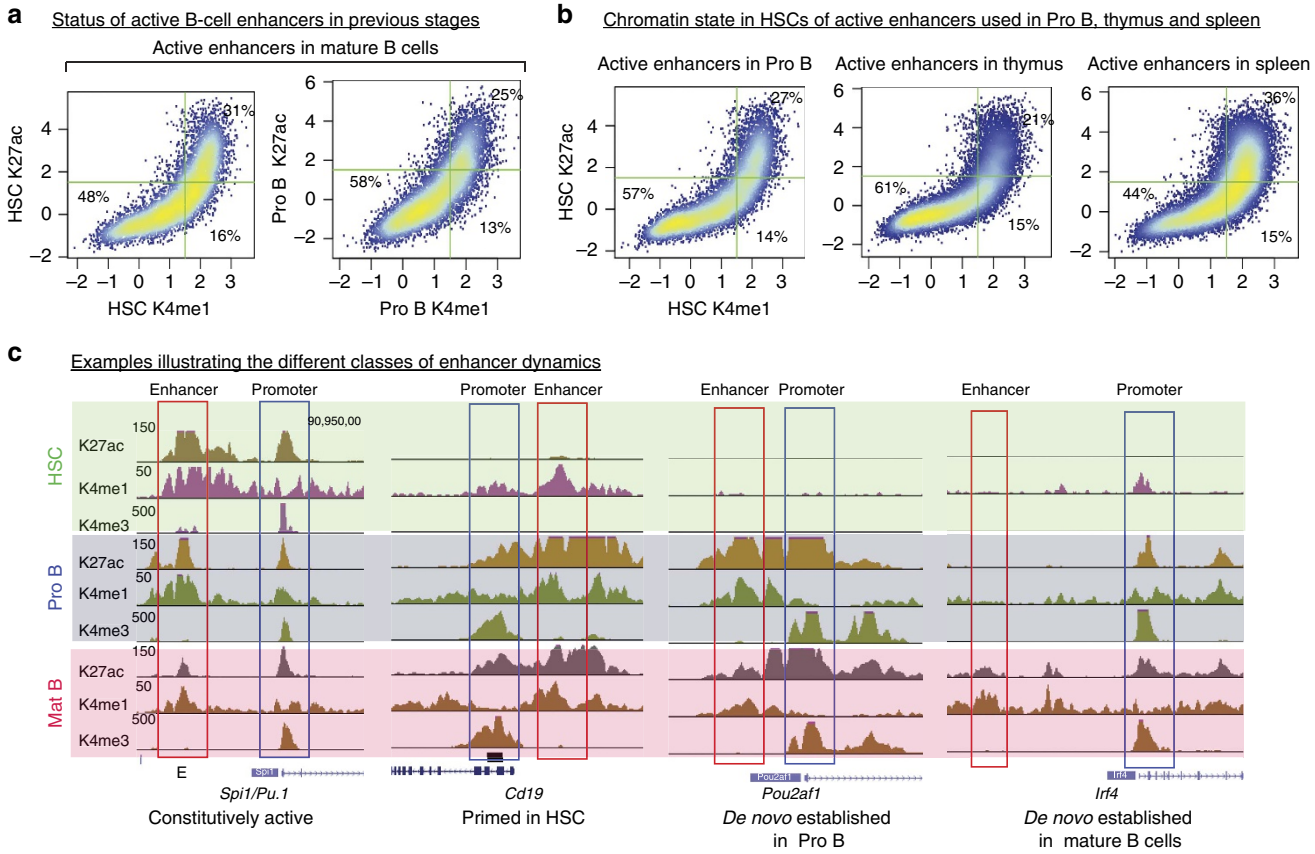


Figure 4 | Most enhancers active in differentiated cells are not primed in stem cells or progenitor stages. (**a**) Chromatin state of active enhancers in mature B cells was investigated in HSCs and Pro B cells. Scatter plots display H3K4me1 and H3K27ac enrichment for HSCs and Pro B cells at the coordinates of active enhancers in mature B cells. The percentage of enhancers in the respective quadrants are indicated. (**b**) Similar to A, the chromatin state of active enhancers in Pro B cells, spleen and thymus was investigated in HSCs. (**c**) Genome browser snapshot illustrating the different classes of enhancer dynamics.

H3K27ac signal is globally similar between LPS-induced and resting B cells (Supplementary Fig. 5A). Furthermore, using the H3K27ac data of resting B cells to identify active enhancers and interrogating their status in Pro B cells and HSCs leads to similar results as for LPS-induced B cells (Supplementary Fig. 5B).

We also considered the possibility that enhancers used in B cell lineages might be primed in an intermediate differentiation stage such as the CLP stage. We therefore used publicly available data sets generated in CLPs[22] and compared how H3K4me1 and H3K27ac change between CLPs, Pro B and mature B cells. The results presented in Supplementary Fig. 5C demonstrate that, in this case also, only a small fraction of the enhancers active in Pro B and mature B cells were primed in CLPs.

Similar observations were also made for active enhancers identified in spleen and thymus when investigated for their status in HSCs (Fig. 4b). Taken together, these results challenge the prevalent model and indicate that early priming events represent an exception rather than the rule: the majority of active enhancers in terminally differentiated cells examined here are *de novo* established in the stage where they are required, with limited priming in previous stages. Fig. 4c presents genome browser views illustrating the different combinations of enhancer dynamics observed from HSCs to mature B cells.

Enhancer usage discriminates close differentiation stages. To address the question whether enhancer usage can distinguish closely related cell stages sharing similar gene expression patterns, we compared gene expression levels for HSCs, Pro B and mature B cells. For simplicity, genes were divided into two classes, expressed or not expressed (see Material and methods and Supplementary Fig. 6) and only genes expressed in at least one stage were considered further. Gene expression patterns are similar between the three stages with only a small fraction found to be differentially expressed (Fig. 5a). Consistent with this, H3K27ac shows only a moderate amount of variability at promoter regions (proximal peaks) between HSCs, Pro B and mature B cells (Fig. 5b, left panel). In contrast, distal H3K27ac peaks (that is, active enhancers) show distinct cell stage-specific patterns (Fig. 5b, right panel), indicating cell stage-specific usage of enhancer elements in cell types sharing highly similar gene expression patterns, in accordance with earlier finding in human cells[34]. This specificity is illustrated by enhancers regulating the *Pax5* gene; while in Pro B cells this gene is controlled by a well-studied enhancer located in the fifth intron (E1)[35], in mature B cells, the *Pax5* gene appears to be associated with two additional enhancers (E2 and E3) located in the sixth intron (Fig. 5c).

Heterochromatin dynamics during differentiation. To investigate whether heterochromatin plays a role in modulating cis-RE dynamics we generated genome-wide maps for H3K9me2, a histone mark characterizing heterochromatin-enriched loci[15]. Visual inspection of H3K9me2 enrichments along the genome revealed a clear separation into large domains of H3K9me2-depleted and H3K9me2-enriched regions (Fig. 6a). To determine the

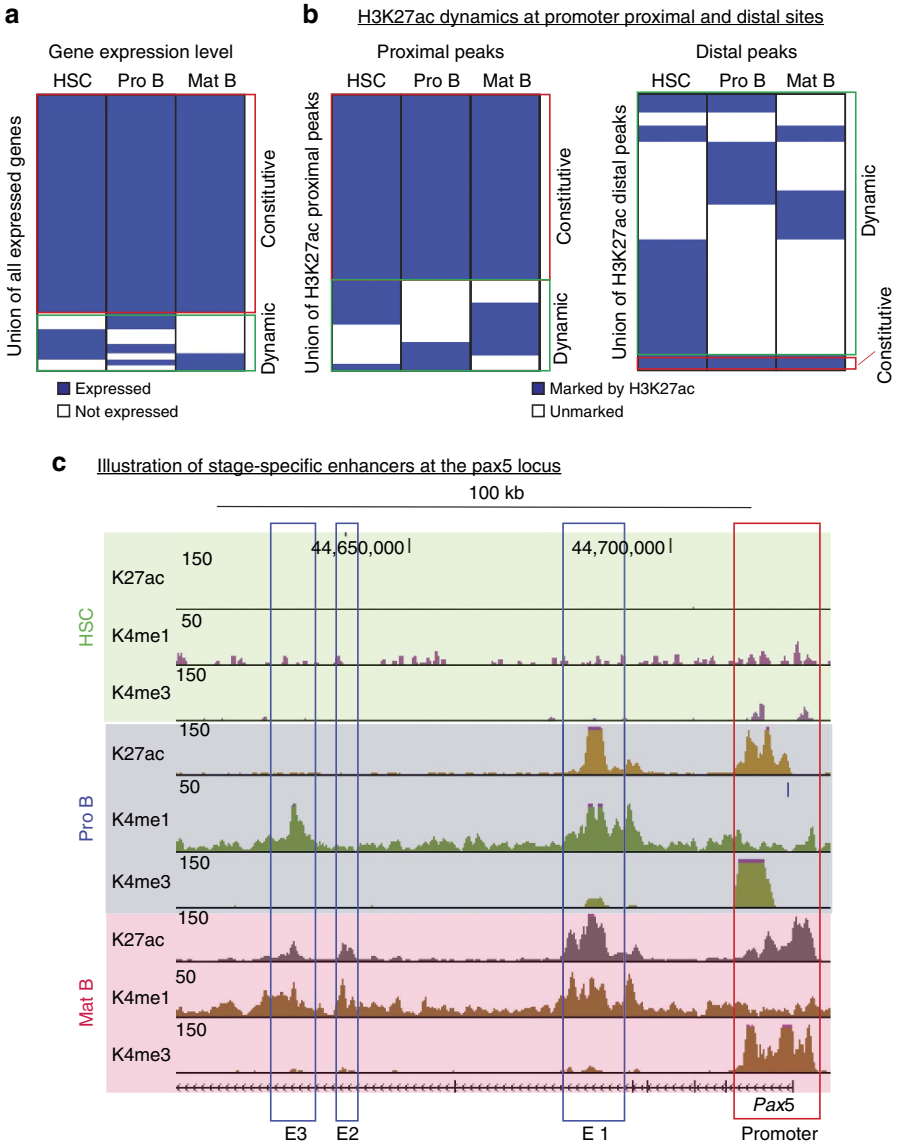


Figure 5 | Enhancers are used in a stage-specific manner. (**a**) Heatmap showing gene expression levels dynamics in HSCs, Pro B and mature B cells. Genes were divided into two categories, expressed (blue) and not expressed (white). (**b**) Heatmap showing H3K27ac signal dynamics at promoters (proximal peaks, left panel) and outside of promoters (distal peaks, right panel). H3K27ac-positive regions are shown in blue and H3K27ac-negative regions in white. (**c**) Browser snapshot illustrating stage-specific enhancer usage for the *Pax5* gene in Pro B and mature B cells. E1, E2 and E3 stand for Enhancers 1, 2 and 3. Enhancer 1 is common to Pro B and mature B cells, enhancer 2 is specific to mature B cells and enhancer 3 is solely marked by H3K4me1 in Pro B cells and acquires H3K27ac in mature B cells.

H3K9me2-enriched domains genome-wide, we used a two-state Hidden Markov Model (HMM, see Material and methods). The identified H3K9me2-enriched heterochromatic domains are shown in red in Fig. 6a for the entire chromosome 19. In all cell types analysed here, the domains cover roughly half of the genome (Fig. 6b). Interestingly, H3K9me2 is largely invariant during differentiation, as evidenced by high correlations at the level of 5 kb tiling windows (Fig. 6c) and promoters (Fig. 6d), as well as by a strong overlap of the HMM-identified domains (Fig. 6e). These findings are in line with previous observations in ES cells and neurons[17]. Despite the global similarity, rare loci show detectable changes in H3K9me2 enrichment (Fig. 6e), for example at the T-cell receptor beta locus, where we observed an overall increase in H3K9me2 signal in Pro B and B cells relative to HSCs (Fig. 6f).

To investigate the relationship between heterochromatin and enhancer dynamics, we monitored the overlap of H3K4me1 and H3K9me2-enriched regions during the differentiation (see Material and methods). As expected, only a small fraction of putative enhancers overlap with H3K9me2-enriched regions in each cell stage (Fig. 6g). Additionally, only a small fraction (~15%) of enhancers that are closed (that is, lose H3K4me1) during the transition from HSCs to Pro B cells become marked by H3K9me2 in Pro B cells (Fig. 6h). Conversely, we found that only ~15% of *de novo* enhancers generated in Pro B cells were marked by H3K9me2 in HSCs (Fig. 6h), indicating that the dynamics of enhancers are largely unrelated to the dynamics of heterochromatin.

Discussion

The development of terminally differentiated cells from pluripotent stem cells is a stepwise process controlled by the complex interaction between TFs and DNA cis-REs embedded in

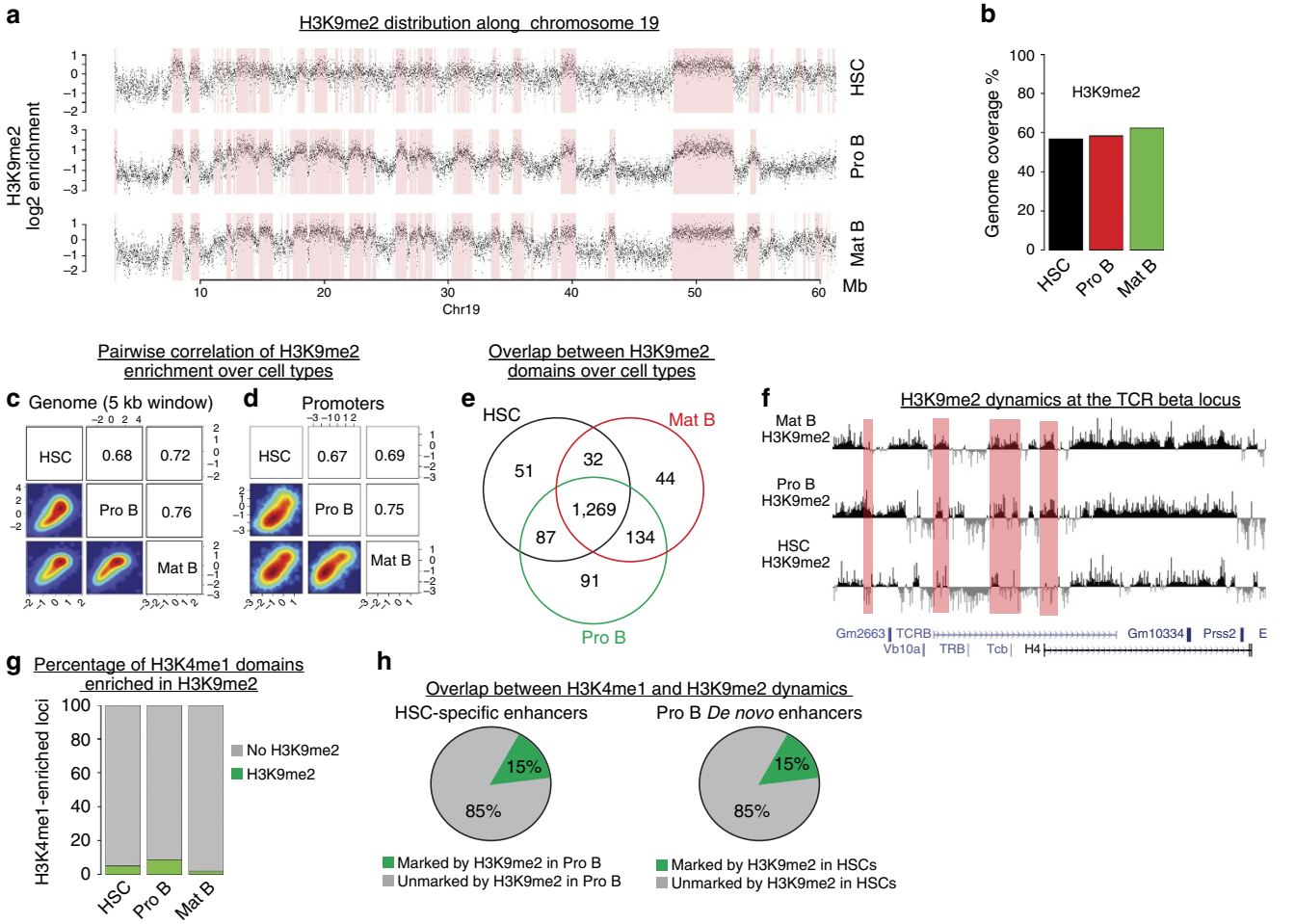


Figure 6 | Heterochromatin distribution is largely stable from HSCs to mature B cells. (**a**) H3K9me2 enrichments in 5-kb windows along the entire chromosome 19 (Chr19) in HSCs, Pro B and mature B cells. Identified heterochromatic domains are indicated by pink rectangles. Mb: megabases. (**b**) Percentage of the genome marked by H3K9me2 in HSCs, Pro B and mature B cells. Only genomic windows with at least one read in any of the input or H3K9me2 samples was considered. (**c**) Pairwise comparison of H3K9me2 enrichment in 5 kb sliding windows for HSCs, Pro B and B cells. Pearson correlations are indicated. (**d**) Similar to C, for promoters (± 1 kb around transcription start sites). (**e**) Venn diagrams showing the number of nucleotides (in Mb) that are common between the identified H3K9me2 domains in HSCs, Pro B and mature B cells. (**f**) Browser snapshot illustrating local changes in H3K9me2 at the T-cell receptor (*TCR*) *beta* locus. (**g**) H3K4me1 marked loci were classified according to their overlap with H3K9me2 domains. (**h**) Enhancers specific to HSCs and closed in Pro B cells were investigated for their overlap with H3K9me2 domains in Pro B cells; similarly, *de novo* enhancers generated in Pro B cells and absent in HSCs were investigated for their overlap with H3K9m2 domains in HSCs. The proportions of H3K9me2-marked and unmarked loci are indicated.

chromatin. Here, we investigated the dynamic changes in the epigenome landscape during B cell development from HSCs. Our study addressed several aspects of epigenetic regulation, ranging from enhancer dynamics to repression mechanisms controlling gene expression. We focused on three main developmental stages: uncommitted MPPs (HSCs), committed progenitors (Pro B) and terminally differentiated cells (mature B cells). We also took advantage of publicly available data generated in other developmental stages or haematopoietic tissues to perform a multilineage comparison.

Several previous studies suggested that enhancers can be activated in a stepwise manner during development[22,35,36]. This activation process involves a primed state that corresponds to a pre-marking of enhancers in early stem cell or progenitor stages before their usage in further differentiated stages. Primed and active states can be discriminated by distinct chromatin signatures: primed enhancers harbour H3K4me1 only, whereas active enhancers harbour acetylation marks such H3K27ac and H3K9ac in addition to H3K4me1. In addition to these two major enhancer states, it was also reported that some enhancers can harbour H3K4me1 and the repressive mark H3K27me3; these enhancers were termed poised enhancers and described primarily in human ES cells[10].

Two very recent studies investigating enhancer dynamics in myeloid cells came to divergent conclusions. Lara-Astiaso *et al.*[22] found that enhancers used in terminally differentiated myeloid cells are not primed in HSCs, but rather in early myeloid progenitors prior to the execution of the RNA expression program. In contrast, Luyten *et al.*[23] concluded that meyloid enhancers are generated *de novo* with only limited priming in earlier stages. These divergent conclusions might be due to the usage of different criteria to select enriched regions for histone marks or to the use of slightly different cell populations.

Although enhancer priming in early stages before their activation was demonstrated for several well-studied enhancers in the haematopoietic system[35–37], the balance in enhancer usage between pre-existing enhancers and *de novo* enhancers had not been addressed genome-wide during B cell differentiation, before

and after commitment. Our data indicate that the majority of active enhancers in differentiated cells are not primed in earlier stages. Indeed, we found that most active enhancers in mature B cells are either already active in the previous stages analysed (Pro B and HSCs) or completely closed (unmarked) in these stages. Furthermore, a differential analysis of H3K4me1 and H3K27ac changes between B cells populations and CLPs stage (public data)[22] indicates that enhancer priming of B cell enhancers is also rare at the CLP stage (Supplementary Fig. 5C). This finding is consistent with the fact that the majority of primed enhancers in HSCs and Pro B cells were not found to become active (that is, positive for H3K27ac mark) in mature B cells. These observations indicate that early priming in previous stages has only a minor contribution to enhancer repertoire establishment during B cell development. Repeating our analysis in the spleen and thymus leads to similar results. This suggests that the minor contribution of priming is a general phenomenon in the haematopoietic lineage and that during cell differentiation from stem cells and progenitors enhancers are dynamically *de novo* generated, with little use of the pre-existing primed enhancer landscape.

Consistent with previous studies, our data also highlight the role of stage-specific TFs in building stage-specific enhancers during B cell development[4,5]. Interestingly, closely related cell stages sharing a highly similar gene expression pattern such as Pro B and mature B cells, have nevertheless highly divergent enhancer repertoires. This suggests cell stage-specific enhancer usage to regulate the same set of genes. Furthermore, our analysis indicates that enhancer patterns may be better descriptors of cellular differentiation stages than commonly used gene expression profiles, and may allow for higher granularity.

The abundance of primed enhancers is thought to reflect cell plasticity of stem cells and early progenitors. Our results challenge this dogma as we found similar proportions of primed enhancers in HSCs, progenitors (Pro B) and mature B cells. Although we have deciphered the usage of primed enhancers in HSCs and Pro B cells during B cell development, it is unclear what the functional role of the H3K4me1 marking in the absence of acetylation (indicative of enhancer activity) is. One possibility is that H3K4me1 deposition in some condition can correspond to 'sterile' binding of TFs, which is not sufficient to induce enhancer activity because of the absence of other binding partners. It is also possible that H3K4me1 represents an epigenetic memory of previous activation events at earlier developmental stages. Nevertheless, we observed that genes associated with primed enhancers show overall a slightly higher expression level than genes not associated with any enhancer, irrespective of the developmental stage considered. Moreover, when these primed enhancers were tested functionally in cell transfection assays, a low but significant activity was observed. This suggests that primed enhancers may already have some weak transcription activation potential. Although we have observed a minor contribution of enhancer priming in earlier stages in building enhancer repertoires in subsequent stages, we cannot exclude the possibility that enhancer priming can occur in a short time window before enhancer activation.

Cell differentiation progresses by repressing pluripotency genes as well as lineage inappropriate genes. Gene silencing is linked to distinct epigenetic mechanisms including DNA methylation, heterochromatin and PcG-mediated repression. In the present study, we focused on the main chromatin-related repression machineries that are heterochromatin and PcG. The role of hetrochromatin prevalence in stem cell plasticity is a matter of debate; although some studies suggest that heterochromatin is more abundant in lineage-restricted cells than in ES cells[14], other investigations showed that heterochromatin prevalence does not discriminate ES cells from neurons[17]. Heterochromatin prevalence and dynamics in adult stem cells, such as HSCs, and their progeny has been poorly investigated so far. We have used H3K9me2 mapping to interrogate the heterochromatin distribution in HSCs, Pro B and mature B cells. Our data indicate that the distribution of heterochromatin is largely invariant between HSCs and lineage-restricted Pro B and mature B cells. Furthermore, we found that heterochromatin dynamics play only a minor role in shaping the enhancer repertoires during B cell differentiation from HSCs.

Overall our study revisits the prevalent models and provides novel insights into the dynamic regulation of the epigenome landscape during haematopoiesis.

Methods

Cells. Nup-Hoxb4 HSCs were generated as previously described in (ref. 24) (see Results section). Pro B cells were sorted as described previously[38] from the BM of 4–5 weeks old mice with the following markers (ckit+, B220+, CD19−, CD25−, IgM−). Pro B cells were expanded in the presence of IL7 and OP9 feeder cells for 1 week. Mature B cells were sorted from the spleen of 8–10-weeks-old mice as CD19-positive cells using MACS kit and CD19 beads (#130-052-201). Mature B cells were cultured in the presence of LPS for 2 days. C57BL/6 mice were housed in a controlled environment and experiments were conducted in accordance with the ordinance provided by Cantonal Veterinary Office, Basel-Stadt, Switzerland.

Antibodies. ChIP: H3K4me3 Millipore #07-473, H3K4me1 abcam #ab8895, H3K27me3 upstate # 07-449, H3K27ac Active Motif #39133, H3K9me2 abcam #1220.
FACS: Anti-Sca1-PerCP Biolegend #108122, Anti-c-Kit-APC Pharmingen #553842, Anti-Flt3-PE Ebioscience #17-1171-83, Anti-Lin-Efluor450 Ebioscience #88-7772-72, B220-FITC BD #553087, CD19 ebioscience #Clone: MB19-1.

Chromatin immunoprecipitation (ChIP). Cells were fixed with 1% formaldehyde in PBS (20 ml volume/ $\sim 1 \times 10^8$ cells) for 10 min, followed by quenching with 2.5 M glycine for 5 min. Cells were lysed for 10 min in 1 ml lysis buffer A (10 mM HEPES, pH 8.0, 10 mM EDTA, 0.5 mM EGTA, 0.25% Triton X-100 + PI). The supernatant was discarded and the nuclei were incubated for 10 min in buffer B (10 mM HEPES, pH 8.0, 200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 0.01% Triton X-100 + PI). Next, the nuclei were suspended in chromatin lysis buffer (50 mM Tris-HCl, pH 8.0, 10 mM EDTA, 0.5% SDS) for 30 min. Sonication was performed using Bioruptor Next Gen (Diagenode) at high output intensity for 25 cycles (30 s on/30 s off). Finally the collected supernatant was diluted 5× with chromatin dilution buffer (250 mM NaCl, 1.67% Triton X-100). The chromatin (25–50 µg) was subjected to the immunoprecipitation with 5 µg of the above mentioned antibodies and incubated overnight. The samples were incubated with either protein A or G coupled to magnetic beads for 1 h. The beads were washed with the following buffers, twice with low salt (0.10% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris pH 8.0, 15 mM NaCl), high salt (0.10% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris pH 8.0, 50 mM NaCl), LiCl (250 mM LiCl, 1 M EDTA, 10 mM Tris pH 8.0, 1% NP-40, 1% Na Deoxychholate) and once with TE (20 mM Tris-HCl, pH8.9, 2 mM EDTA). The protein-DNA complexes were eluted from the beads with 500 µl elution buffer (10 mM Tris-HCl, pH7.5, 1 mM EDTA, 1% SDS, 100 mM NaHCO3) and reverse crosslinked overnight at 65 °C with Proteinase K. ChIPed DNA was isolated by phenol/ chloroform extraction and ethanol precipitated. 10 ng of ChIPed DNA were processed for Illumina high-seq analyser according to the manufacturer's protocol.

RNA isolation and sequencing. Cells sorted from three different mice were pooled and cultured as described above. Total RNA was extracted using RNAeasy kit from QIAGEN #74104 with on-column DNaseI treatment. 10 ng of RNA were used for library preparation using ScriptSeq v2 RNA-Seq Library Preparation Kit and sequenced according to the Illumina protocol.

Library preparation for ChIP sequencing. Sequencing libraries were prepared using bar-coded adapters following the standard Illumina library preparation protocol. Four samples carrying different barcodes were pooled at equal molar ratios and subjected to sequencing on Illumina HiSeq 2000 sequencer according to Illumina standards.

Reporter assays. Loci corresponding to different categories of enhancers were amplified from mouse genomic DNA based on H3K4me1 and H3K27ac peak locations and cloned into pNL1.1 vector (Promega) upstream of the nano-luciferase gene and minimal promoter. 0.4 µg of cloned constructs were electroporated into 2.5×10^5 cells using SF solution (PBC2-00675) and Amaxa 4D-NucleofectorX device. Luciferase activity was assessed 24 h post electroporation using the Nano-Glo Luciferase Assay System #N1110 from Promega and Berthold LB device.

A construct containing only the minimal promoter (enhancer-less construct) was used as a negative control. Genomic Coordinates of the cloned sites are provided in Supplementary Table 1

Computational analyses. Sequencing reads were mapped to the mouse genome (mm9) using the qAlign function, which internally uses Bowtie (http://bowtie-bio.sourceforge.net/), from the QuasR R package[39] (http://www.bioconductor.org/packages/release/bioc/html/QuasR.html) with default settings. Alignments were shifted by 60 bases, corresponding to an estimated fragment length of 120 bp. For window-based analyses, a sliding window of 1 kb length and a step-size of 0.5 kb was used. For all samples, reads were counted in each window using the qCount function in QuasR (with shift = 60) to generate a matrix with read counts for each sample in the defined genomic intervals. The total read counts per sample were normalized to the mean total number of reads across all samples. We calculate the enrichment over input using the following formula (after library-size normalization): (number_of_reads_ChIP + 8)/(number_of_reads_input + 8). We thus add a pseudo-count of 8 to the tag counts within a given region, which results in an effective reduction of enrichments when the counts are low.

We generated four biological replicates for H3K4me1 in HSCs and Pro B cells and two replicates for this mark in mature B cells (Supplementary Figs 3A and 8). For H3K27ac, we generated two biological replicates per stage (Supplementary Fig. 3B). For H3K27me3, we generated two biological replicates in Pro B and mature B cells and used publicly available data for HSCs generated in the same cellular system used in the current study[24]. For H3K4me3, we generated one replicate per stage and for H3K9me2 two replicates for HSCs and one replicate for Pro B and B cells. Enrichments for sample replicates were averaged. The cut-offs to select regions enriched for H3K27ac and H3K27me3 were set to 1.5 (log2 scale), which separates the two modes of the bimodal distribution of these signals at promoters and thus is a sensible choice for genome-wide classification. For H3K4me3, regions with an enrichment value below 1 (log2 scale) were considered as depleted from this mark and region with enrichment value higher than 3 were considered as positive. The cut-offs for the H3K4me3 mark were estimated from its distribution at promoters where negative and positive population can easily be discriminated. For H3K4me1 several cut-offs were tested, ranging from a log2 enrichment of 1 to a log2 enrichment of 2. Upon visual inspection, a cut-off of 1 (2-fold) was found to be too low for confident detection of enriched regions. We therefore used a more conservative cut-off of 1.5 (~3 fold). For all marks, overlapping enriched regions were merged to one region.

Peak detection of H3K4me3 and H3K27ac using the MACS software[40] gave quantitatively similar results to the above described windows-based method.

To determine K9me2-enriched domains, the entire genome was tiled into consecutive windows of 5 kb and for each cell type H3K9me2 enrichments over input were calculated (analogously to the other marks as described above). Only windows with reads in either K9me2 or input samples in any of the three stages were retained (~95% of all windows). Visual inspection of the enrichments along the genome indicated a clear separation into two classes of broad domains with either positive (K9me2-enriched) or negative (K9me2-depleted) values. To determine these domains genome-wide, we first trained, separately for each cell type, a two-state HMM with Gaussian emission distributions on the enrichment values of chromosome 1 using standard expectation maximization as implemented in the R package mhsmm[41]. As starting values, we used an initial probability of 1 to be in the enriched state, symmetric transition probabilities with 0.9 probability to stay in the same state and 0.1 probability to transit to the opposite state, and Gaussian emission distributions with means −1 and 1 and variances 0.5 and 0.5 for the H3K9me2-depleted and H3K9me2-enriched states, respectively. We then predicted the most-likely domains genome-wide using the Viterbi Algorithm as implemented in the predict function of the mhsmm package.

Promoters were defined as non-overlapping −1 to +1 kb intervals around transcription start sites as defined in the refGene.txt file downloaded from the UCSC genome browser (http://hgdownload.cse.ucsc.edu/goldenPath/mm9/database/refGene.txt.gz, downloaded April 20, 2012). ChIP enrichments at promoters were calculated as described for the windows-based analysis.

TF motif discovery was performed using the HOMER software[4] using the following command (findMotifsGenome.pl input_list mm9 output_directory -size 1000). The size of the regions used to search for motifs was set to 1 kb. The lists of TF motifs identified were filtered by expression levels.

ChIP sequencing signals were visualized using library-size normalized wig files generated with the qExportWig function (default parameters) from the QuasR package and uploaded as custom tracks to the UCSC Genome browser. Except for H3K9me2, where they represent enrichment over input in 1 kb windows, wig files represent read counts in 100 nt tiling windows of the genome.

Gene expression levels were determined by summing, for each gene, all RNA-seq reads that map to any of the annotated exons of each gene (using the qCount function of QuasR). Annotations were taken from the UCSC refGene.txt file (see above). To compare expression levels between stages, the RNA-seq libraries were normalized to the mean total number of reads and to the average exonic length. The exonic length of each gene was defined as the total number of unique exonic bases stemming from any annotated transcript.

Enhancers were assigned to the closest promoter allowing for a maximum distance of 500 kb. First, active enhancers were assigned to their closest promoters; the promoters associated with active enhancers were excluded from the list of promoters used to assign primed and poised enhancers.

References

1. Inlay, M. A. *et al.* Ly6d marks the earliest stage of B-cell specification and identifies the branchpoint between B-cell and T-cell development. *Genes Dev.* **23,** 2376–2381 (2009).
2. Welner, R. S., Pelayo, R. & Kincade, P. W. Evolving views on the genealogy of B cells. *Nat. Rev. Immunol.* **8,** 95–106 (2008).
3. Matthias, P. & Rolink, A. G. Transcriptional networks in developing and mature B cells. *Nat. Rev. Immunol.* **5,** 497–508 (2005).
4. Heinz, S. *et al.* Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38,** 576–589 (2010).
5. Lin, Y. C. *et al.* A global network of transcription factors, involving E2A, EBF1 and Foxo1, that orchestrates B cell fate. *Nat. Immunol.* **11,** 635–643 (2010).
6. Choukrallah, M. A. & Matthias, P. The interplay between chromatin and transcription factor networks during B cell development: who pulls the trigger first? *Front. Immunol.* **5,** 156 (2014).
7. Bulger, M. & Groudine, M. Functional and mechanistic diversity of distal transcription enhancers. *Cell* **144,** 327–339 (2011).
8. Calo, E. & Wysocka, J. Modification of enhancer chromatin: what, how, and why? *Mol. Cell* **49,** 825–837 (2013).
9. Creyghton, M. P. *et al.* Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc. Natl Acad. Sci. USA* **107,** 21931–21936 (2010).
10. Rada-Iglesias, A. *et al.* A unique chromatin signature uncovers early developmental enhancers in humans. *Nature* **470,** 279–283 (2011).
11. Cao, R. *et al.* Role of histone H3 lysine 27 methylation in polycomb-group silencing. *Science* **298,** 1039–1043 (2002).
12. Shen, X. *et al.* EZH1 mediates methylation on histone H3 lysine 27 and complements EZH2 in maintaining stem cell identity and executing pluripotency. *Mol. Cell* **32,** 491–502 (2008).
13. Tachibana, M. *et al.* Histone methyltransferases G9a and GLP form heteromeric complexes and are both crucial for methylation of euchromatin at H3-K9. *Genes Dev.* **19,** 815–826 (2005).
14. Hawkins, R. D. *et al.* Distinct epigenomic landscapes of pluripotent and lineage-committed human cells. *Cell Stem Cell* **6,** 479–491 (2010).
15. Wen, B., Wu, H., Shinkai, Y., Irizarry, R. A. & Feinberg, A. P. Large histone H3 lysine 9 dimethylated chromatin blocks distinguish differentiated from embryonic stem cells. *Nat. Genet.* **41,** 246–250 (2009).
16. Meshorer, E. *et al.* Hyperdynamic plasticity of chromatin proteins in pluripotent embryonic stem cells. *Dev. Cell* **10,** 105–116 (2006).
17. Lienert, F. *et al.* Genomic prevalence of heterochromatic H3K9me2 and transcription do not discriminate pluripotent from terminally differentiated cells. *PLoS Genet.* **7,** e1002090 (2011).
18. Mozzetta, C. *et al.* The histone H3 lysine 9 methyltransferases G9a and GLP regulate polycomb repressive complex 2-mediated gene silencing. *Mol. Cell* **53,** 277–289 (2014).
19. Revilla, I. D. R. *et al.* The B-cell identity factor Pax5 regulates distinct transcriptional programmes in early and late B lymphopoiesis. *EMBO J.* **31,** 3130–3146 (2012).
20. Mercer, E. M. *et al.* Multilineage priming of enhancer repertoires precedes commitment to the B and myeloid cell lineages in hematopoietic progenitors. *Immunity* **35,** 413–425 (2011).
21. Cui, K. *et al.* Chromatin signatures in multipotent human hematopoietic stem cells indicate the fate of bivalent genes during differentiation. *Cell Stem Cell* **4,** 80–93 (2009).
22. Lara-Astiaso, D. *et al.* Immunogenetics. Chromatin state dynamics during blood formation. *Science* **345,** 943–949 (2014).
23. Luyten, A., Zang, C., Liu, X. S. & Shivdasani, R. A. Active enhancers are delineated de novo during hematopoiesis, with limited lineage fidelity among specified primary blood cells. *Genes Dev.* **28,** 1827–1839 (2014).
24. Vigano, M. A. *et al.* An epigenetic profile of early T-cell development from multipotent progenitors to committed T-cell descendants. *Eur. J. Immunol.* **44,** 1181–1193 (2013).
25. Ruedl, C., Khameneh, H. J. & Karjalainen, K. Manipulation of immune system via immortal bone marrow stem cells. *Int. Immunol.* **20,** 1211–1218 (2008).
26. Heintzman, N. D. *et al.* Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat. Genet.* **39,** 311–318 (2007).
27. Shen, Y. *et al.* A map of the cis-regulatory sequences in the mouse genome. *Nature* **488,** 116–120 (2012).
28. Hagman, J., Gutch, M. J., Lin, H. & Grosschedl, R. EBF contains a novel zinc coordination motif and multiple dimerization and transcriptional activation domains. *EMBO J.* **14,** 2907–2916 (1995).
29. Lin, H. & Grosschedl, R. Failure of B-cell differentiation in mice lacking the transcription factor EBF. *Nature* **376,** 263–267 (1995).

30. Emslie, D. *et al.* Oct2 enhances antibody-secreting cell differentiation through regulation of IL-5 receptor alpha chain expression on activated B cells. *J. Exp. Med.* **205,** 409–421 (2008).
31. Matsuyama, T. *et al.* Targeted disruption of IRF-1 or IRF-2 results in abnormal type I IFN gene induction and aberrant lymphocyte development. *Cell* **75,** 83–97 (1993).
32. Gerondakis, S. & Siebenlist, U. Roles of the NF-kappaB pathway in lymphocyte development and function. *Cold Spring Harb Perspect. Biol.* **2,** a000182 (2010).
33. Yue, F. *et al.* A comparative encyclopedia of DNA elements in the mouse genome. *Nature* **515,** 355–364 (2014).
34. Heintzman, N. D. *et al.* Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature* **459,** 108–112 (2009).
35. Decker, T. *et al.* Stepwise activation of enhancer and promoter regions of the B cell commitment gene Pax5 in early lymphopoiesis. *Immunity* **30,** 508–520 (2009).
36. Walter, K., Bonifer, C. & Tagoh, H. Stem cell-specific epigenetic priming and B cell-specific transcriptional activation at the mouse Cd19 locus. *Blood* **112,** 1673–1682 (2008).
37. Maier, H. *et al.* Early B cell factor cooperates with Runx1 and mediates epigenetic changes associated with mb-1 transcription. *Nat. Immunol.* **5,** 1069–1077 (2004).
38. Yamaguchi, T. *et al.* Histone deacetylases 1 and 2 act in concert to promote the G1-to-S progression. *Genes Dev.* **24,** 455–469 (2010).
39. Gaidatzis, D., Lerch, A., Hahne, F. & Stadler, M. B. QuasR: quantification and annotation of short reads in R. *Bioinformatics* **31,** 1130–1132 (2015).
40. Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9,** R137 (2008).
41. O'connell, J. & Hojsgaard, S. Hidden Semi Markov Models for Multiple Observation Sequences: The mhsmm Package for R. *J. Stat. Softw.* **39,** 1–22 (2011).

Acknowledgements

We thank Sophie Dessus-Babus and Tim Roloff for library preparation, Hubertus Kohler for cell sorting, Alessandra Vigano for sharing data, Benjamin Herquel, Tugce Aktas Ilik and Michael Stadler for critical reading of the manuscript. This work was supported by the Novartis Research Foundation and Systems X RTD grant Cell Plasticity from the Swiss National Science Foundation (SNF) to P.M.

Author contributions

M.A.C. and P.M. designed the project; M.A.C. performed all the experiments and the majority of bioinformatic analyses; S.S. contributed to the reporter assay experiments, L.B. performed and supervised the bioinformatics analysis; A.G.R. provided help with cultures of HSCs; P.M. supervised the project. M.A.C. wrote the paper with all authors contributing to the final version.

Additional information

Accession codes: All data sets have been deposited in GEO under the accession number GSE60005.

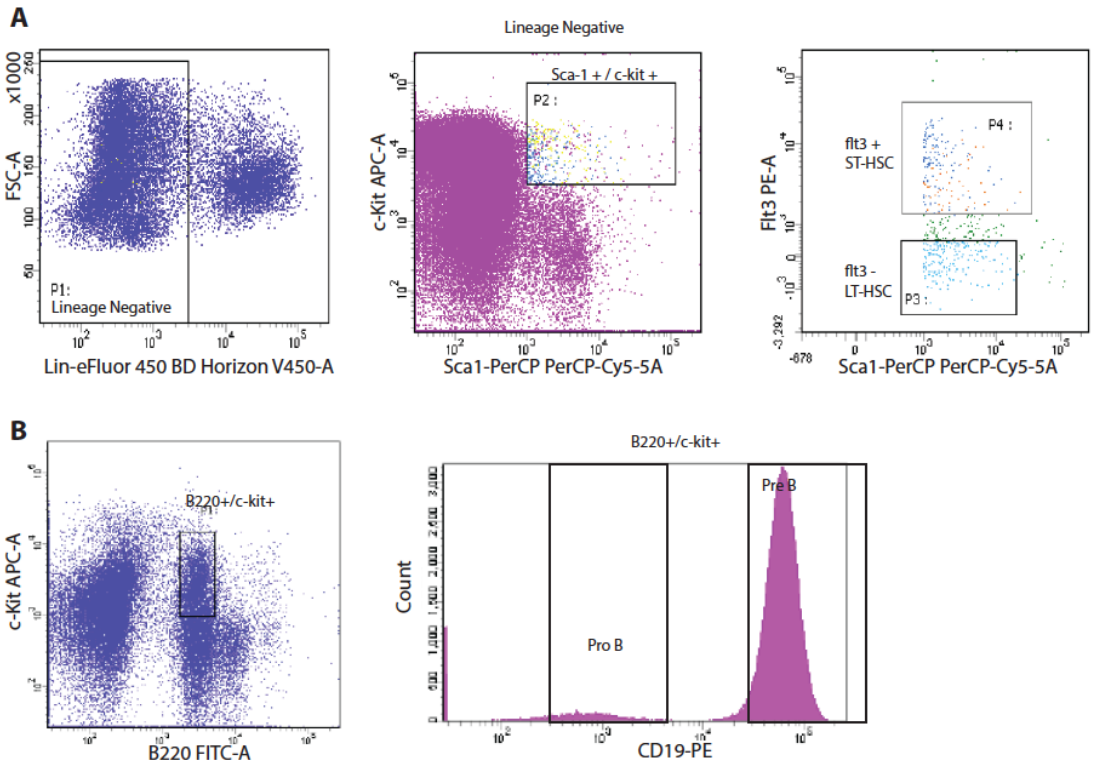
Supplementary Information accompanies this paper at http://www.nature.com/naturecommunications

Competing financial interests: The authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

How to cite this article: Choukrallah, M.-A. *et al.* Enhancer repertoires are reshaped independently of early priming and heterochromatin dynamics during B cell differentiation. *Nat. Commun.* 6:8324 doi: 10.1038/ncomms9324 (2015).

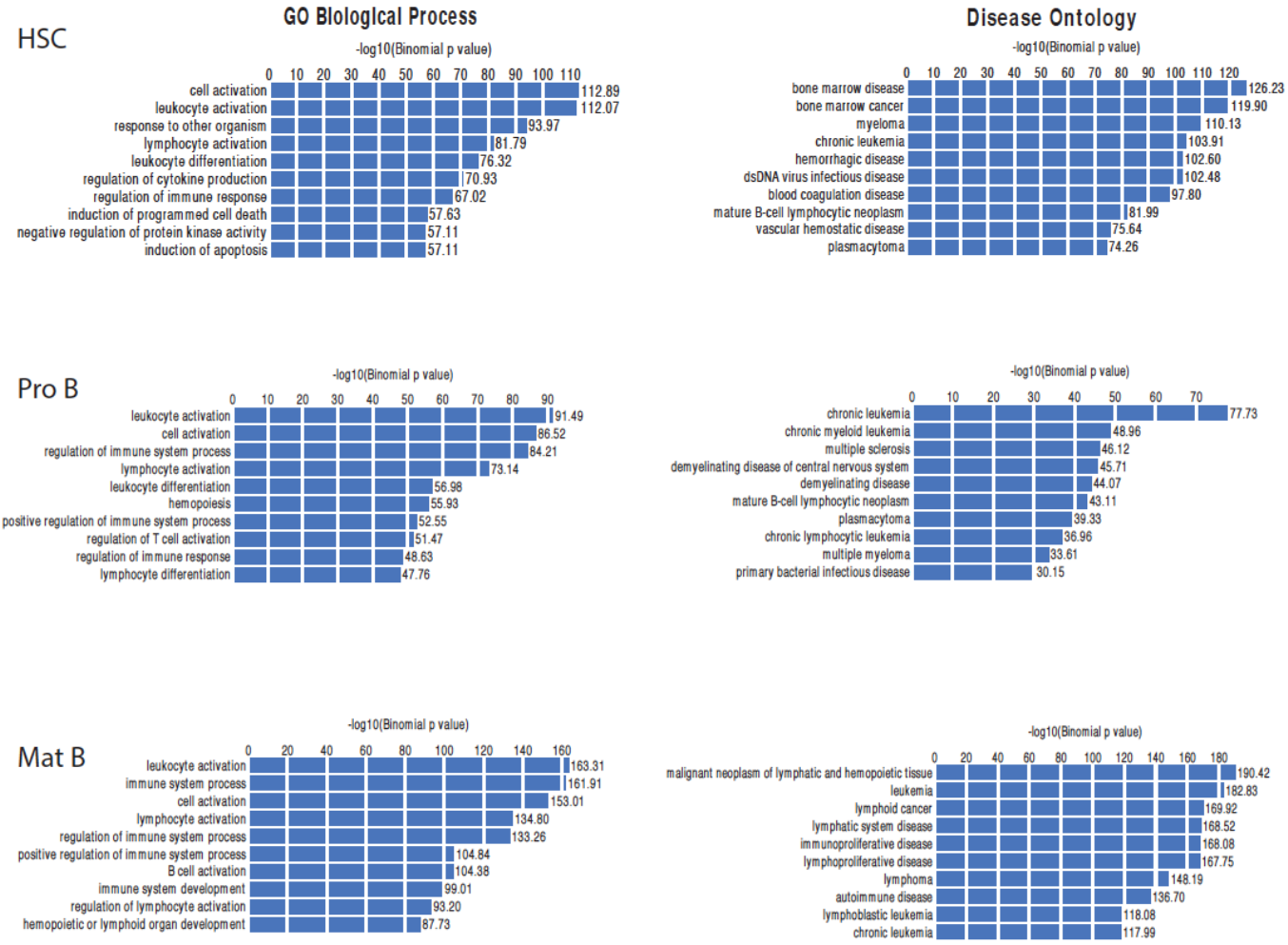
Supplementary Figures



Supplementary Fig. 1: Nup-Hoxb4-HSCs are able to repopulate the hematopoietic system of irradiated Rag2 deficient mice.

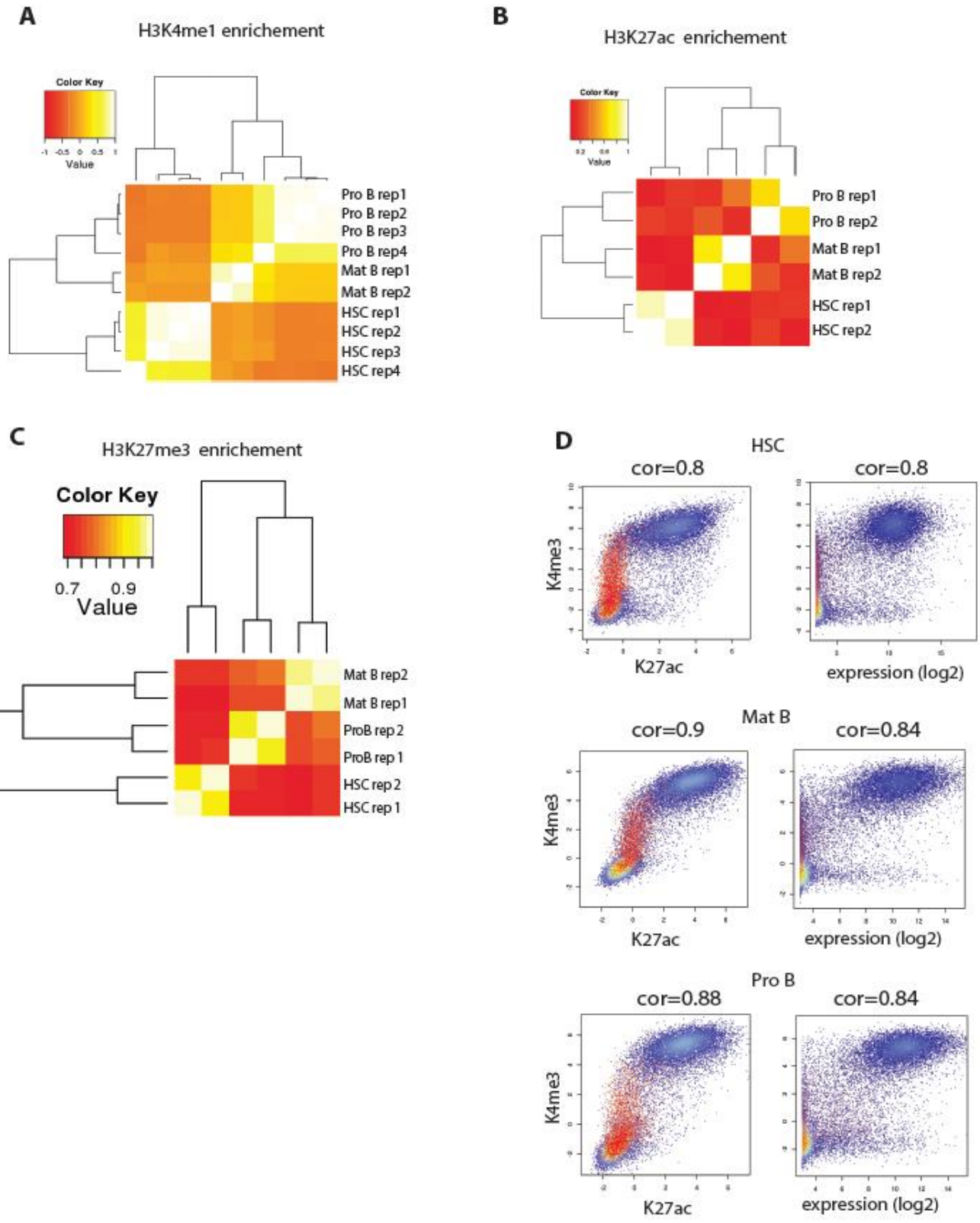
A. Nup-hox4-HSCs were injected into irradiated Rag2 deficient mice. After 4 weeks, LT-HSCs (Lineage-, Sca+, c-kit+, flt3-) and ST-HSCs (Lineage-, Sca+, c-kit+, flt3+) were isolated from BM of recipient mice.

B. Pro B cells (B220+, c-kit+, CD19-) and Pre B cells (B220+, c-kit+, CD19+) were also isolated from recipient mice.



Supplementary Fig. 2: Biological functions and diseases associated with target genes of active enhancers in HSCs, Pro B and mature B cells.

Genomic coordinates of active enhancers were associated to their putative target genes; the resulting set of genes was submitted to gene annotation terms for biological processes and diseases using the Great software (http://bejerano.stanford.edu/) with default settings. Top ten hits and their associated p-values are shown.

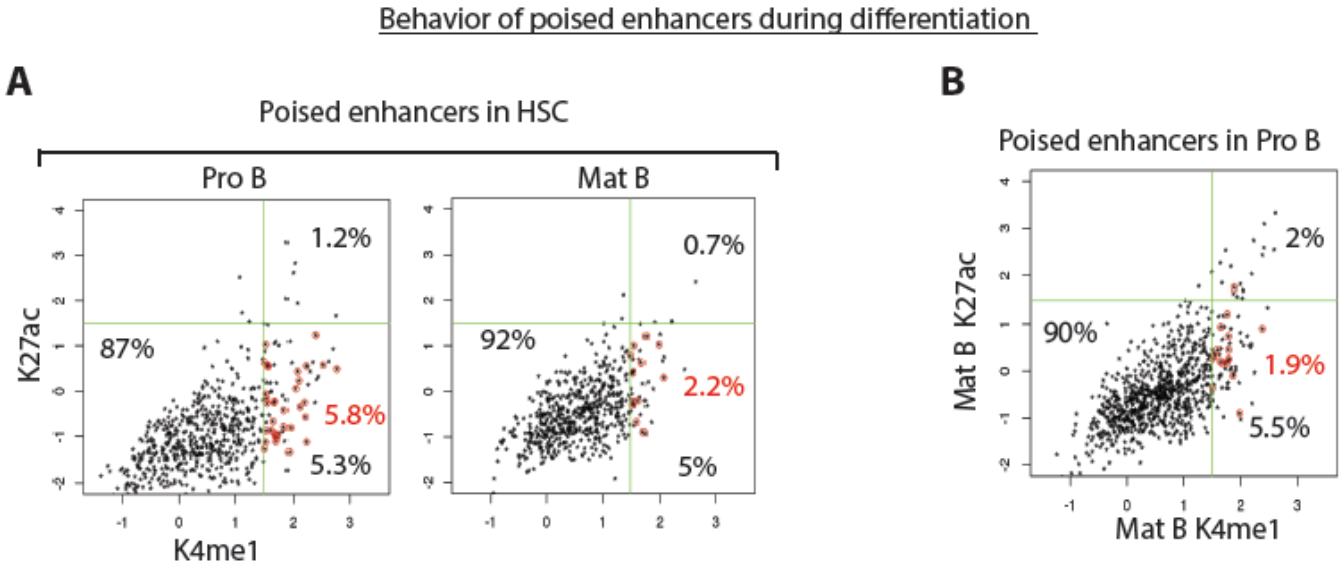


Supplementary Fig. 3: Correlation between replicates across samples.

A and B. Clustering of H3K4me1 (A) and H3K27ac (B) enrichment at enhancer elements based on Pearson correlation. Biological replicates for each stage cluster together.

C. Clustering of H3K27me3 enrichment at promoter regions.

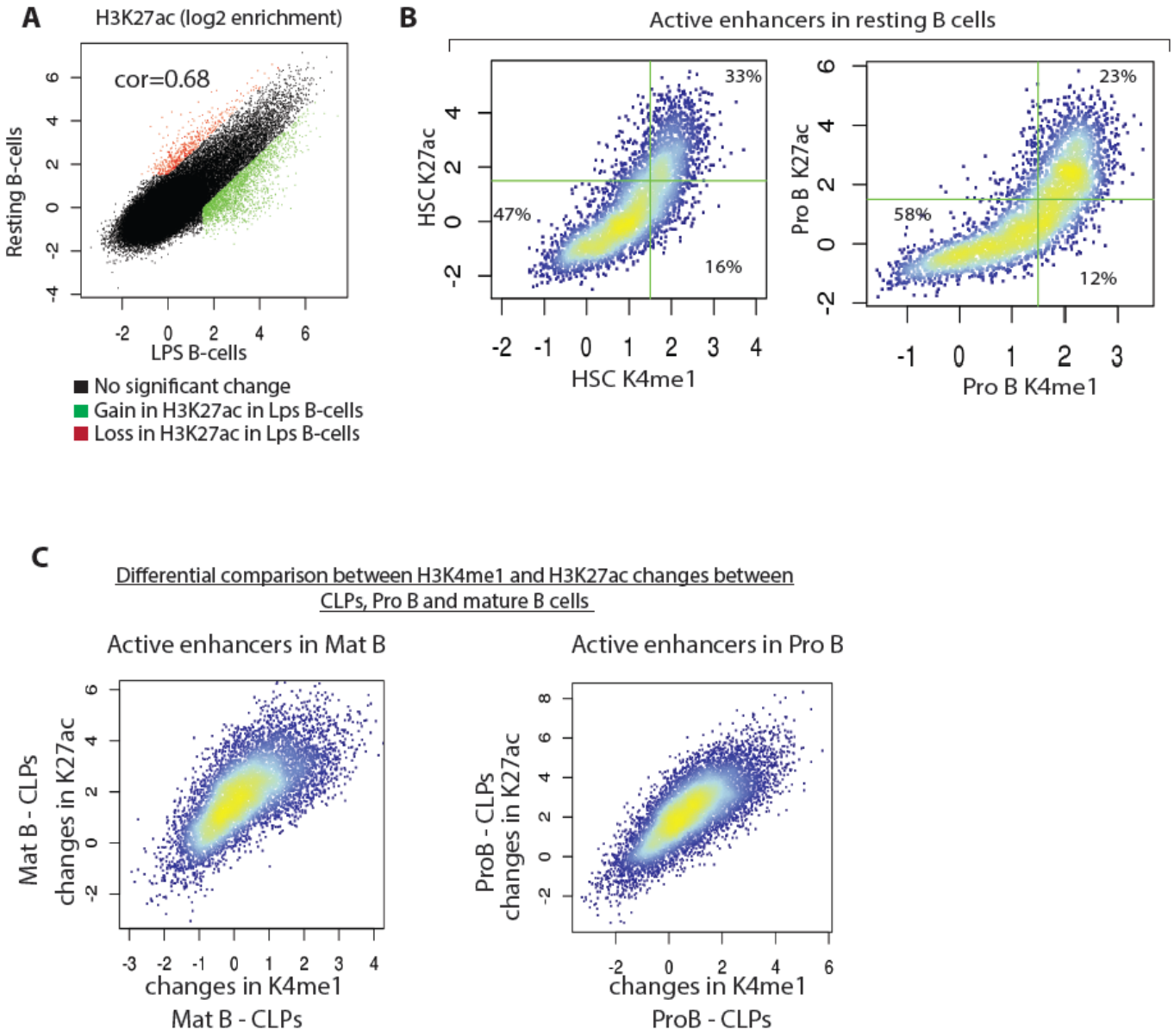
D. Scatter plots of H3K4me3, H3K27ac and RNA sequencing signals at promoter regions. Promoters enriched in H3K27me3 are colored in red. Cor stands for Pearson correlation coefficient.



Supplementary Fig. 4: Behavior of poised enhancers during differentiation.

A. The chromatin state of enhancers poised in HSCs was investigated in Pro B and mature B cells. H3K4me1 and H3K27ac enrichment in the indicated cell types were calculated at genomic coordinates corresponding to poised enhancers in HSCs. Green lines indicate cut-offs used to select enriched regions for each signal. The proportions of the different population are indicated in the scatter plots. Poised enhancers in the indicated stages are marked by red circles and their proportions are indicated in red.

B. Similar to A, the chromatin state of enhancers poised in Pro B cells was investigated in mature B cells.

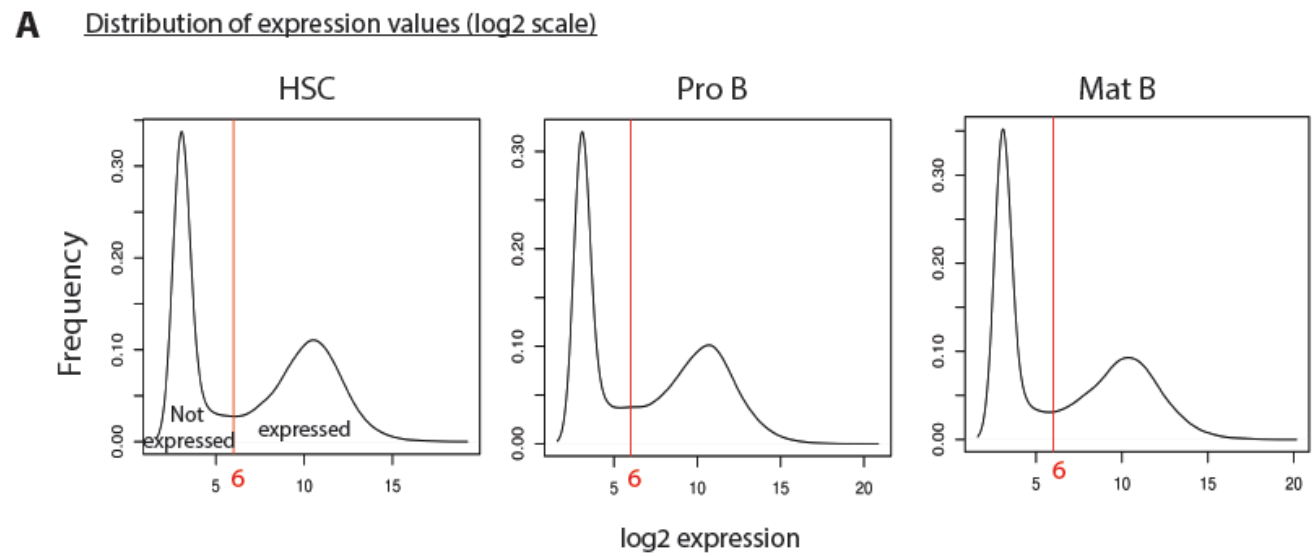


Supplementary Fig. 5: Comparison of active enhancers in resting and LPS-activated B cells.

A. H3K27ac enrichment (over input) in 1kb sliding windows across the genome for LPS-stimulated and resting B cells. Loci with no change in H3K27ac signal between the two conditions are colored in black, loci showing an increase in H3K27ac signal in LPS stimulated cells compared to resting cells are colored in green and loci with a decreased H3K27ac enrichment are colored in red. A 1.5 fold cut-off was used to select differentially enriched loci.

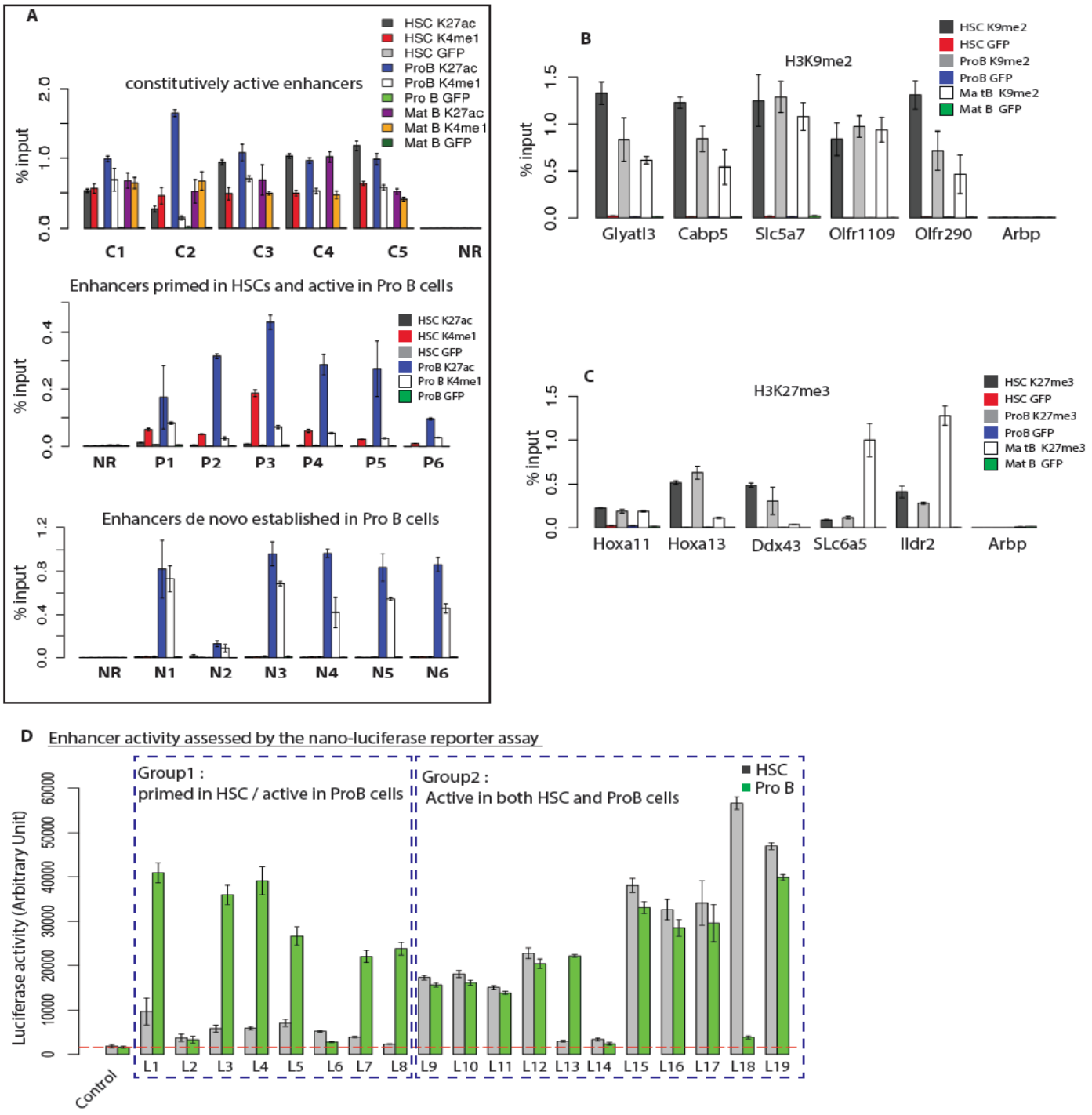
B. The chromatin state of active enhancers in resting B cells (identified based on H3K27ac signal) was investigated in HSCs (left panel) and Pro B cells (right panel).

C. Scatter plots for changes in H3K27ac and H3K4me1 between mature B cells and CLPs (left panel) and between Pro B cells and CLPs (right panel).



Supplementary Fig. 6: Gene classification according to expression level.

A. Histograms showing the distribution of the log2 expression levels as determined by RNA-sequencing experiments (see Material and Methods) in HSCs, Pro B and mature B cells. The cut-off to classify a gene as expressed was set to 6 (red horizontal line).

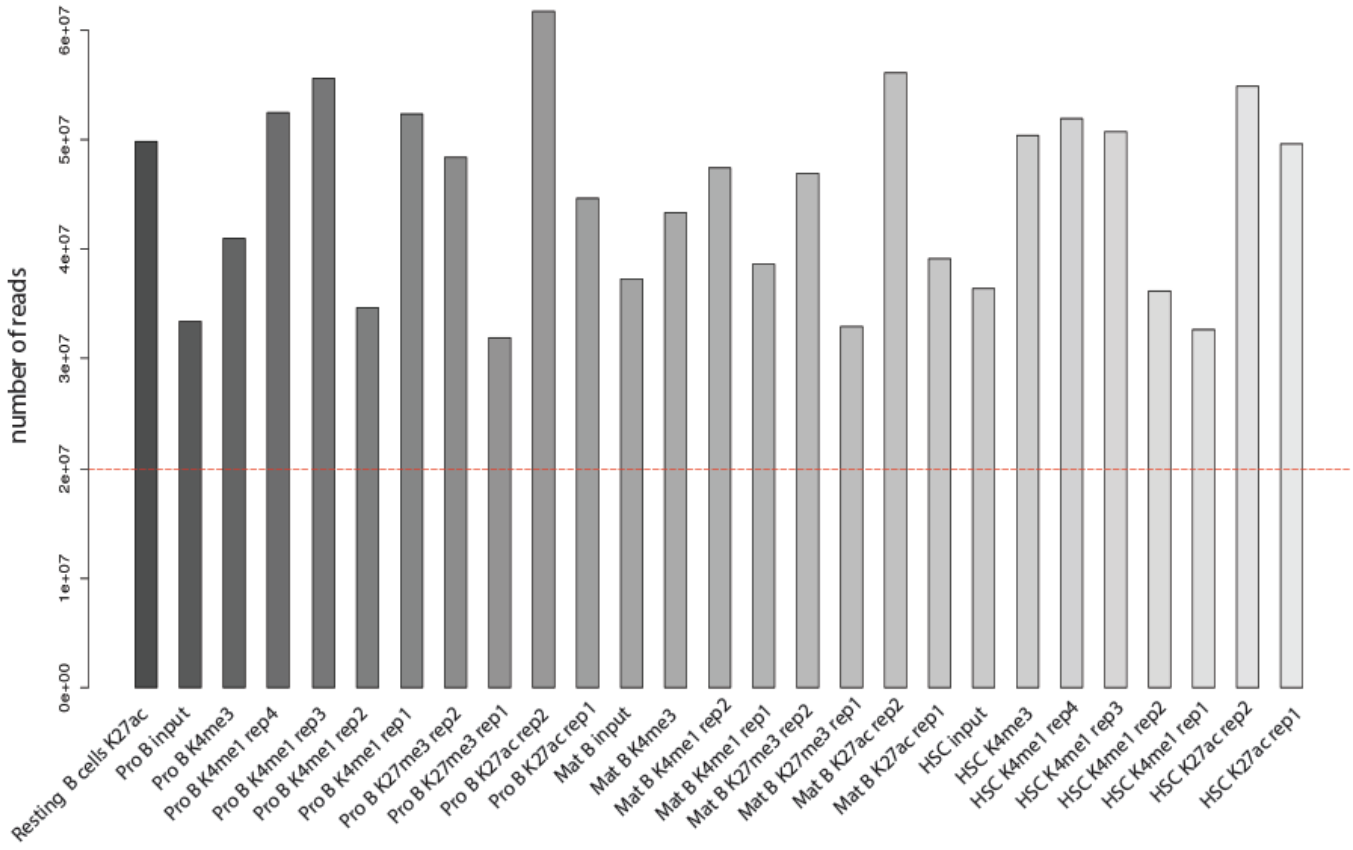


Supplementary Fig. 7: ChIP qPCR and reporter assays validation of ChIP sequencing data.

A. A subset of enhancers from different categories were validated by independent ChIP qPCR experiments; the first panel represents constitutively active enhancers (C1 to C5), the second panel represents enhancers primed in HSCs and active in Pro B cells (P1 to P6) and the third panel represents enhancers de novo generated in Pro B cells (N1 to N6). The testis specific locus Protamine1 was used as a negative control. Genomic Coordinates of these loci are provided into supplementary Table 3.

B and C. A subset of genes enriched in H3K9me2 (B) and H3K27me3 (C) marks were investigated by ChIP qPCR; the highly expressed ribosomal protein gene *Arbp* was used as a negative control.

D. Luciferase activity for a subset of primed and active enhancers in HSCs and Pro B cells. The coordinates of the loci used in this figure are given in the table S1.



Supplementary Fig. 8: Barplot representing the number of reads for each ChIP-sequencing experiment.

Supplementary Tables

Supplementary Table 1: Coordinates of enhancers used in reporter assays (Figure 1D, Supplementary Fig. 7D).

Locus Name	Coordinates
L1	chr11: 100591638-100592913
L2	chr2:118324225-118325286
L3	chr19: 47612048- 47613191
L4	chr1: 129108567-129109172
L5	chr2: 90371166-90372676
L6	chr1:130571787-130572987
L7	chr7:133559137-133561153
L8	chr10:94605090-94606684
L9	chr2:144056107-144057375
L10	chr11: 100591638-100592913
L11	chr2:118324225-118325286
L12	chr9:45081937-45082864
L13	chr4:9601957-9602736
L14	chr4: 9603272- 9603860
L15	chr17:47658318-47659541
L16	chr9:13897323-13898563
L17	chr7:73378119-73379155
L18	chr16:92620823-92621911
L19	chr2:90921161-90925642

Supplementary Table 2: Public data sets used in this study.

GEO sample ID	Sample
GSM1441284	H3K27Ac_CLP
GSM1441300	H3K4me1_CLP
GSM769031	Spleen H3K4me1
GSM1000138	Spleen H3K27ac
GSM769036	Spleen H3K4me3
GSM769037	Spleen input
GSM918705	Thymus input
GSM1000101	Thymus H3K4me3
GSM1000102	Thymus H3K4me1
GSM1000103	Thymus H3K27ac
GSM1164635	HSC H3K27me3 rep1
GSM1164636	HSC H3K27me3 rep2

Supplementary Table 3: Coordinates and qPCR primers for enhancers verified by ChIP-qPCR in Supplementary Fig. 7A.

Locus Name	Coordinates	fw primers	rev primers
C1	chr2:144073306-144073986	ttcagctggacacacacaca	ggctgtcctggaactctgtc
C2	chr13:103482392-103483524	gtggtgtccattgtgagtgc	atggcctactgcatttcctg
C3	chr9:45081937-45082864	ctggggaacagatttgcatt	aattctccctttccctccaa
C4	chr17:47658919-47660312	agttggctcaatgggaacac	aaccactgttcaggggtcag
C5	chr2:90921161-90925642	aggtcatcaagttccgcaga	ctgagtgggtgggttctgat
P1	chr18:35974566-35975392	gattgcccagagtgtggagt	gcccacctggatctaactca
P2	chr11:100589876-100591397	gcaggggagtcagcagtaag	aagactgttctccggcttca
P3	chr2:118324225-118325286	ggaagctgtgtggaaagctc	gtctcctaaggcagcaggtg
P4	chr19:47612048-47613191	tcagcctccctcctgtagaa	atatcgatcggccttgaatg
P5	chr1:129108567-129109172	acgcagtggaaggagaagaa	agctgatgatcccatgatcc
P6	chr2:90371166-90372676	accatctcggggaaaactct	gctggatggtggcagtaaat
N1	chr7:108617991-108619209	tgaggtggcaatgaaatgaa	caggtcccacacagtcattg
N2	chr13:73847304-73848770	gccagccagctgtttcttac	gctaaggatcgcagtcaagc
N3	chr11:45351843-45353015	ttctgcaacatcctcactgc	catgaaagcggagacacaga
N4	chr5:137154497-137155688	gtaaaagctgtgggctgagg	acacccacaggtgagactcc
N5	chr5:83382743-83384187	gaatgccccacgttaagaga	atgttctcccagcaaaccac
N6	chr4:44683297-44688461	agcgagttgtgaaggtcaga	catcacgcagcagaactctc

Supplementary Table 4: Primers of genes verified by qPCR in Supplementary Fig. 7B and C.

Gene name	fw primers	rev primers
Glyatl3	ctgggtttcccttctcatca	accccgagactgacctacct
Cabp5	aggaggagttggcaaaggat	gaaactgcattggagcaggt
Slc5a7	cccccaaaaacatcaaatct	tgcatataatggttccaaaaca
Olfr1109	tccccagttacccacaaaaa	catggacccagggaatcata
Olfr290	tggggtcattactcccatgt	accaaaagacacaggccaac
Arbp (rplp0)	gtcgatggaaccagccaata	cctcccacaacaaaacaacc
Hoxa11	aggagccttcctttcagctc	ggccctttcaaagtctttcc
Hoxa13	ccccttccatgttcttgttg	cttcacttcttgggggcttt
Ddx43	cattagcccgtcatgaacct	gctagtgtctggctgggaac
Slc6a5	cattagcccgtcatgaacct	gctagtgtctggctgggaac
Ildr2	tcttatcgtgccagctgatg	acgaaggtggagtggaacac

2.3 Appendix

2.3.1 The reversed regulatory effects between PU.1 and OBF1

We found that OCT factors and OBF1 colocalize with several TFs from ETS family, and genome-wide colocalization between OCT1/OCT2/OBF1 and PU.1/ETS1 has been confirmed experimentally (**Result 2.1**). To uncover genes co-regulated by these TFs, we compared the differentially expressed genes (DEGs) between *Pou2af1*$^{-/-}$ and *Spi1*$^{-/-}$ B cells stimulated with LPS or anti-CD40/IL4. We found that OBF1 showed a much stronger effect than PU.1, as more DEGs were discovered in *Pou2af1*$^{-/-}$ B cells than in *Spi1*$^{-/-}$ B cells under both stimulations (Figure 15A-15D). In fact, only a minor overlap of DEGs from *Pou2af1*$^{-/-}$ and *Spi1*$^{-/-}$ B cells was discovered (Figure 15E and 15F). Surprisingly, OBF1 and PU.1 seem to play contrasting roles in *in vitro* stimulated B cells, in that the genes downregulated and upregulated in *Spi1*$^{-/-}$ B cells are upregulated and downregulated in *Pou2af1*$^{-/-}$ B cells, respectively (Figure 15G). As discussed in **3.4**, PU.1 negatively regulates PC differentiation[253], while OBF1 is one of the master regulators that control the expression of BLIMP1 and promote PC differentiation[239,254]. Moreover, PU.1 and OBF1 play reversed roles in controlling follicular B cell signature genes: PU.1 is important in maintaining follicular B cell homeostasis[253], while follicular genes are upregulated in the absence of OBF1.

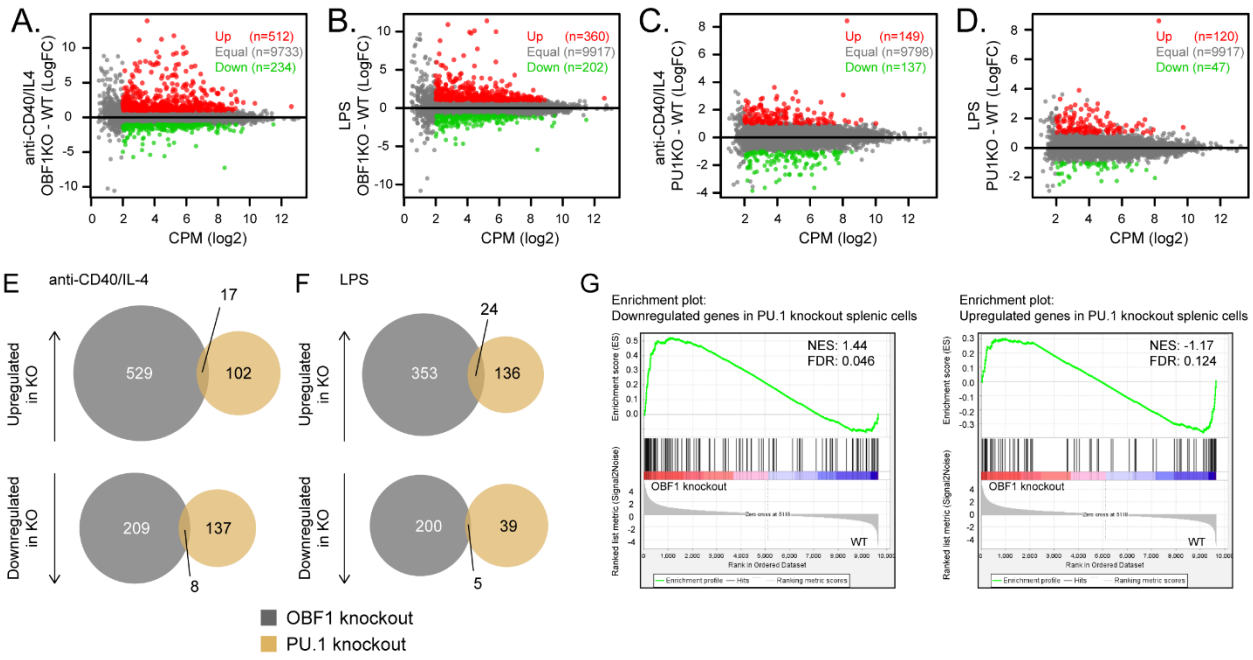


Figure 15. Comparison of the transcriptome regulated by OBF1 and PU.1.

A. Average expression (log2 Counts Per Million, x-axis) and change of expression (log2 Fold Change, y-axis) of genes in WT and *Pou2af1*$^{-/-}$ CD19^{+} splenic B cells stimulated with anti-CD40/IL4. Differentially expressed genes are color-coded (DEGs fold change ≥ 1, FDR < 0.01, logCPM > 2): red data points represent genes with higher expression level in *Pou2af1*$^{-/-}$ cells, green data points represent genes with higher expression in WT cells.

B. Average expression (log2 counts per million, x-axis) and change of expression (log2 Fold Change, y-axis) of genes in WT and *Pou2af1*$^{-/-}$ CD19^{+} splenic B cells stimulated with LPS. Differentially expressed genes are color-coded (DEGs fold change ≥ 1, FDR < 0.01, logCPM > 2): red data points represent genes with higher expression level in *Pou2af1*$^{-/-}$ cells, green data points represent genes with higher expression in WT cells.

C. Average expression (log2 Counts Per Million, x-axis) and change of expression (log2 Fold Change, y-axis) of genes in WT and *Spi1*$^{-/-}$ CD19^{+} splenic B cells stimulated with anti-CD40/IL4. Differentially expressed genes are color-coded (DEGs fold change ≥ 1, FDR < 0.01, logCPM > 2): red data points represent genes with higher expression level in *Spi1*$^{-/-}$ cells, green data points represent genes with higher expression in WT cells.

D. Average expression (log2 Counts Per Million, x-axis) and change of expression (log2 Fold Change, y-axis) of genes in WT and *Spi1*$^{-/-}$ CD19^{+} splenic B cells stimulated with LPS. Differentially expressed genes are color-coded (DEGs fold change ≥ 1, FDR < 0.01, logCPM > 2): red data points represent genes with higher expression level in *Spi1*$^{-/-}$ cells, green data points represent genes with higher expression in WT cells.

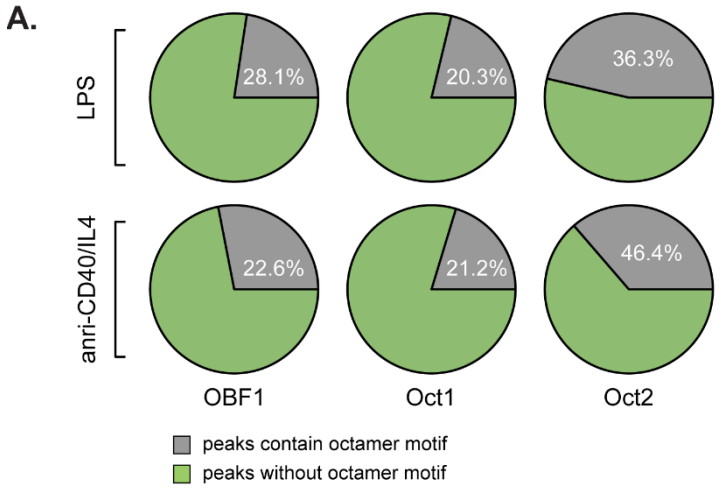
E. Venn diagrams showing the overlaps between significantly upregulated (upper panel) and downregulated (bottom panel) genes identified in *Pou2af1*$^{-/-}$ and *Spi1*$^{-/-}$ B cells under anti-CD40/IL4 treatment.

F. Venn diagrams showing the overlaps between significantly upregulated (upper panel) and downregulated (bottom panel) genes identified in *Pou2af1*$^{-/-}$ and *Spi1*$^{-/-}$ B cells under LPS treatment.

G. Gene set enrichment analysis of relative gene expression in *Pou2af1*$^{-/-}$ versus WT anti-CD40/IL4 stimulated splenic B cells against the gene set identified as downregulated (left) and upregulated (right) upon *Spi1*$^{-/-}$ splenic B cells. FDR, false discovery rate; NES, normalized enrichment score.

2.3.2 Octamer motif density in peaks

We performed Bio-ChIP-seq for OCT1, OCT2 and OBF1 under LPS and anti-CD40/IL4 stimulation, and reported that the octamer motif is among the top enriched motifs by HOMER in each case. Next, we would like to know the percentage of the peaks that contain octamer motifs. Therefore, we annotated octamer motif (JASPAR: MA0785.1) to the binding regions of each factor under both stimulations. Surprisingly, we found that octamer motifs only occur in a relatively small percentage of peaks of OCT1/OCT2/OBF1, even if the octamer motif is overrepresented in motif analysis using HOMER (as discussed in **3.5)** (Figure 16A). However, we found that the frequency of octamer motifs seems to be correlated with peak binding enrichment, as we found a high percentage of top enriched peaks contain octamer motifs (Figure 16B). The overall low frequency of octamer motif occurrences is not understood yet. This observation indicates uncover binding characteristics of OCT factors, including novel binding sites, degenerative binding sites, or unknown interacting DNA binding proteins, which warrants further study.



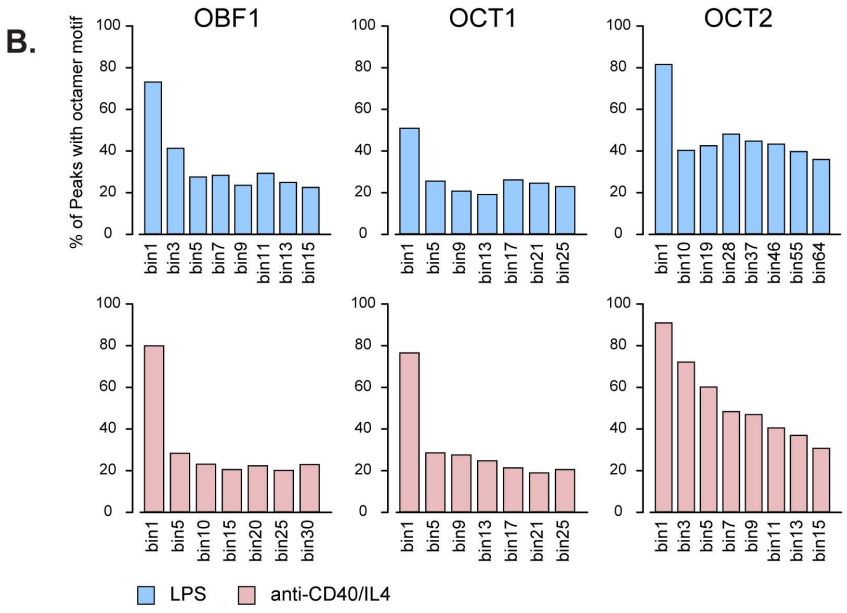


Figure 16. The frequency of octamer motif in OCT1, OCT2 and OBF1 binding regions.

A. Pie charts show the frequency of octamer motifs in the total peaks of OBF1, OCT1 and OCT2 identified in B cells under the corresponding stimulation.

B. Peaks of OBF1, OCT1 and OCT2 identified in B cells under corresponding stimulations are ranked and binned based on the ChIP enrichment signal (log_2 ChIP/Input) with each bin of 500 ranked peaks. The percentage of the octamer motif in the selected ranked bins are visualized in barplot.

2.3.3 Cell cycle analysis of OBF1-depleted lymphoma cells

We found that the proliferation rate of Raji and Ramos cells is dramatically reduced when OBF1 is depleted by knockdown. We then performed cycle analysis for WT and OBF1-depleted Raji and Ramos cells by DAPI staining and found that cells are arrested in G1 phase when OBF1 is reduced (Figure 17).

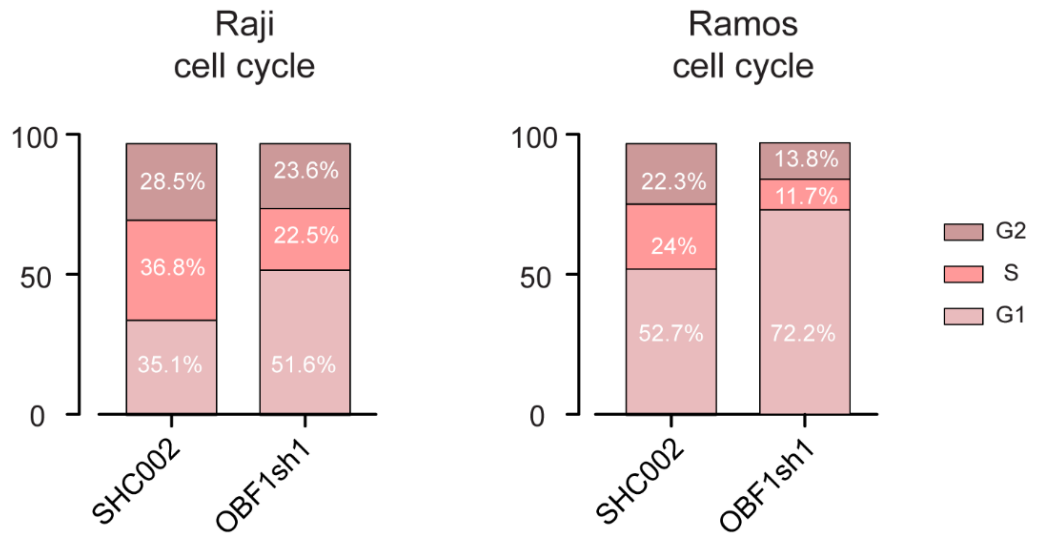


Figure 17. Cell cycle analysis of OBF1-depleted Raji and Ramos cells.

Left, cell cycle analysis of Raji cells infected with control shRNA (SHC002) and OBF1-specific shRNA (OBF1sh1) using DAPI staining by FACS; percentage of cells of different stages are indicated.

Right, cell cycle analysis of Ramos cells infected with control shRNA (SHC002) and OBF1-specific shRNA (OBF1sh1) using DAPI staining by FACS; percentage of cells of different stages are indicated.

2.3.4 Interaction between OBF1 and PU.1

I found that OCT factors and OBF1 are extensively colocalized with PU.1 (**2.1**). To understand whether PU.1 and OCT factors or OBF1 interact with each other, I performed immunoprecipitation with anti-OBF1 antibody. By western blot, I detected a very weak PU.1 band in OBF1 immunoprecipitated product (Figure 18). However, it is difficult to conclude if PU.1 forms a complex with OBF1. Because OBF1 or OCT factors and PU.1 are colocalized on DNA and I cannot exclude the presence of some genomic DNA in this experiment; thus, the observed PU.1 signal could be ascribed to the possibility that PU.1 is co-immunoprecipitated with OBF1 by binding to the same DNA fragments instead of forming a protein complex.

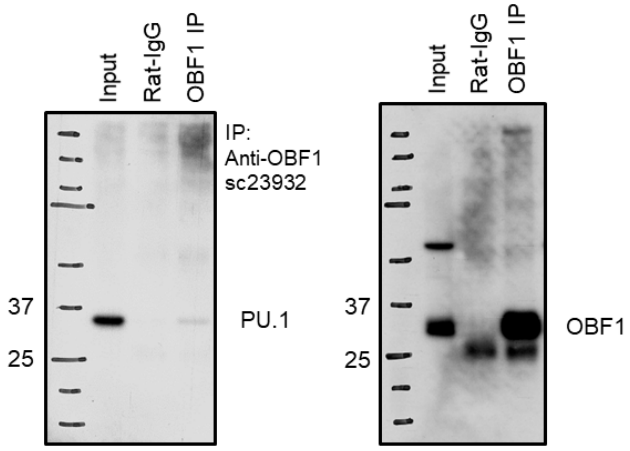


Figure 18. The frequency of octamer motif in OCT1, OCT2 and OBF1 binding regions.

Cell lysates of PD31 cells were immunoprecipitated with anti-OBF1 antibody followed by immunoblotting with anti-PU.1 antibody (left) and anti-OBF1 antibody (right).

2.3.5 OBF1 facilitates immune evasion of GC-derived B lymphoma

The proliferation of GC-derived B lymphoma cells is OBF1-dependent ([236,251,252]and this thesis). Mechanistically, OBF1 maintains the proliferation of B lymphoma cells by activating BCL6 and repressing IRF4. In addition, genes associated with favorable prognosis were upregulated in OBF1-deficient Raji cells. In line with this, we found that OBF1 is engaged in immune evasion by repressing the expression of genes that are essential in immunosurveillance, natural killer cell function, and apoptosis, as knocking down OBF1 relieving the repression of these genes (Figure 19).

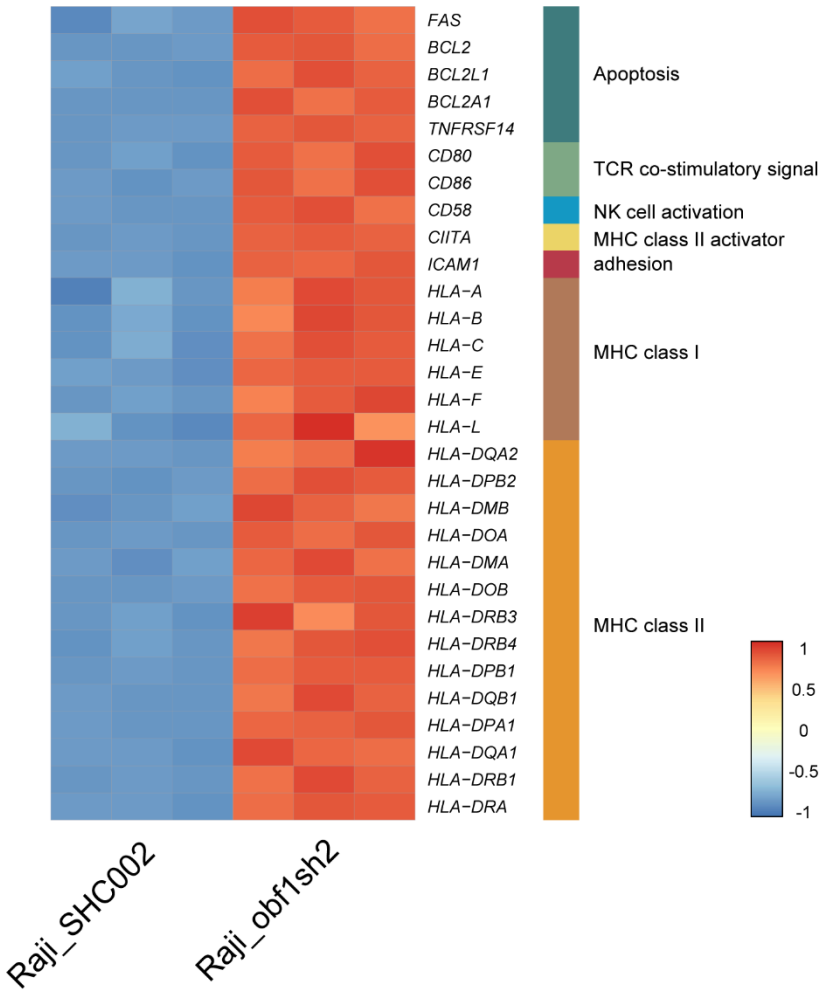


Figure 19. Upregulation of immune evasive genes in OBF1-deficient Raji cells.

RNA-seq analysis of OBF1-deficient Raji cells versus control Raji cells. Heatmap represents row z-scores of RPKM normalized read counts for differentially expressed genes. Genes involved in immune evasion were selected; the immunological functions of these genes are indicated.

Discussion

Chapter 3 Discussion

3.1 OBF1 stabilizes the binding of Oct factors

Former studies showed that, although OBF1 is a coactivator and lacks DNA binding domain, it seems to stabilize the binding of OCT factors on octamer sites [255,256]. Studies performed on selected binding sites had concluded that OBF1 does not or only weakly stabilizes the binding of OCT factors on the DNA[162,200,257]. In contrast, OBF1 was found to enhance the binding of OCT1 POU domain on the octamer motif in the histone 2 promoter[256]. Another study demonstrated that OBF1 not only significantly stabilizes the dimer formation of OCT1 POU domains on PORE motif and enhances the transcriptional activity of OCT1, but also potentiates the transcriptional activity of OCT1 on PORE motifs harboring various mismatches[255]. In another study, Strubin et al claimed that OBF1 failed to stabilize the binding of OCT1 on a canonical octamer site in an off-rate EMSA study[196]. Thus, whether or not OBF1 stabilizes the binding of OCT factors on the DNA was not conclusively resolved.

Instead of analyzing the stabilizing effect of OBF1 on individual DNA sequences, our results provide a holistic platform to interrogate whether OBF1 is able to stabilize the genomic binding of OCT factors or not. By comparing the Bio-ChIP-seq data of OCT1 in WT and *Pou2af1*$^{-/-}$ B cells, we conclusively demonstrated that OBF1 plays a fundamental role in stabilizing the binding of OCT1 across the genome (**2.1** Fig. 4F). While OBF1 also seems to be effective in stabilizing the genomic binding of OCT2 (**2.1** Fig. 4C and S3B), it is difficult to reach a definitive conclusion experimentally under our project setting, as the protein level of OCT2 decreased significantly in *Pou2af1*$^{-/-}$ B cells (**2.1** Fig. S3C); therefore, reduced binding might also be due to the lower OCT2 protein level. By contrast, PU.1, which often co-localized with OCT1 and OCT2, does not stabilize their binding.

Whether the stabilizing effects of OBF1 on OCT1 binding is motif-dependent or not is still not entirely clear but seems likely. EMSA studies with individual DNA sequences resembling canonical and mutant octamer motifs revealed that OBF1 facilitates the binding of OCT1 monomer on classic octamer motif[256], and enables OCT1 dimer formation on PORE motifs that were originally considered to preclude the binding[255]. OBF1 functionally recruits POU_S and POU_H subdomains to promoter DNA and acts as a molecular clamp through the interaction with

both subdomains of the POU domain, as well as the DNA[209]. In my analysis, motif discovery suggested that the octamer motif enrichment is highly different between peaks that are weaker when OBF1 is ablated (as in OBF1 KO B cells) and those that are equal between both genotypes. The octamer motif is highly enriched in the OCT1 peaks which exhibit significantly reduced enrichment when OBF1 is ablated, while it is barely detectable in OCT1 peaks which are insignificantly changed between WT and *Pou2af1*$^{-/-}$ B cells (Figure 20). Taken together, OBF1 significantly stabilizes the global binding of OCT1 on its binding regions containing octamer motifs.

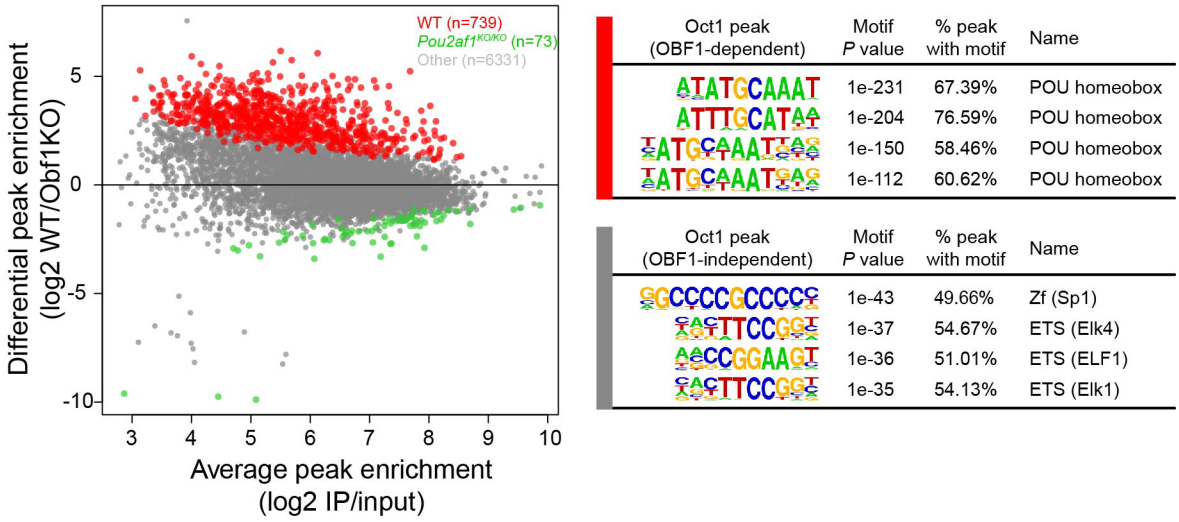


Figure 20. Motif analysis for OCT1 peaks in WT and *Pou2af1*$^{-/-}$ B cells. Left, differential OCT1 binding regions in WT and *Pou2af1*$^{-/-}$ CD19^{+} splenic B cells stimulated by LPS (red dots, OCT1 binding regions showing significantly weaker binding in *Pou2af1*$^{-/-}$ samples than in WT; green dots, OCT1 binding regions showing significantly stronger binding in *Pou2af1*$^{-/-}$ samples than in WT). Right, motif analysis of regions differentially bound by OCT1 in WT and *Pou2af1*$^{-/-}$ B cells using HOMER.

In addition to the stabilizing effects of OBF1 at a structural level, the modulation of chromatin organization by OBF1 adds another layer of functional relevance in promoting the binding of OCT factors. H3K27ac level is reduced on the OCT1 binding regions which showed weaker enrichment of OCT1 upon the depletion of OBF1, indicating that chromatin structure is affected by the loss of OBF1, whereby the accessibility of OCT1 is restricted. Moreover, another study showed that OBF1 and HDAC3 compete with each other to bind to TFII-I at promoter regions.

In this way, the acetylation level of local chromatin is modulated. As a result, loss of OBF1 is coupled with reduced H4K5ac[240]. It has also been demonstrated that OBF1 is critical for the formation of long-range chromatin interaction between the *IgH* promoter and 3' *IgH* enhancer[240], by which it relays the transcriptional activity of enhancer-bound OCT factors to the target promoters. This proves to be essential for driving the expression of IgH in MPC11 cells, and maintaining the proliferation of GC-derived B lymphoma cells with translocations that place oncogenes to the proximity of 3' *Igh* enhancer, such as the activation of t(14;18) translocated *Bcl2* gene[258,259]. The activity of OBF1 in facilitating chromatin loops provides an explanation on why octamer motifs exist in only a portion of the OCT factor peaks (Figure 14) because OCT factors could be recruited to distal sites that contain no octamer motif through chromatin looping.

3.2 Colocalization between ETS factors and OCT factors

It is a widely accepted concept that significant co-enrichment of motifs of certain TFs in the binding regions of a particular TF indicates a possible colocalization between these two groups of TFs. A previous study found that octamer motif was discovered in the PU.1 binding regions[260]. Similarly, *de novo* motif analyses of OCT1, OCT2 and OBF1 peaks revealed strong enrichment of binding motifs for ETS factors (**2.1** Fig. S2D) in splenic B cells. In agreement with this finding, ChIP-seq revealed extensive colocalization between OCT factors/OBF1 and PU.1 peaks.

ETS factors are critical regulators that regulate B cell fate commitment, B cell development, and immunological functions, as well as B cell malignancies [261-263]. ETS motifs are highly conserved in terms of the core sequence (5'-GGAA-3'), it was important to investigate how many ETS factors, beyond PU.1, colocalize with OCT factors. A simple statistical model by comparing the original and randomized motif counts of various ETS factors in peak regions of OCT factors might provide a smaller set of ETS factors with a higher probability in co-binding with OCT factors or OBF1.

In an attempt to understand if additional ETS factors are colocalized with OCT factors, we mined our ChIP-seq data by counting the occurrences of the motifs of ETS factors that are expressed in B cells by RNA-seq data. Then, we counted the occurrences of original and column-randomized motif weight matrices at each nucleotide position of a 500bp window surrounding peak summits.

The *P*-values for each original weight matrix over randomized controls were calculated with the 95th percentile counts of original and randomized motif weight matrices using the following formula:

$$P = (\text{sum } (95^{th}\text{-random-count} > 95^{th}\text{-motif-count}) + 1) / (500 + 1).$$

In this way, we found that *P*-values of ETS1 and ETV6 were highly significant and that these additional ETS factors might colocalize with OCT factors and OBF1. Indeed, by comparing to publicly available ChIP-seq data of ETS1 in mature B cells (GSE83797)[264], we observed a considerable overlap between ETS1 and OCT1/OCT2/OBF1 even to a larger extent than PU.1. This result supports the results from computational analysis, and it could be used as a preliminary approach to screen the binding of a particular TF from a family of TFs that share similar binding motifs. Although ETV6 ChIP-seq data is not available yet, it is fair to assume that this factor may behave similarly to PU.1 and ETS1. Thus, the significant colocalization between OCT1/OCT2/OBF1 and several ETS factors in B cells and the extent of overlap between these TF networks had not been appreciated.

One issue in this analysis that randomizes the columns of positional weight matrices (PWM) might introduce biases to the final results when the core binding sequence is highly conserved. In this case, 5'-GGAA-3' is the core sequence in each ETS motif. Shuffling the PWM might create a number of randomized motifs containing exactly the identical core sequence, thus resulting in higher *P*-value, which then leads to false-negative rejections. Two improvements can be considered for future analysis. One approach is to randomize the DNA sequence of each window, and count the occurrence of ETS motif in peak regions of original and randomized DNA sequences. Another approach is to randomize the PWM not only by columns but also by rows, so that the weight of each position is changed and the chance of generation randomized motifs with core sequence is minimized.

Although we found the extensive colocalization between ETS factors and OCT1/OCT2/OBF1, we did not detect physical interaction between them. We tried immunoprecipitation of OBF1 in WT mouse B cell line PD31, but failed to detect PU.1 or ETS1 in western blot. Similar results were observed in 293T cells with ectopic expression of both OBF1 and PU.1. Therefore, although ETS1 and PU.1 colocalize with OCT factors, these results suggest that they do not appear to function in a complex.

The colocalization between ETS factors and OCT1/OCT2/OBF1 suggests that genes bound by these factors might be coregulated. Genes that are coregulated by ETS factors and TFs from other families have been uncovered by previous studies. *Tal1* and *Lyl1*, important genes for the specification and function of HSCs, including *Fli1*, *Erg*, and *Elf1*, are coregulated by ETS factors and GATA2 [265-268]. Another study identified genes that are coregulated by GABP and YY1 in several cell types[259]. We performed RNA-seq analysis with splenic B cells isolated from WT, *Spi1*$^{-/-}$ and *Pou2af1*$^{-/-}$ mice. However, only minor overlaps were identified among the differentially expressed genes of *Spi1*$^{-/-}$ or *Pou2af1*$^{-/-}$ B cells (24 commonly upregulated and 5 downregulated genes) (**2.1** Fig. 3F-3H). While the functional relevance of these observations will require further systematic analysis, given the number of different ETS factors involved in co-localization with OCT1/OCT2/OBF1, it seems likely that in some cases these factors will be co-regulating their targets

Together, our results indicate that OCT1/OCT2/OBF1 generally tends to colocalize with ETS factors in B cells, but the functional consequence is unclear.

3.3 Reversed regulatory roles of PU.1 and OBF1

Unexpectedly, GSEA analysis, a method to compare established gene sets associated with diseases of biological processes with RNA-seq or microarray data, using signatures in *Pou2af1*$^{-/-}$ or *Spi1*$^{-/-}$ B cells, revealed a reversed transcriptional correlation between the transcriptomes of these two genotypes under both LPS and anti-CD40/IL4 stimulation (Figure 15). Genes upregulated or downregulated in *Pou2af1*$^{-/-}$ B cells are downregulated or upregulated in *Spi1*$^{-/-}$ B cells, respectively. This inversed transcriptional pattern might stem from the functional distinctions of OBF1 and PU.1 on PC differentiation.

PU.1 was reported to negatively regulate the differentiation of PCs under *in vivo* and *in vitro* settings[253]. In keeping with this finding, we performed GSEA analysis and found that *Spi1*$^{-/-}$ B cells exhibit a partial upregulation of PC signatures. Genes involved in PCs, such as *Eaf2*, *Prdm1* and *Xbp1,* are upregulated when PU.1 is deleted. Conversely, OBF1 serves to promote PC differentiation by activating BLIMP1[239,254]. BLIMP1 (encoded by the *Prdm1* gene) is a master regulator controlling GC B to PC differentiation[269,270]. In the absence of OBF1, BLIMP1 fails to be upregulated, thus impeding the final stages of antibody-secreting cell development[239].

Consistently, BLIMP1 was downregulated upon OBF1 knockdown in our RNA-seq data (**2.1** Fig. 7B). Moreover, OBF1 has also been shown to play critical roles in *IgH* expression in plasmacytoma cells[240]. This result indicates that the PC program is blocked when OBF1 is missing. Moreover, GSEA analysis showed that follicular B cell signatures are upregulated upon OBF1 depletion. Notably, enforced expression of OBF1 leads to greatly reduced numbers of follicular B cells[271]. In contrast, PU.1, along with Spi-B, controls follicular B cell homeostasis[253]. While the functional relevance of these observations will require further systematic analysis, the different regulatory roles of PU.1 and OBF1 on PC development and follicular B cell maintenance might shed light on the reverse correlation between the transcriptome of *Spi1*$^{-/-}$ and *Pou2af1*$^{-/-}$ B cells.

3.4 Low occurrences of octamer motif in OCT1/OCT2/OBF1 binding regions

We found that the octamer motif is strongly enriched in the binding regions of OBF1, OCT1 and OCT2 in *de novo* motif analysis by HOMER and MEME-ChIP. However, when octamer motifs were annotated at peak level, we surprisingly found that only a relatively small fraction of peaks contain one or more octamer motifs. Figure 14A shows that a low percentage of OBF1, OCT1 and OCT2 peaks contain discernible octamer motifs. Ranking the peaks based on ChIP enrichment signals, we found that octamer motifs are more frequently identified in peaks with top enrichment (Figure 14B).

One possible explanation for the low percentage of octamer motif positive peaks is that the octamer motif negative peaks may be generated by chromatin loops. A TF that binds to its cognate site can be recruited to a distal locus without a binding motif through a chromatin loop, whereby a peak of this TF can then be detected at this motif-free distal locus. In MPC11 cells, a plasmacytoma cell line, OCT1/OCT2 and OBF1 have been shown to mediate the long-range *Igh* promoter-enhancer loop (Figure 21). These factors directly bind to the 3' *Igh* enhancer and form indirect interaction to *Igh* promoter through the interaction with promoter-bound TFII-I complex[240]. Moreover, depletion of OBF1 leads to significantly weakened promoter-enhancer interaction, as well as a drastic reduction in IgG production. In another study, researchers showed that OCT2 and OBF1 can also mediate the interaction between the *Bcl2* promoter and 3' *Igh* enhancer in follicular lymphoma cells harboring t(14;18) translocations which positions *Bcl2* gene and Ig enhancer in proximity[258]. OCT2 was demonstrated to bind to the promoters of *Bcl2*

and *Igh* enhancers. Since no octamer motif can be identified in *Bcl2* promoters, it is likely that OCT2 binds to *Igh* enhancer, and then physically associates with the P1 and P2 promoters of *Bcl2*. Therefore, chromatin loops seem to be a contributing reason why most peaks of OCT factors and OBF1 contain no octamer motifs. The roles of OCT factors and OBF1 in mediating chromatin loops require more detailed research, as well as the abundance of chromatin loops associated with these factors.

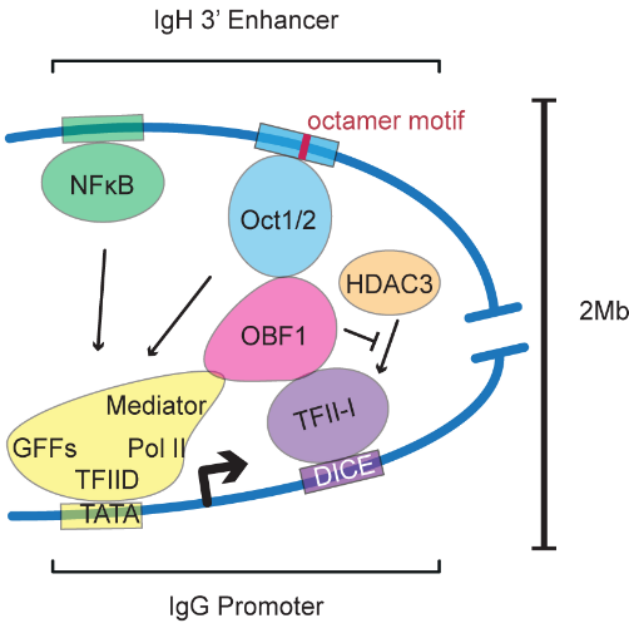


Figure 21. **A scheme showing the long-range chromatin loop between *Igh* promoter and 3' *Igh* enhancer mediated by OBF1.** First, OBF1 and HDAC3 compete with each other to gain interaction with TFII-I, which removes the repression of the target locus. Second, the established interaction between OBF1 and TFII-I facilitates the communication between *Igh* promoter and enhancer elements which places the two regulatory regions into juxtaposition. Eventually, with the recruitment of relevant factors, enhancer is able to exert the functional role on *Igh* promoter. This figure is adapted from[240].

OCT factors and OBF1 can also be recruited to binding sites by other chromatin binding factors regardless of a lack of octamer motifs. This might provide another insight for the low octamer positive peak fraction of the three OCT factors. In this way, the binding of TFs to the chromatin can be identified in those chromatin loci that contain no discernible motifs. Translocation of the *BCL2* gene (18q21) to *IGH* gene region (14q32) is the hallmark of follicular lymphoma, a

prevalent form of non-Hodgkin's lymphoma. This translocation places the *BCL2* gene under the transcriptional control of the *IGH* enhancer, which leads to enhanced survival of lymphoma cells. OCT2 and OBF1 are critical factors that bind to the alternative *BCL2* promoter and sustain the high-level expression of *BCL2* in t(14;18) lymphoma cells. As octamer motifs are missing in this alternative promoter, and their binding are largely dependent on C/EBP and Cdx-binding sites, OCT2 and OBF1 are believed to be recruited by C/EBPα, Cdx2, or A-Myb[272]. We found that ETS motifs are strongly enriched in the peaks regions of these factors. Several ETS factors, particularly PU.1 and ETS1, are highly colocalized with OCT factors and OBF1 at genomic scale. Therefore, it is highly possible that OCT factors can be recruited to chromatin directly or indirectly by ETS factors without the prerequisites for octamer motifs. The binding of a factor to the regulatory regions that contain no cognate motifs through the recruitment by alternative TFs seems to be a common phenomenon.

Low octamer motif occurrences can also be ascribed to the biases intrinsic to current program-based motif search, as low-affinity motifs or degenerated motifs might be overlooked during the in the *in silico* motif discovery process. For example, Kemler et al found that in addition to octamer motif, OCT1 and OCT2 also bind a heptamer element, 5'-CTCATGA-3', which showed reduced binding affinity for OCT factors compared to the classic octamer motif[214]. Another example is the dyad OCT-binding sequence discovered in the imprinted *H19*/*Igf2* locus between *Igf2* and the *H19* lncRNA[273]. In addition, octamer motifs containing point mutation exhibit lower binding affinities[255,274] for OCT factors, as well as reduced transcriptional activity[275], and therefore are also under-represented in motif analysis. Still, although many studies on the binding sites of OCT1 and OCT2 have been done, they mainly focus on the canonical octamer motifs, as well as the mutant derivatives. Therefore, it is reasonable to speculate that additional binding motifs with lower genomic abundance and affinities have not yet been appreciated. Lack of a complete understanding of their binding motifs seems also to be another factor that contributes to the underestimated frequency of octamer motif occurrence.

Pioneer factors are able to access their target DNA binding site located in compact silent chromatin regions. They are capable of independently recognize and bind to the partially exposed motifs. Soufi et al proposed that OCT4, as well as Sox2 and Klf4, can target nucleosomes through a partial or degenerate binding motif[276]. OCT4, OCT1 and OCT2 recognize similar binding motifs, and all of them use POU_S and POU_H subdomain for DNA binding.

Therefore, it is reasonable to presume that OCT1 and OCT2 could also bind to loci with partial or degenerate binding motifs, and these motifs might be largely overlooked by motif analysis. Besides degenerate OCT-binding motifs, OCT factors have been shown to bind to the composite binding motifs; mutations can be tolerated in the cases of compound motifs, and binding on sub-optimal binding sites could be stabilized by the presence of a second OCT factor, or OBF1, as shown in our ChIP-seq (**2.1** Fig. 4B, 4D and 4F). Since compound binding sites have not been observed in *de novo* motif analysis, the *in silico* motif analysis of the motif is limited. In addition, OCT factors also exhibit context-specific binding on the cognate sites. OCT1 binding to MORE motifs is induced by oxidative and genotoxic stress, which leads to the phosphorylation of the conserved Ser385 residue in OCT1 POU_H sub-domain[277]. Therefore, the non-canonical octamer binding motifs seem to account for a considerable number of binding events.

In sum, the low frequency of canonical octamer motif identified in the peaks of OCT1, OCT2 and OBF1 reflect the unique binding characteristics of OCT1 and OCT2. While further studies are needed to thoroughly characterize their genomic binding, they bind to chromatin, in addition to canonical octamer motifs, through the engagement in functional long-range chromatin loops, recruitment by chromatin-bound factors and interactions with non-canonical or derivatives of octamer motifs. The multi-dimensional complexity of their binding seems to be associated with low frequency of motif occurrence.

3.5 Functional roles of OCT1 in B cells

We generated transgenic mice using CRISPR/Cas9 technology by inserting AviTag-coding sequence endogenously to *Pou2f1*, *Pou2f2* and *Pou2af1* gene loci, and then performed Bio-ChIP-seq to map the genomic location of OCT1, OCT2 and OBF1 in LPS-stimulated splenic B cells. This allowed us to obtain the genomic binding data of these factors with much greater depth than possible with antibodies. Antibody-based ChIP-seq identified much fewer peaks; yet, all of the peaks identified with antibodies are also present in Bio-ChIP-seq results (**2.1** Fig. S1C and S1D). Motif analysis of the peak regions showed that octamer motifs are among the top enriched motifs, which indicates that the identified peaks are specific and authentic.

Previous ChIP-seq studies with antibodies examined the binding of OCT2 only [278] or OCT2 and OBF1[92] in B cells, but did not investigate the functional role and genomic binding of OCT1. Although studies showed that OCT1, OCT2 and OBF1 are interaction partners, and OBF1 binds to DNA indirectly via interaction with OCT1 and OCT2[193,210], a systematic comparison at genomic scale was missing. Here, we found extensive overlap between OCT1 and OCT2 peaks, and over 90% of OBF1 peaks overlapped with either OCT1 or OCT2 peaks. This is in accord with the notion that OBF1 binds to DNA through the recruitment by OCT factors. Motif analysis of the different subgroups of peaks (e.g. sites bound by OBF1 and OCT factors, or by OCT factor alone) did not reveal significant differences in terms of octamer-like motifs, and the known consensus motif 5'-ATGCAAAT-3' or variants thereof were ranked at the top of the enriched motif list.

OCT2 and OBF1 have been identified early on as critical factors for the GC reaction, although the mechanisms underlying their functions are unclear[199-202,232]. Until now, the role of OCT1 in GC reaction has not been characterized. Similar to OBF1 and OCT2, we showed here that *de novo* motif analysis of regions differentially bound by OCT1 under anti-CD40/IL4 showed that motifs for GC-related TFs were highly enriched. Importantly, all three factors bound to multiple genes encoding TFs important for the initiation phase of GCs (**2.1** Fig. 5C and S4D). In human B lymphoma cells, OCT1 controls the fast proliferation of Raji cells through modulating the expression level of OBF1, as the expression level of OBF1 is decreased when OCT1 is depleted in Raji cells. Although OCT1 function remains to be studied in conditional *Pou2f1*$^{-/-}$ mouse models, this factor seems to be critical for GC reaction, which is similar to OBF1 and OCT2.

In sum, OCT1, OCT2 and OBF1 extensively colocalize at genomic scale. While further studies are needed to establish the functional role of OCT1 in GC reaction and B lymphoma cells, it seems to be another critical factor controlling GC reaction and B lymphoma proliferation.

3.6 The functional hierarchy of OCT1, OCT2 and OBF1 in the context of GC

OCT2 and OBF1 are critical factors for GC reaction[199-202,232]. However, the detailed mechanisms have remained largely unknown. The requirement of OCT2 in the GC reaction is antigen-specific and somewhat controversial. *Pou2f2*$^{-/-}$ mice completely failed to develop GCs when immunized with 4-hydroxy-3-nitro-phenyl-acetyl-ovalbumin (NP-OVA)[199]. However, GCs formation is

normal in *Pou2f2*$^{-/-}$ mice when challenged with influenza virus[235]. In another study, NP-KLH immunization successfully induced GC formation in *Pou2f2*$^{-/-}$ mice. Nonetheless, when compared to the WT counterparts, these GC B cells exhibited several defects such as impaired cellular proliferation, dramatically reduced serum level of antigen-specific IgGs, and retarded PC development. Therefore, the functional role of OCT2 in controlling GC formation remains debatable, and the development GCs in *Pou2f2*$^{-/-}$ mice is likely to be antigen-specific. In contrast to OCT2, OBF1 is unequivocally essential for GC reaction in an antigen-independent manner[199-202,232]. Up to now, *Pou2af1*$^{-/-}$ mice exhibit a complete deficiency in GC development in response to all antigens tested, including NIP-OVA[200], DNP-KLH (my own data) and sheep red blood cells (SRBCs) (my own data). In yet another study, although GC formation has not been characterized, T-cell dependent response was almost absent when immunized with LCMV and VSV in *Pou2af1*$^{-/-}$ mice, as manifested by a reduced level of class-switched antibodies[232]. Therefore, while both OCT2 and OBF1 are critical for GC formation, the latter shows a stronger effect than OCT2.

In our study, we found that GC reaction-related genes were upregulated after anti-CD40/IL4 stimulation (**2.1** Fig. 5A) compared to LPS stimulation. OBF1 showed stronger binding enrichment to the promoters of multiple GC-associated genes (**2.1** Fig. 5B) specifically under anti-CD40/IL4 stimulation, such as *Bcl6*, *Myc*, *Foxo1*, *Mef2b* and *Spi1*. Motifs for GC-related TFs were highly enriched in the regions differentially bound by OBF1 under anti-CD40/IL4 stimulation (contrast: anti-CD40/IL4 versus LPS). We observed similar patterns for OCT1 and OCT2 (**2.1** Fig. S4B and S4C). GSEA analysis showed that the GC signature is lost in anti-CD40/IL4 stimulated *Pou2af1*$^{-/-}$ B cells. Moreover, knocking down OCT1, OCT2 or OBF1 in Raji cells leads to a substantially decreased cellular proliferation rate. However, this defect is largely dependent on OBF1, because OBF1 expression level is downregulated when OCT1 or OCT2 is depleted by specific shRNAs, and enforced expression of OBF1 completely restored the proliferation of OCT1 and OCT2-depleted Raji cells. Therefore, combining our findings with the observations from other groups, OBF1 is likely to be the true master regulator for the GC reaction. The defects observed in OCT1 or OCT2-depleted cells are due to the concomitant downregulation of OBF1.

3.7 OBF1 controls the balance between GC program and post-GC differentiation

The GC reaction is a process tightly controlled by a network of TFs and other enzymes. Here, we showed that, in the absence of OBF1, the GC signature is downregulated in GC-derived B lymphoma cells, and that genes associated with PC differentiation are upregulated. Therefore, OBF1 is critical for the integrity of the GC transcriptome.

BCL6 is a master regulator of the GC reaction[279], and it is critical for the maintenance of the GC program[280]. *Bcl6*$^{-/-}$ mice are unable to undergo GC reaction[257]. Likewise, GCs are missing in *Pou2af1*$^{-/-}$ mice when immunized with various antigens[199-202,232]. Moreover, the GC transcriptional program is severely perturbed in the absence of BCL6[279,280], which is also very similar to OBF1-depleted B cells. These highly similar phenotypes indicate these two factors might function in the same pathway. Indeed, BCL6 is downregulated when OBF1 is depleted in both anti-CD40/IL4 stimulated mouse B cells and GC-derived human B lymphoma cells. Genes repressed by BCL6 are upregulated (**2.1** Fig. S6A). These observations suggest that OBF1 likely regulates GC reaction through modulating *Bcl6*.

OBF1 maintains the GC program, as GC-related genes are downregulated upon the depletion of OBF1 (**2.1** Fig. 7B). These cells with reduced OBF1 start to adopt a PC signature, as genes involved in plasma differentiation are upregulated, such as *IRF4*, *SLAMF7*, *SCAMP5*, *SORBS3* and others. This indicates at least a partial initiation of the post-GC differentiation. Using CUT&RUN assays, we found that OBF1 binds to the regulatory regions of the majority of genes associated with GC reaction, such as *BCL6*, *BACH2*, *FOXO1*, *AICDA* and *PAX5*, in primary murine and human GC B cells, as well as human GC-derived B cell lymphoma cells (including Burkitt's lymphoma and DLBCL). These evidence strongly points to the fact that OBF1 maintains the entire transcriptome of GC through modulating GC-related genes.

OBF1 governs the transition between the GC program and post-GC differentiation. OBF1 and BCL6 are co-expressed in centroblasts at a high level, and both of these factors show reduced expression levels in centrocytes[281,282]. The expression of IRF4 is induced by CD40 signaling pathway in centrocytes that express high-affinity antibodies; this leads to repression of *Bcl6* expression and thus promotes the GC B cell to PC differentiation [92,282,283]. Notably, high level of IRF4 is essential to induce OBF1 and BCL6 at pre-GC stage, and subsequent repression of IRF4

by elevated BCL6 allows the initiation of GC reaction[284]. Together with our data, this indicates that OBF1 and BCL6 are expressed in a precisely synchronized manner, while OBF1 and IRF4 expression are mutually exclusive. More recently, OBF1 has been reported to regulate BCL6 in $CD4^+$ T_{FH} cells[238] and we showed here that it directly regulates the BCL6 promoter in primary human and mouse GC B cells. Taken into consideration of its transcription activation function on *BCL6* as well as other GC-related genes in B cells discovered by us, OBF1 is the key factor controlling the switch between GC reaction and further differentiation.

Depletion of OBF1 in GC-derived lymphoma cells leads to reduced BCL6 expression and impaired cellular proliferation. In these OBF1-depleted cells, GSEA analysis showed that GC signatures are lost, while PC differentiation signatures are upregulated. Consistently, IRF4 is strongly upregulated when OBF1 is depleted. Therefore, a seemingly contradictory issue emerges that OBF1 and IRF4 are co-expressed at high levels in PCs and both of them are closely engaged in PC differentiation and function, despite our findings that suggest a transcriptional repressive relationship between the two factors. Figure 22 summarizes the functional mechanism of OBF1 in controlling the balance between GC maintenance and GC exit.

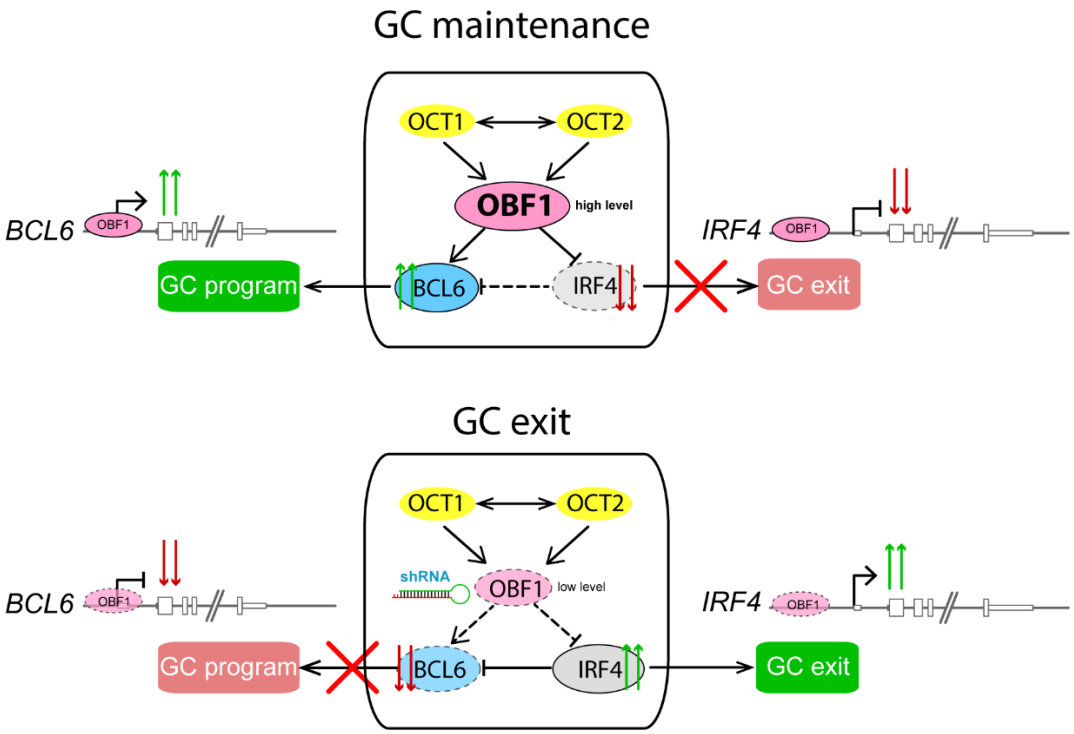


Figure 22 Scheme of the proposed molecular model. The expression of OBF1 is controlled by OCT1 and OCT2. High level of OBF1 activates the expression BCL6 while represses IRF4, thus maintaining the integrity of GC program and inhibiting GC exit. Upon the depletion of OBF1, BCL6 expression is silenced and IRF4 expression is activated, which terminates GC program and promotes GC exit process. (This graph is the graphic abstract of our accepted paper in **2.1**).

In fact, OBF1 and IRF4 are co-expressed in PCs, where both factors are critical for plasma differentiation [239,281,284,285]. In PCs, both IRF4 and BLIMP1 activate XBP1, which in turn drives the strong expression of OBF1[92,286]. Therefore, the coexistence of OBF1 and IRF4 in PCs indicates that OBF1 exerts repression on IRF4 expression specifically in the GC context: once the integrity of GC program has been disrupted and B cells subsequently have adopted a PC signature, OBF1 is no longer able to repress IRF4. Similarly, BCL6 is regulated by OBF1 under GC conditions, as BCL6 and IRF4 expression was found to remain largely unchanged in *Pou2af1*$^{-/-}$ splenic cells when stimulated with anti-CD40/IL4 *in vitro*[237] (and our data not shown).

In summary, OBF1 maintains the GC program by activating BCL6 and repressing IRF4 expression. Once OBF1 expression is lost, the BCL6 level is reduced and IRF4 is elevated, which in turn shuts down the GC program and permits post-GC differentiation. The functional role of OBF1 on the expression of BCL6 and IRF4 appears to be GC-restricted.

3.8 OBF1 as a therapeutic target for GC-derived B cell lymphoma

A CRISPR/Cas9 screen identified OBF1 as an essential gene for Burkitt's lymphoma cell lines[252], and it has been reported to play key roles in maintaining the proliferation of an array of GC-derived B lymphoma cells, including Burkitt's lymphoma and DLBCL cells [236,251]. Up until now, OCT1 and OCT2 are the only known two interaction partners of OBF1, however, the functional roles of OCT1 and OCT2, and the functional hierarchy of OCT1/OCT2/OBF1 in B cell lymphomas have remained unclear. Here, we showed that the downregulation of OCT1 and OCT2 also leads to the strongly impaired proliferation of B lymphoma cells (2.1 Fig. 6B-6C). Notably, OBF1 level is simultaneously reduced when OCT1 and OCT2 are depleted, and the proliferation of OCT1-/OCT2-depleted Raji cells can be fully rescued when OBF expression is

restored. Therefore, OBF1 is likely to be the critical factor for the proliferation of lymphoma cells.

Recent studies provide possible functional mechanisms underlying the importance of OBF1 for lymphoma proliferation. The OBF1 gene has been discovered to be controlled by a super-enhancer in DLBCL cells[251], and Burkitt's lymphoma and DLBCLs are highly dependent on the high expression level of OBF1[236,251,252]. JQ1 is a BET inhibitor with the highest specificity for BRD4[287], and it blocks the binding of the latter to acetylated chromatin, which in turn leads to reduced activity of enhancers or promoters[287]. The proliferation of DLBCL cells is impaired following JQ1 administration, as the activity of a super-enhancer controlling OBF1 expression is reduced, which in turn leads to reduced level of OBF1[251]; moreover, ectopic expression of MYC failed to rescue the anti-proliferative effects of JQ1 in DLBCL cells. Recently, functional importance has been ascribed to the OCT2-OBF1 interaction interface[236], as mutations that undermine the protein-protein interaction between these two proteins leads to reduced proliferation of B lymphoma cells. However, the specific downstream target(s) of OCT2 or OBF1 remained to be identified. Here, we found that BCL6 and IRF4, as well as genes involved in GC maintenance and post-GC differentiation, are all critical direct targets of OBF1 in Burkitt's lymphoma and DLBCL, as revealed by RNA-seq and bioinformatics analysis. Although BCL6 is a master regulator of GC maintenance and integrity, ectopic expression of BCL6 failed to rescue the proliferation of OBF1-depleted B lymphoma cells (2.1 Fig. S5G). This observation indicates that the identity of Raji cells has changed when OBF1 is depleted, which points to the possible disruption of the GC program and initiation of post-GC differentiation. Indeed, we found that IRF4, a negative regulator of Burkitt's lymphoma cell proliferation[288] and a master regulator of PC differentiation[285], was upregulated when OBF1 is abrogated in both Burkitt's lymphoma and DLBCL cell lines. In line with this, GSEA analysis confirmed upregulation of PC signature genes and perturbed BCL6 target genes. Strikingly, additional knockdown of IRF4 rescued the proliferation of OBF1-depleted GC-derived B lymphoma cells. Moreover, OBF1 also binds to the regulatory regions of the majority of GC-related genes, particularly the IRF4 and BCL6 loci, in primary murine or human GC B cells and GC-derived lymphoma cells. These observations further corroborate the functional role of OBF1 in GC maintenance, post-GC differentiation and proliferation of GC-derived B lymphoma cells. OBF1,

therefore, controls the proliferation of GC-derived B lymphoma cells by maintaining the GC transcriptional program and repressing post-GC differentiation.

OBF1 is associated with lymphomagenesis. Reciprocal chromosomal translocations that involve the IGH locus are a hallmark of GC-derived B cell lymphomas[289]. This is frequently found in DLBCL, Burkitt's lymphoma and plasmacytoma cells with most commonly t(3;14)(q27;q32), t(14;18)(q32;q21), and t(8;14)(q24;q32), which result in the deregulation of proto-oncogenes BCL6, BCL2 and c-MYC, respectively[290,291]. The juxtaposition of proto-oncogenes in the proximity of the IGH locus brings them under the powerful transcriptional control of the 3' IGH enhancer[289]. The transcriptional activity of the 3' IGH enhancer is itself dependent on OBF1; OBF1 mediates a long-range chromatin loop between IGH promoter and 3' enhancer which is essential for the high expression of class-switched antibodies in MPC11 cells, a plasmacytoma cell line[240]. OCT2 is able to bind to the alternative promoter region of BCL2 in t(14;18) translocation lymphoma, although no discernable octamer motifs could be identified in this region[272]. In addition, OCT2 and OBF1 have been found to mediate the interaction between the BCL2 promoter and the 3' IGH enhancer[258]. Since it governs the transcription of 3' IGH enhancer and translocated BCL2 gene, OBF1 presumably regulates any translocated gene which are juxtaposed to 3' IGH enhancer in B lymphoma cells. Indeed, we found that OBF1 maintains the fast proliferation of several GC-derived B lymphoma cells, including Burkitt's lymphoma and DLBCL, with various reciprocal chromatin translocation types. Depletion of OBF1 leads to a substantially impaired proliferation of Raji (c-MYC-IGH translocation[292]), SUDHL4 (BCL2-IGH translocation)[293], Daudi (c-MYC-IGH translocation[294]) and Ramos (c-MYC-IGH translocation[295]), which lend strong support to the notion that OBF1 exhibit a general transcriptional control of genes translocated in the proximity of the IGH locus.

Notably, although it is responsible for the expression of proto-oncogenes translocated to the IGH locus, OBF1 also binds to and controls the germline proto-oncogenes in GC-derived B lymphoma cells. For example, Daudi, Ramos and Raji cells harbor only germline BCL6[296,297], and yet OBF1 binds to and activates the transcription of these two genes. In the absence of OBF1, the BCL6 level was downregulated. This observation suggests that OBF1 controls the proliferation of GC-derived B cell lymphomas also using mechanisms beyond its functions on the 3' IGH enhancers.

Surprisingly, GSEA analysis revealed that B lymphoma cells with depleted OBF1 adopt signature genes that are linked to favorable prognosis, indicating that OBF1 could be used as an indicator for clinic lymphoma severity classification. Consistent with its pivotal roles in B lymphoma cells, OBF1 is highly expressed in lymphoma[281] and widely expressed in B cells. Moreover, although OBF1 is expressed throughout B cell development, it is more important for the function of late B cells[271]. Therefore, OBF1 may be a novel therapeutic target for treating patients diagnosed with B lymphomas.

In summary, OBF1 is an essential regulator for GC-derived B lymphoma cells and closely related to lymphomagenesis. OBF1 might serve as an indicator for lymphoma diagnosis, lymphoma severity classification and prognosis. Finally, it may be a novel therapeutic target for the clinic treatment of B lymphoma with presumably minor side effects.

3.9 OBF1 licenses the GC-B-to-PC/Bmem differentiation

GC to plasma or Bmems differentiation is a complex and tightly regulated process. Transcriptional regulation underlying this differentiation is under extensive study, and the detailed mechanisms governing this process are yet to be elucidated. The GC program has to be shut down before the initiation of downstream differentiation[92], this is termed “GC exit”.

IRF4 is a master regulator that promotes the GC-to-PC differentiation[285]. Although IRF4 is largely silent or weakly expressed during the GC reaction, it is highly expressed during normal PC differentiation and function[284]. Ectopic expression of IRF4 leads to the differentiation of PCs[256]. We showed that high expression level of OBF1 represses IRF4 expression. *Prdm1* is a target of OBF1[239]. Consistently, using RNA-seq, we found that the expression level of *Prdm1* is significantly reduced in the absence of OBF1.

Moreover, genes involved in vesicle transport and ER-to-Golgi transition – which are important for the function of secretory plasma cells- are upregulated in Raji cells with depleted OBF1, indicating that these cells are undergoing the differentiation necessary for setting up the secretion machinery. Therefore, this observation suggests that plasma differentiation has been initiated when OBF1 is reduced.

OBF1 has been reported to control the Bmem development in GC reaction[242], which is not supported by our data. Levels et al found that an elevated level of OBF1 activates ABF1 and

represses BCL6, unlike what we have seen in our studies, as I found that OBF1 activates BCL6, at least in GC-derived B lymphoma cells. Moreover, genes upregulated in Bmems are upregulated when OBF1 is decreased in GC-derived lymphoma cells. ZBTB32 is highly expressed in memory B cells[298], *ZBTB32* is upregulated in OBF1-depleted Raji cells, and its promoter is bound by OBF1 in both human and mouse GC B cells, suggesting a repressive role of OBF1 in regulating this gene.

In sum, OBF1 controls the post-GC differentiation by regulating multiple marker genes in PCs and Bmems.

3.10 Targeting OBF1

OBF1 emerges as an important target in GC-derived B lymphomas. Therefore, finding a way to specifically perturb the level of OBF1 is of a great imperative for clinical practice.

Small interfering RNAs (siRNAs) are a promising category of RNA-based oligonucleotide drugs, and great efforts have been made in the development of siRNA molecules for clinical use. The main problems that limit the clinical use of siRNAs are the low efficiency of delivery and the extensive degradation of siRNA molecules when administered. Although strategies have been developed to overcome these problems, they bring new issues by themselves. Incorporation of 2'F modification shields the modified siRNAs from nucleases which elongates the half-life of siRNA in the body, however, it also increases toxicity. Therefore, small chemical compounds are a better choice for targeting OBF1.

Traditionally, it is considered to be difficult for small molecules to directly regulate TF function by interfering with their DNA binding abilities [299]. DNA binding domains (DBDs) of TFs have been generally considered "undruggable", as these domains are flat and similar surface areas [300,301], which tend to preclude the proper and specific binding of inhibitors. Therefore, this issue renders DBDs unattractive candidates for drug discovery. In contrast, successful studies showed that small molecules are effective in perturbing protein-protein interfaces (PPIs) with the engagement of cofactors, dimerization of TFs, and complex forming subunits. This strategy should be effective in targeting OBF1 by blocking the interaction interface between OBF1 and OCT1/2, the only known interaction partners of OBF1. One study suggested that interfering with

the binding between OBF1 and OCT2 could reduce the proliferation of GC-derived lymphoma cells [236]. In our study, I found that decreasing the protein level of OBF1 is critical and effective to compromise the proliferation of lymphoma cells, as well as the integrity of GC program. Therefore, a targeting strategy that reduces OBF1 protein level is more straightforward comparing to targeting the interaction interface.

Recently, a new technology, which specifically directs the proteolysis of the targeted protein by proteolysis-targeting chimeras (PROTACs)[302,303], shed light on specifically targeting OBF1 at protein level. PROTAC technology leads to target protein degradation by taking advantage of the ubiquitin-proteasome system (UPS), the main intracellular machinery that is responsible for the degradation of proteins that are harmful or no longer needed. A PROTAC agent is a fusion molecule that consists of a small-molecule compound of the target protein, a linker carbon chain and a ligand of an E3 ubiquitin ligase. PROTAC technology has been tested and proven to interfere with a broad range of proteins in various subcellular compartments, including the cytosol, the nucleus and so forth[304-306]. Therefore, PROTAC technology serves as an ideal system for targeting OBF1 at protein levels.

However, the drawback of PROTAC technology in targeting OBF1 lies in identifying the chemical compounds (recruiting element) that specifically recognize OBF1, and the efficiencies of these compounds that translate the specific recognition into specific protein degradation. Therefore, the screening system has to be carefully developed with consideration to the particular cell lines to use, the reporter systems with clear readouts to monitor the specific degradation of OBF1 and exhibit the resulting desired phenotypes. Although it is difficult to design and perform the screening of chemical compounds that specifically target and degrade OBF1, it is of high value and imperative to obtain these compounds so that patients with dangerous or recalcitrant GC-derived B cell lymphomas can be greatly relieved or even cured.

3.11 OBF1-dependent immune evasion of B cell lymphomas

Immune evasion is an important strategy employed by tumor cells, including GC-derived B cell lymphomas. This is achieved by downregulation of proteins involved in antigen presentation (MHCI and MHCII, CIITA), TCR co-stimulation (CD80, CD86), cell-cell adhesion (ICAM1) and

inactivation of apoptotic pathways (FAS, TNF receptor family, BCL2/BCL2l1), thereby preventing the recognition and attack of tumor cells by the immune system.

The levels of MHCI and MHCII complexes are lost in B lymphoma cells, which is associated with reduced T cell infiltration and adverse prognosis[307-311]. CIITA is a master regulator of the MHCII complex, which is also downregulated in B lymphoma cells. Therefore, mutations and translocations of *CIITA* gene lead to the silencing of MHCII genes[312]. In addition, the MHCII genes and *CIITA* genes are inhibited by BCL6, which is a potent therapeutic target in DLBCL[313]. Restored expression of MHCII complex results in enhanced apoptosis of DLBCL cell by T cells[314]. Here, in this thesis, we found that loss of OBF1 in Raji cells leads to the downregulation of *BCL6*, and upregulation of *CIITA* and subunits of the MHCI/II complexes.

The *CD58* gene belongs to the Ig superfamily and is important in the adhesion and activation of $CD8^+$ T cells and natural killer (NK) cells[307]. Loss of MHCI complex leads to the activation of NK cells in the presence of cellular surface CD58, therefore B lymphoma cells with low levels of both MHCI and CD58 are exempted from NK recognition and attack[307]. Interestingly, we found that *CD58* and MHCI genes were significantly upregulated upon the OBF1 depletion, thereby reduced OBF1 leaves B lymphoma cells vulnerable in front of NK cells.

CD86 belongs to the B7 co-stimulatory family, whose expression is reduced in most DLBCL cells[315]. Importantly, loss of CD86 associates with reduced tumor-infiltrating lymphocytes, which facilitates immune evasion of B lymphoma cells[316]. We found that *CD86* was upregulated in OBF1-depleted Raji cells, therefore the presence of OBF1 is associated with immune evasion of B lymphoma cells.

Moreover, *ICAM1*, involved in tumor cell adhesion and immune synapse formation[317], and *CD95*, *TNFRSF14*, *BCL2* and *BCL2L1*, which are involved in apoptotic pathway[318-320], were upregulated in Raji cells when OBF1 is depleted.

These findings showed that OBF1 is involved in multiple aspects of immunoevasive maneuvers, which is in line with the prognosis favorable phenotype observed upon OBF1 depletion.

Materials and methods

Chapter 4 Materials and Methods

Mouse strains and cell lines

All strains were maintained on the C57BL/6J background. Animal experiments were carried out according to valid project licences, approved and regularly controlled by the Swiss Cantonal Veterinary Office of Basel-Stadt. *Pou2af1*$^{-/-}$ mice have been described[200]. *Rosa26*$^{BirA/BirA}$ mouse line[321] was obtained from M. Busslinger (IMP, Vienna). *Spi1*$^{-/-}$ mouse line was provided by S. Nutt (WEHI, Melbourne). Transgenic mouse strains harbouring endogenously AviTagged *Pou2f1*, *Pou2f2* and *Pou2af1* were generated with CRISPR/Cas9 technology, as discussed below.

293T cells were cultured in DMEM medium, 10% FBS (Sigma), at 37℃ and 5% CO_2. Ramos and Daudi were received from S. Junker (Aarhus University), and HT and SUDHL4 were purchased from DSMZ. B Lymphoma cell lines were maintained in RPMI1640 medium (10% FBS, 50 μM 2-mercaptoethanol) and passaged every two to three days, at 37℃ and 5% CO_2. Splenic B cells were isolated using MACS and stimulated with Lipopolysaccharide (LPS) or anti-CD40/Interleukin-4 (IL4).

Generation of transgenic mice by CRISPR/Cas9

Gene-specific sgRNAs were designed using the online tool (https://zlab.bio/guide-design-resources). Top and bottom strands of sgRNAs were annealed and ligated into the BbsI site of pX458 plasmid. T7 promoter sequence was introduced by PCR, and sgRNAs were transcribed and purified. The resulting sgRNAs were stored immediately in -80 ℃ for zygote microinjection. The sequences of each guide RNA are listed:

Pou2af1: 5'-CCTCTCCGTGGAGGGCTTTT-3'

Pou2f1: 5'-CTCAGCTCACTGTGCCTTGG-3'

Pou2f2: 5'-GCCCTGGCTGGCGGGCAGCA-3'

Gene specific single-stranded oligo DNA nucleotides (ssODNs) were used as repair templates for inserting the coding sequence of AviTag-FLAG specifically to *Pou2f1*, *Pou2f2* and *Pou2af1* loci. Each ssODN contains coding sequence for AviTag and FLAG located between the

upstream and downstream 40nt homologous arms. They were purchased from Integrated DNA Technologies, and reconstituted with ultrapure water at a final concentration of 1 μg/μL. Sequences of ssODNs are listed below:

Pou2f1:

5'-CACAGTCGCCTCTGCCAGTGGGCCTGCTTCCACCACCACAGCTGCAT

CCAAGGCACAGGGCGGCCTGAACGACATCTTCGAGGCTCAGAAAAT

CGAATGGCACGAAGGTGACTACAAAGACGATGACGATAAATGAGCT

GAGTGCAGAGCTGGGTTGCCACAGGCTTTCCTCACTACA-3'

Pou2f2:

5'-GGCGGGGAGATGACACAGTTGTTCCCCAGCCCTGGCTGGCGGGCAG

CATGATGGGCCTGAACGACATCTTCGAGGCTCAGAAAATCGAATGG

CACGAAGTTCATTCCAGCATGGGGGCTCCAGGTAAGAGGCTGATGC

CTTCT-3'

Pou2af1:

5'-AACACGTACGAGCTCAACCACACCCTCTCCGTGGAGGGCTTTGGCG

GCCTGAACGACATCTTCGAGGCTCAGAAAATCGAATGGCACGAAGG

TGACTACAAAGACGATGACGATAAATGAGGCTGGCTTGCATCTAAC

AGATGTTTCACCCATAGCTGAG-3'

Microinjection was performed following the described protocol[322]with some modifications. For each zygote, a solution containing 25 ng/μL sgRNA (in-house), 25 ng/μL Cas9 protein (Toolgen), 50 ng/μL Cas9 mRNA (Sigma) and 7 ng/μL ssODN was micro-injected into the pronucleus. in total roughly 200 zygotes were injected for each gene. Zygotes post injection were immediately transplanted to the pseudopregnant foster mothers. In 3 weeks, gene-edited pups were delivered.

To genotype the resulting transgenic pups, genomic DNA was extracted from CRISPR/Cas9 treated pups using DNA extraction buffer (10% Chelex 100 in ultrapure water). The targeted

regions of the genome were amplified by PCR. To determine the correct integration of AviTag coding sequence, PCR amplicons were TA-cloned (A1360, Promega), and sent for Sanger sequencing. Primers used in genotyping for genome-edited mice are listed below.

Genotyping polymerase chain reaction (PCR)

To determine the genotype of mouse strains used in this thesis, genomic DNA was isolated from ear clipping biopsies and used as a template for PCR.

The PCR assay for WT and *Pou2af1*$^{-/-}$ loci were carried out with the following primer combinations, respectively:

for: 5'-GCT CCC TGA CCA TTG AC-3',
rev: 5'-TCC TGT CCC ATC CCC CTG TAA-3'

and

for: 5'-GCG TGC AAT CCA TCT TGT TCA ATG G-3',
rev: 5'-GGC TTA GAT AAC AAA GCG TGT GTC C-3'

The size of the PCR products for the WT allele were 600 bp and for the mutant allele 900 bp.

AviTagged OCT1, OCT2 and OBF1 mice was determined using the following primer combinations, respectively:

for: 5'-CTG TTC TGG GCT CTT TGT CT-3'
rev: 5'-AGG AAG CCA ATG ACA GGT GT-3'

and

for: 5'-ACA CTC ATG ATC TGG GGT CG-3'
rev: 5'-ACA CAC ACC CCA TGC CTA AT-3'

and

for: 5'-TCC CTG ACC ATT GAC AAG CT-3'
rev: 5'-TTG CTT CCC CAA TGA TGT GC-3'

The size of the PCR products for the WT/OCT1 were 433/508 bp, WT/OCT2 were 427/475 bp, and WT/OBF1 were 409/484 bp, respectively .

WT and transgenic *Rosa26-BirA* loci were determined using the following primer combinations, respectively:

for: 5'-GTG TAA CTG TGG ACA GAG GAG-3'
rev: 5'-GGA GAG CTT AAA TCG GGC GA-3'

and

for: 5'-GGA GAG CTT AAA TCG GGC GA-3'
rev: 5'-CCC TGT TTG TCT ATT CCG CG-3'

The size of the PCR products for the WT allele were 367 bp and for the mutant allele 593 bp.

RNA isolation and Quantitative Real-Time PCR (RT-qPCR)

RNA was extracted using RNeasy Micro kit (QIAGEN), complementary DNA was generated using ImProm-II Reverse Transcription System (Promega), and quantitative PCR performed using FastStart Universal SYBR Green Master (Roche); human *β-ACTIN* was used as a reference gene. The RT-qPCR signal was measured with StepOne Real-Time PCR System (Applied Biosystems). The following primers were used in RT-qPCR:

POU2AF1:

for: 5'-CAC CAA TGT CAC GAC AAG AAG C-3',
rev: 5'-ACG GGA AAT AGG TGA GGG GT-3'

POU2F1:

for: 5'-ATG AAC AAT CCG TCA GAA ACC AG-3',
rev: 5'-GAT GGA GAT GTC CAA GGA AAG C-3'

POU2F2:

for: 5'-GAG GAG CCC AGT GAT CTG GA-3',
rev: 5'-GAA GCG GGA AAT GGT CGT C-3'

BCL6:

for: 5'-CGC AAC TCT GAA GAG CCA CCT GCG-3',
rev: 5'-TTT GTG ACG GAA ATG CAG GTT A-3'

IRF4:

for: 5'-ACA GCA GTT CTT GTC AGA G-3',
rev: 5'-GAG GTT CTA CGT GAG CTG-3'

β-ACTIN:

for: 5'-AGC CTC GCC TTT GCC GAT C-3',
rev: 5'-AGC GGG CGA TAT CAT CAT CC-3'

Splenic B cell stimulation

Splenic B cells were isolated from the spleens using MACS, following the procedure of the cell separation protocol from Miltenyi Biotec. Splenic single cell suspension was prepared and washed in fresh MACS buffer. B cells were stained with anti-CD19 microbeads for 15min at 4 °C. $CD19^+$ mature B cells were isolated using LS column (Cat# 130-042-401, Miltenyi Biotec). Purified $CD19^+$ mature B cells were subjected to *in vitro* culture with desired stimulants for experiments.

For T-dependent stimulation, B cells were cultured under anti-CD40/IL4 in RPMI1640 medium (10% FBS, 50 μm 2-mercaptoethanol, 1 ug/mL anti-murine-CD40, 20 ng/mL IL4, 2 mM Glutamine, 1% ITS, 1% MEM, 1 mM Sodium Pyruvate). For T-independent stimulation, B cells were cultured under LPS in RPMI1640 medium (10% FBS, 50 μm 2-mercaptoethanol, 50 μg/mL LPS, 2 mM Glutamine, 1% ITS, 1% MEM, 1mM Sodium Pyruvate). Culture medium of the corresponding stimulation was changed every day.

For each stimulation, B cells were collected after 72 hours. The stimulated B cells were crosslinked with 1% formaldehyde in PBS for 10 min at room temperature and then stored in -80 °C for Bio-ChIP-seq.

Purification of primary murine and human GC B cells

Murine GCs were induced by sheep red blood cells (SRBCs) immunization as described in[323]. 100μL 10% SRBC PBS suspension was injected intraperitoneally into each mouse (8-12 weeks old, *Pou2af1*^{AviF} mice). Seven days after injection, GC B cells were harvested for Bio-ChIP-seq or FACS analysis.

To isolate murine GC B cells for Bio-ChIP-seq, seven days post SRBC immunization, mice were sacrificed, and GC B cells were purified from spleen following procedures from PNA MicroBead Kit (Germinal Center B Cell (PNA) MicrsoBead Kit, 130-110-479, Miltenyi Biotec). Purified GC B cells were crosslinked with 1% formaldehyde in PBS for 10 min at room temperature and then stored in -80 ℃ for Bio-ChIP-seq.

Human GC B cells were isolated from human tonsillectomy specimens with chronic tonsillitis. Tonsils were collected and anonymized during standard surgical pathology work-up. All patients had given informed consent, in accordance with the Swiss Federal Act on Research involving Human Beings, art. 38 and in accordance with the declaration of Helsinki. GC B cells were purified by fluorescence-activated cell sorting (FACS) based on surface expression of CD19 (CD19-PE-cy7, 557835, BD Biosciences), CD38 (CD38-PE, 555460, BD Biosciences) and IgD (IgD-BV605, 563313, BD Biosciences). Purified GC B cells were crosslinked with 1% formaldehyde in PBS for 10 min at room temperature and then stored in -80 ℃ for CUT&RUN.

Protein extraction and western blot

B cells were lysed by RIPA buffer (10 mM Tris-pH7.4, 100 mM NaCl, 1 mM EDTA, 1 mM EGTA, 1% Triton X-100, 0.1% SDS, 0.5% sodium deoxycholate, 'complete' protease inhibitor cocktail (Roche)). Protein concentrations were determined with BCA protein assay kit (23227, ThermoFisher Scientific). Antibodies used in this thesis: OBF1 (33483, Cell Signaling Technology), OCT2 (ab179808, Abcam), OCT1 (8157, Cell Signaling Technology), MCM7 (ab2360, Abcam) and H3 (ab1791, Abcam).

Lentivirus production and gene knockdown by shRNAs

B lymphoma cells with lentiviruses were infected for 5 days at 37℃ and 5% CO2. Post infection, stable integrations were selected for 7 days in the presence of 3 μg/mL Puromycin (GIBCO) or Blasticidin (GIBCO). Non-target shRNA (Cat# SHC002, Sigma) was used as a control in shRNA-mediated gene knockdown experiments. After selection, stably integrated B lymphoma cells were used for downstream assays. For the proliferation measurements, cells treated with shRNAs were quantified by flow cytometry using Vi-CELL XR Cell Viability Analyzer (Beckman Coulter) at the indicated time points.

To prepare lentivirus, 7 μg backbone plasmid, as well as helper plasmids expressing Rev, Vsv-g, Tat and Gag/Pol, were transfected to 1.3×10^6 293T cells that were seeded 24 hours before. Supernatant containing lentivirus was collected at 24, 48 and 72 hours post the transfection. Supernatants of different time point were collected, centrifuged at 1500g for 45 min, and filtered through 0.4 μm membrane. Lentiviruses were precipitated by adding 5 mL LentiX concentrator to 15 mL filtered lentivirus-containing supernatant, and the solution was incubated for 24 hours at 4 ℃ followed by centrifugation at 1500g for 45 min. Lentivirus pellet was then resuspend in 100~150 μL growth medium, which was ready for infection.

To infect B lymphoma cells, 10 μL lentivirus suspension was added to each well of 6-well plate with 2×10^5 cells cultured in RPMI1640 growth medium containing 8 μg/mL polybrene for 5 days. For shRNA mediated gene knockdown experiments, stable integrations were selected in the presence of 3 μg/mL puromycin or blasticidin. Non-mammalian shRNA (Cat# SHC002, Sigma) was used as a control in knockdown experiments. Stably integrated B lymphoma cells were used for downstream assays. For the proliferation measurements, cells treated with shRNAs were stained with trypan blue and quantified using Vi-CELL XR Cell Viability Analyzer (Beckman Coulter) at the indicated time points.

Chromatin immunoprecipitation (ChIP)

B cell chromatin immunoprecipitation was performed as described in the manuscript Song *et al* and[324]. The following antibodies were used: H3K4me3 (07-473, Millipore), H3K4me1 (ab8895, Abcam), H3K27ac (39133, Active Motif) and PU.1 (2258, Cell Signaling Technology).

Bio-ChIP-seq

B cell chromatin immunoprecipitation was performed as described in the manuscript Song *et al* and[325] with some modifications. $CD19^{+}$ splenic B cells from *Pou2f1*^{AviF}, *Pou2af1*^{AviF} and *Pou2f2*^{Avi} were stimulated with either LPS or anti-CD40/IL4 stimulation, and primary GC B cells were subjected to Bio-ChIP-seq.

Gene Set Enrichment Analysis (GSEA)

GSEA was carried out as described in the manuscript Song *et al*. TMM normalization was applied to raw count table of RNA-seq data. In the case of murine data, mouse gene symbols were converted to human counterparts if existed. Gene sets used in the GSEA were Molecular Signatures Database v7.2 (c7: immunologic signature gene sets) and lymphoma signature database (https://lymphochip.nih.gov/signaturedb/) from the L. Staudt laboratory[326]. FDR less than 0.25 was used as a cutoff for significant enrichment.

CUT&RUN

CUT&RUN experiments were performed following the protocol as described [327,328], as well as on protocol.io (https://www.protocols.io/view/cut-amp-run-targeted-in-situ-genome-wide-profiling-zcpf2vn?step=68), with some modifications for crosslinked cells. Antibodies used in CUT&RUN were: OBF1 (33483, Cell Signaling Technology), H3K27ac (39133, Active Motif) and IgG (I5006-10MG, Sigma).

Sequencing library preparation procedures for CUT&RUN DNA using NEBNext Ultra II DNA Library Prep Kit aimed at preserving small DNA fragments (30-80 bp) was performed as described[329]. Stepwise protocol was elaborated at protocol.io (https://doi.org/10.17504/protocols.io.wvgfe3w). Briefly, after CUT&RUN, the purified DNA fragments were incubated at 50 °C for 1 h so that short DNA fragments were retained. Seven times diluted NEB adapter was used in order to minimize adapter dimer formation. PCR amplification was performed for at least 8 cycles, and the resulting libraries were purified by two step 1.1× (0.8× + 0.3×) AMPure beads incubation. In case of the resistance presence of dimers, additional 1.1× AMPure beads purification might need to be performed. Libraries were pooled

and subjected for sequencing following NextSeq High Output (75 cycles) (2 × 38 bp, 6-bp index) protocol. In this thesis, B lymphoma cells (Raji and SUDHL4) and primary human tonsillar GC B cells were subjected to CUT&RUN to map the genomic binding regions of OBF1.

ChIP-seq data processing

Detail ChIP-seq data processing procedures and parameters were specified in manuscript Song *et al*.

RNA-seq data processing

Detail RNA-seq data processing procedures and parameters were specified in manuscript Song *et al*.

CUT&RUN data processing

Detail CUT&RUN data processing procedures and parameters were specified in manuscript Song *et al*. CUT&RUN signals from different samples were normalized using *E. coli* carry-over DNA as was introduced with pA/G-MNase[330].

Data availability

ChIP-seq and RNA-seq data in this thesis are available at GEO under accession number GSE142040.

References

References

1. Tsuneto M, Tokoyoda K, Kajikhina E, et al. B-cell progenitors and precursors change their microenvironment in fetal liver during early development. *Stem Cells*. 2013;31(12):2800-2812.

2. Ohmura K, Kawamoto H, Lu M, et al. Immature multipotent hemopoietic progenitors lacking long-term bone marrow-reconstituting activity in the aorta-gonad-mesonephros region of murine day 10 fetuses. *J Immunol*. 2001;166(5):3290-3296.

3. Cumano A, Godin I. Ontogeny of the hematopoietic system. *Annu Rev Immunol*. 2007;25:745-785.

4. Yokomizo T, Ng CE, Osato M, Dzierzak E. Three-dimensional imaging of whole midgestation murine embryos shows an intravascular localization for all hematopoietic clusters. *Blood*. 2011;117(23):6132-6134.

5. Melchers F. Checkpoints that control B cell development. *J Clin Invest*. 2015;125(6):2203-2210.

6. Sitnicka E, Brakebusch C, Martensson IL, et al. Complementary signaling through flt3 and interleukin-7 receptor alpha is indispensable for fetal and adult B cell genesis. *J Exp Med*. 2003;198(10):1495-1506.

7. Decker T, Pasca di Magliano M, McManus S, et al. Stepwise activation of enhancer and promoter regions of the B cell commitment gene Pax5 in early lymphopoiesis. *Immunity*. 2009;30(4):508-520.

8. Revilla IDR, Bilic I, Vilagos B, et al. The B-cell identity factor Pax5 regulates distinct transcriptional programmes in early and late B lymphopoiesis. *EMBO J*. 2012;31(14):3130-3146.

9. Shimizu T, Mundt C, Licence S, Melchers F, Martensson IL. VpreB1/VpreB2/lambda 5 triple-deficient mice show impaired B cell development but functional allelic exclusion of the IgH locus. *J Immunol*. 2002;168(12):6286-6293.

10. Schuh W, Meister S, Roth E, Jack HM. Cutting edge: signaling and cell surface expression of a mu H chain in the absence of lambda 5: a paradigm revisited. *J Immunol*. 2003;171(7):3343-3347.

11. Galler GR, Mundt C, Parker M, Pelanda R, Martensson IL, Winkler TH. Surface mu heavy chain signals down-regulation of the V(D)J-recombinase machinery in the absence of surrogate light chain components. *J Exp Med*. 2004;199(11):1523-1532.

12. Grawunder U, Leu TM, Schatz DG, et al. Down-regulation of RAG1 and RAG2 gene expression in preB cells after functional immunoglobulin heavy chain rearrangement. *Immunity*. 1995;3(5):601-608.

13. Lu R, Neff NF, Quake SR, Weissman IL. Tracking single hematopoietic stem cells in vivo using high-throughput sequencing in conjunction with viral genetic barcoding. *Nat Biotechnol*. 2011;29(10):928-933.

14. Peschon JJ, Morrissey PJ, Grabstein KH, et al. Early lymphocyte expansion is severely impaired in interleukin 7 receptor-deficient mice. *J Exp Med*. 1994;180(5):1955-1960.

15. von Freeden-Jeffry U, Vieira P, Lucian LA, McNeil T, Burdach SE, Murray R. Lymphopenia in interleukin (IL)-7 gene-deleted mice identifies IL-7 as a nonredundant cytokine. *J Exp Med*. 1995;181(4):1519-1526.

16. Rolink A, Kudo A, Karasuyama H, Kikuchi Y, Melchers F. Long-term proliferating early pre B cell lines and clones with the potential to develop to surface Ig-positive, mitogen reactive B cells in vitro and in vivo. *EMBO J*. 1991;10(2):327-336.

17. Alt FW, Yancopoulos GD, Blackwell TK, et al. Ordered rearrangement of immunoglobulin heavy chain variable region segments. *EMBO J*. 1984;3(6):1209-1219.

18. Jung D, Giallourakis C, Mostoslavsky R, Alt FW. Mechanism and control of V(D)J recombination at the immunoglobulin heavy chain locus. *Annu Rev Immunol*. 2006;24:541-570.

19. Rolink A, Ghia P, Grawunder U, et al. In-vitro analyses of mechanisms of B-cell development. *Semin Immunol*. 1995;7(3):155-167.

20. Llorian M, Stamataki Z, Hill S, Turner M, Martensson IL. The PI3K p110delta is required for down-regulation of RAG expression in immature B cells. *J Immunol*. 2007;178(4):1981-1985.

21. Geier JK, Schlissel MS. Pre-BCR signals and the control of Ig gene rearrangements. *Semin Immunol*. 2006;18(1):31-39.

22. Melchers F. The pre-B-cell receptor: selector of fitting immunoglobulin heavy chains for the B-cell repertoire. *Nat Rev Immunol*. 2005;5(7):578-584.

23. Parker MJ, Licence S, Erlandsson L, et al. The pre-B-cell receptor induces silencing of VpreB and lambda5 transcription. *EMBO J*. 2005;24(22):3895-3905.

24. Karnowski A, Cao C, Matthias G, et al. Silencing and nuclear repositioning of the lambda5 gene locus at the pre-B cell stage requires Aiolos and OBF-1. *PLoS One*. 2008;3(10):e3568.

25. Flemming A, Brummer T, Reth M, Jumaa H. The adaptor protein SLP-65 acts as a tumor suppressor that limits pre-B cell expansion. *Nat Immunol*. 2003;4(1):38-43.

26. Jumaa H, Bossaller L, Portugal K, et al. Deficiency of the adaptor SLP-65 in pre-B-cell acute lymphoblastic leukaemia. *Nature*. 2003;423(6938):452-456.

27. Maki R, Kearney J, Paige C, Tonegawa S. Immunoglobulin gene rearrangement in immature B cells. *Science*. 1980;209(4463):1366-1369.

28. Perry RP, Kelley DE, Coleclough C, Kearney JF. Organization and expression of immunoglobulin genes in fetal liver hybridomas. *Proc Natl Acad Sci U S A*. 1981;78(1):247-251.

29. Melchers F, Yamagami T, Rolink A, Andersson J. Rules for the rearrangement events at the L chain gene loci of the mouse. *Adv Exp Med Biol*. 2007;596:63-70.

30. Yamagami T, ten Boekel E, Andersson J, Rolink A, Melchers F. Frequencies of multiple IgL chain gene rearrangements in single normal or kappaL chain-deficient B lineage cells. *Immunity*. 1999;11(3):317-327.

31. Ehlich A, Schaal S, Gu H, Kitamura D, Muller W, Rajewsky K. Immunoglobulin heavy and light chain genes rearrange independently at early stages of B cell development. *Cell*. 1993;72(5):695-704.

32. Grawunder U, Haasner D, Melchers F, Rolink A. Rearrangement and expression of kappa light chain genes can occur without mu heavy chain expression during differentiation of pre-B cells. *Int Immunol*. 1993;5(12):1609-1618.

33. Li YS, Hayakawa K, Hardy RR. The regulated expression of B lineage associated genes during B cell differentiation in bone marrow and fetal liver. *J Exp Med*. 1993;178(3):951-960.

34. ten Boekel E, Melchers F, Rolink A. The status of Ig loci rearrangements in single cells from different stages of B cell development. *Int Immunol*. 1995;7(6):1013-1019.

35. Pelanda R, Schaal S, Torres RM, Rajewsky K. A prematurely expressed Ig(kappa) transgene, but not V(kappa)J(kappa) gene segment targeted into the Ig(kappa) locus, can rescue B cell development in lambda5-deficient mice. *Immunity*. 1996;5(3):229-239.

36. Collins AM, Watson CT. Immunoglobulin Light Chain Gene Rearrangements, Receptor Editing and the Development of a Self-Tolerant Antibody Repertoire. *Front Immunol*. 2018;9:2249.

37. Martensson IL, Ceredig R. Review article: role of the surrogate light chain and the pre-B-cell receptor in mouse B-cell development. *Immunology*. 2000;101(4):435-441.

38. Tiegs SL, Russell DM, Nemazee D. Receptor editing in self-reactive bone marrow B cells. *J Exp Med*. 1993;177(4):1009-1020.

39. Gay D, Saunders T, Camper S, Weigert M. Receptor editing: an approach by autoreactive B cells to escape tolerance. *J Exp Med*. 1993;177(4):999-1008.

40. Nemazee D, Weigert M. Revising B cell receptors. *J Exp Med*. 2000;191(11):1813-1817.

41. Rolink A, Grawunder U, Haasner D, Strasser A, Melchers F. Immature surface Ig+ B cells can continue to rearrange kappa and lambda L chain gene loci. *J Exp Med*. 1993;178(4):1263-1270.

42. Nemazee DA, Burki K. Clonal deletion of B lymphocytes in a transgenic mouse bearing anti-MHC class I antibody genes. *Nature*. 1989;337(6207):562-566.

43. Hartley SB, Cooke MP, Fulcher DA, et al. Elimination of self-reactive B lymphocytes proceeds in two stages: arrested development and cell death. *Cell*. 1993;72(3):325-335.

44. Duong BH, Ota T, Aoki-Ota M, et al. Negative selection by IgM superantigen defines a B cell central tolerance compartment and reveals mutations allowing escape. *J Immunol*. 2011;187(11):5596-5605.

45. Spanopoulou E, Roman CA, Corcoran LM, et al. Functional immunoglobulin transgenes guide ordered B-cell differentiation in Rag-1-deficient mice. *Genes Dev*. 1994;8(9):1030-1042.

46. Chen C, Nagy Z, Radic MZ, et al. The site and stage of anti-DNA B-cell deletion. *Nature*. 1995;373(6511):252-255.

47. Vela JL, Ait-Azzouzene D, Duong BH, Ota T, Nemazee D. Rearrangement of mouse immunoglobulin kappa deleting element recombining sequence promotes immune tolerance and lambda B cell production. *Immunity*. 2008;28(2):161-170.

48. Sandel PC, Monroe JG. Negative selection of immature B cells by receptor editing or deletion is determined by site of antigen encounter. *Immunity*. 1999;10(3):289-299.

49. Oettinger MA, Schatz DG, Gorka C, Baltimore D. RAG-1 and RAG-2, adjacent genes that synergistically activate V(D)J recombination. *Science*. 1990;248(4962):1517-1523.

50. Hozumi N, Tonegawa S. Evidence for somatic rearrangement of immunoglobulin genes coding for variable and constant regions. *Proc Natl Acad Sci U S A*. 1976;73(10):3628-3632.

51. Roth DB. V(D)J Recombination: Mechanism, Errors, and Fidelity. *Microbiol Spectr*. 2014;2(6).

52. Early P, Huang H, Davis M, Calame K, Hood L. An immunoglobulin heavy chain variable region gene is generated from three segments of DNA: VH, D and JH. *Cell*. 1980;19(4):981-992.

53. Schilling J, Clevinger B, Davie JM, Hood L. Amino acid sequence of homogeneous antibodies to dextran and DNA rearrangements in heavy chain V-region gene segments. *Nature*. 1980;283(5742):35-40.

54. Benedict CL, Gilfillan S, Thai TH, Kearney JF. Terminal deoxynucleotidyl transferase and repertoire development. *Immunol Rev*. 2000;175:150-157.

55. Schatz DG, Ji Y. Recombination centres and the orchestration of V(D)J recombination. *Nat Rev Immunol*. 2011;11(4):251-263.

56. Tonegawa S. Somatic generation of antibody diversity. *Nature*. 1983;302(5909):575-581.

57. Mostoslavsky R, Alt FW, Rajewsky K. The lingering enigma of the allelic exclusion mechanism. *Cell*. 2004;118(5):539-544.

58. Alt FW, Enea V, Bothwell AL, Baltimore D. Activity of multiple light chain genes in murine myeloma cells producing a single, functional light chain. *Cell*. 1980;21(1):1-12.

59. Hieter PA, Korsmeyer SJ, Waldmann TA, Leder P. Human immunoglobulin kappa light-chain genes are deleted or rearranged in lambda-producing B cells. *Nature*. 1981;290(5805):368-372.

60. Li S, Liu W, Li Y, et al. Contribution of secondary Igkappa rearrangement to primary immunoglobulin repertoire diversification. *Mol Immunol*. 2016;78:193-206.

61. Rolink AG, Andersson J, Melchers F. Characterization of immature B cells by a novel monoclonal antibody, by turnover and by mitogen reactivity. *Eur J Immunol*. 1998;28(11):3738-3748.

62. ten Boekel E, Melchers F, Rolink AG. Changes in the V(H) gene repertoire of developing precursor B lymphocytes in mouse bone marrow mediated by the pre-B cell receptor. *Immunity*. 1997;7(3):357-368.

63. Wardemann H, Nussenzweig MC. B-cell self-tolerance in humans. *Adv Immunol*. 2007;95:83-110.

64. Yurasov S, Wardemann H, Hammersen J, et al. Defective B cell tolerance checkpoints in systemic lupus erythematosus. *J Exp Med*. 2005;201(5):703-711.

65. Victora GD, Nussenzweig MC. Germinal centers. *Annu Rev Immunol*. 2012;30:429-457.

66. Batista FD, Harwood NE. The who, how and where of antigen presentation to B cells. *Nat Rev Immunol*. 2009;9(1):15-27.

67. Okada T, Miller MJ, Parker I, et al. Antigen-engaged B cells undergo chemotaxis toward the T zone and form motile conjugates with helper T cells. *PLoS Biol*. 2005;3(6):e150.

68. Qi H, Cannons JL, Klauschen F, Schwartzberg PL, Germain RN. SAP-controlled T-B cell interactions underlie germinal centre formation. *Nature*. 2008;455(7214):764-769.

69. Kerfoot SM, Yaari G, Patel JR, et al. Germinal center B cell and T follicular helper cell development initiates in the interfollicular zone. *Immunity*. 2011;34(6):947-960.

70. Kitano M, Moriyama S, Ando Y, et al. Bcl6 protein expression shapes pre-germinal center B cell dynamics and follicular helper T cell heterogeneity. *Immunity*. 2011;34(6):961-972.

71. De Silva NS, Klein U. Dynamics of B cells in germinal centres. *Nat Rev Immunol*. 2015;15(3):137-148.

72. Coffey F, Alabyev B, Manser T. Initial clonal expansion of germinal center B cells takes place at the perimeter of follicles. *Immunity*. 2009;30(4):599-609.

73. MacLennan IC. Germinal centers. *Annu Rev Immunol*. 1994;12:117-139.

74. Allen CD, Okada T, Cyster JG. Germinal-center organization and cellular dynamics. *Immunity*. 2007;27(2):190-202.

75. Jacob J, Kelsoe G. In situ studies of the primary immune response to (4-hydroxy-3-nitrophenyl)acetyl. II. A common clonal origin for periarteriolar lymphoid sheath-associated foci and germinal centers. *J Exp Med*. 1992;176(3):679-687.

76. Benson MJ, Erickson LD, Gleeson MW, Noelle RJ. Affinity of antigen encounter and other early B-cell signals determine B-cell fate. *Curr Opin Immunol*. 2007;19(3):275-280.

77. Dal Porto JM, Haberman AM, Shlomchik MJ, Kelsoe G. Antigen drives very low affinity B cells to become plasmacytes and enter germinal centers. *J Immunol*. 1998;161(10):5373-5381.

78. Dal Porto JM, Haberman AM, Kelsoe G, Shlomchik MJ. Very low affinity B cells form germinal centers, become memory B cells, and participate in secondary immune responses when higher affinity competition is reduced. *J Exp Med*. 2002;195(9):1215-1221.

79. Paus D, Phan TG, Chan TD, Gardam S, Basten A, Brink R. Antigen recognition strength regulates the choice between extrafollicular plasma cell and germinal center B cell differentiation. *J Exp Med*. 2006;203(4):1081-1091.

80. Shih TA, Meffre E, Roederer M, Nussenzweig MC. Role of BCR affinity in T cell dependent antibody responses in vivo. *Nat Immunol*. 2002;3(6):570-575.

81. de Vinuesa CG, Cook MC, Ball J, et al. Germinal centers without T cells. *J Exp Med*. 2000;191(3):485-494.

82. Delves PJ, Roitt IM. The immune system. First of two parts. *N Engl J Med*. 2000;343(1):37-49.

83. Samardzic T, Gerlach J, Muller K, et al. CD22 regulates early B cell development in BOB.1/OBF.1-deficient mice. *Eur J Immunol*. 2002;32(9):2481-2489.

84. Allen CD, Ansel KM, Low C, et al. Germinal center dark and light zone organization is mediated by CXCR4 and CXCR5. *Nat Immunol*. 2004;5(9):943-952.

85. Victora GD, Schwickert TA, Fooksman DR, et al. Germinal center dynamics revealed by multiphoton microscopy with a photoactivatable fluorescent reporter. *Cell*. 2010;143(4):592-605.

86. Fyfe G, Cebra-Thomas JA, Mustain E, Davie JM, Alley CD, Nahm MH. Subpopulations of B lymphocytes in germinal centers. *J Immunol*. 1987;139(7):2187-2194.

87. Klein U, Tu Y, Stolovitzky GA, et al. Transcriptional analysis of the B cell germinal center reaction. *Proc Natl Acad Sci U S A*. 2003;100(5):2639-2644.

88. Hogerkorp CM, Borrebaeck CA. The human CD77- B cell population represents a heterogeneous subset of cells comprising centroblasts, centrocytes, and plasmablasts, prompting phenotypical revision. *J Immunol*. 2006;177(7):4341-4349.

89. Nakayama Y, Stabach P, Maher SE, et al. A limited number of genes are involved in the differentiation of germinal center B cells. *J Cell Biochem*. 2006;99(5):1308-1325.

90. Allen CD, Okada T, Tang HL, Cyster JG. Imaging of germinal center selection events during affinity maturation. *Science*. 2007;315(5811):528-531.

91. Caron G, Le Gallou S, Lamy T, Tarte K, Fest T. CXCR4 expression functionally discriminates centroblasts versus centrocytes within human germinal center B cells. *J Immunol*. 2009;182(12):7595-7602.

92. Song S, Matthias PD. The Transcriptional Regulation of Germinal Center Formation. *Front Immunol*. 2018;9:2026.

93. Schwickert TA, Lindquist RL, Shakhar G, et al. In vivo imaging of germinal centres reveals a dynamic open structure. *Nature*. 2007;446(7131):83-87.

94. Khalil AM, Cambier JC, Shlomchik MJ. B Cell Receptor Signal Transduction in the GC Is Short-Circuited by High Phosphatase Activity. *Science*. 2012;336(6085):1178-1181.

95. Bannard O, Horton RM, Allen CD, An J, Nagasawa T, Cyster JG. Germinal center centroblasts transition to a centrocyte phenotype according to a timed program and depend on the dark zone for effective selection. *Immunity*. 2013;39(5):912-924.

96. Hauser AE, Junt T, Mempel TR, et al. Definition of germinal-center B cell migration in vivo reveals predominant intrazonal circulation patterns. *Immunity*. 2007;26(5):655-667.

97. Allen CD, Cyster JG. Follicular dendritic cell networks of primary follicles and germinal centers: phenotype and function. *Semin Immunol*. 2008;20(1):14-25.

98. Hardie DL, Johnson GD, Khan M, MacLennan IC. Quantitative analysis of molecules which distinguish functional compartments within germinal centers. *Eur J Immunol*. 1993;23(5):997-1004.

99. Mandel TE, Phipps RP, Abbot AP, Tew JG. Long-term antigen retention by dendritic cells in the popliteal lymph node of immunized mice. *Immunology*. 1981;43(2):353-362.

100. Szakal AK, Gieringer RL, Kosco MH, Tew JG. Isolated follicular dendritic cells: cytochemical antigen localization, Nomarski, SEM, and TEM morphology. *J Immunol*. 1985;134(3):1349-1359.

101. Muramatsu M, Kinoshita K, Fagarasan S, Yamada S, Shinkai Y, Honjo T. Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. *Cell*. 2000;102(5):553-563.

102. Forster R, Mattis AE, Kremmer E, Wolf E, Brem G, Lipp M. A putative chemokine receptor, BLR1, directs B cell migration to defined lymphoid organs and specific anatomic compartments of the spleen. *Cell*. 1996;87(6):1037-1047.

103. Haynes NM, Allen CD, Lesley R, Ansel KM, Killeen N, Cyster JG. Role of CXCR5 and CCR7 in follicular Th cell positioning and appearance of a programmed cell death gene-1high germinal center-associated subpopulation. *J Immunol*. 2007;179(8):5099-5108.

104. Kopf M, Herren S, Wiles MV, Pepys MB, Kosco-Vilbois MH. Interleukin 6 influences germinal center development and antibody production via a contribution of C3 complement component. *J Exp Med*. 1998;188(10):1895-1906.

105. Wu Y, El Shikh ME, El Sayed RM, Best AM, Szakal AK, Tew JG. IL-6 produced by immune complex-activated follicular dendritic cells promotes germinal center reactions, IgG responses and somatic hypermutation. *Int Immunol*. 2009;21(6):745-756.

106. Nishikawa Y, Hikida M, Magari M, et al. Establishment of lymphotoxin beta receptor signaling-dependent cell lines with follicular dendritic cell phenotypes from mouse lymph nodes. *J Immunol*. 2006;177(8):5204-5214.

107. Victoratos P, Lagnel J, Tzima S, et al. FDC-specific functions of p55TNFR and IKK2 in the development of FDC networks and of antibody responses. *Immunity*. 2006;24(1):65-77.

108. Winter DB, Gearhart PJ. Dual enigma of somatic hypermutation of immunoglobulin variable genes: targeting and mechanism. *Immunol Rev*. 1998;162:89-96.

109. Lebecque SG, Gearhart PJ. Boundaries of somatic mutation in rearranged immunoglobulin genes: 5' boundary is near the promoter, and 3' boundary is approximately 1 kb from V(D)J gene. *J Exp Med*. 1990;172(6):1717-1727.

110. Cattoretti G, Buttner M, Shaknovich R, Kremmer E, Alobeid B, Niedobitek G. Nuclear and cytoplasmic AID in extrafollicular and germinal center B cells. *Blood*. 2006;107(10):3967-3975.

111. Moldenhauer G, Popov SW, Wotschke B, et al. AID expression identifies interfollicular large B cells as putative precursors of mature B-cell malignancies. *Blood*. 2006;107(6):2470-2473.

112. Papavasiliou FN, Schatz DG. Somatic hypermutation of immunoglobulin genes: merging mechanisms for genetic diversity. *Cell*. 2002;109 Suppl:S35-44.

113. Teng G, Papavasiliou FN. Immunoglobulin somatic hypermutation. *Annu Rev Genet*. 2007;41:107-120.

114. Larson ED, Cummings WJ, Bednarski DW, Maizels N. MRE11/RAD50 cleaves DNA in the AID/UNG-dependent pathway of immunoglobulin gene diversification. *Mol Cell*. 2005;20(3):367-375.

115. Wilson TM, Vaisman A, Martomo SA, et al. MSH2-MSH6 stimulates DNA polymerase eta, suggesting a role for A:T mutations in antibody genes. *J Exp Med*. 2005;201(4):637-645.

116. Odegard VH, Schatz DG. Targeting of somatic hypermutation. *Nat Rev Immunol*. 2006;6(8):573-583.

117. Pavri R, Gazumyan A, Jankovic M, et al. Activation-induced cytidine deaminase targets DNA at sites of RNA polymerase II stalling by interaction with Spt5. *Cell*. 2010;143(1):122-133.

118. Basu U, Meng FL, Keim C, et al. The RNA exosome targets the AID cytidine deaminase to both strands of transcribed duplex DNA substrates. *Cell*. 2011;144(3):353-363.

119. Rada C, Milstein C. The intrinsic hypermutability of antibody heavy and light chain genes decays exponentially. *EMBO J*. 2001;20(16):4570-4576.

120. Bachl J, Carlson C, Gray-Schopfer V, Dessing M, Olsson C. Increased transcription levels induce higher mutation rates in a hypermutating cell line. *J Immunol*. 2001;166(8):5051-5057.

121. Goyenechea B, Klix N, Yelamos J, et al. Cells strongly expressing Ig(kappa) transgenes show clonal recruitment of hypermutation: a role for both MAR and the enhancers. *EMBO J*. 1997;16(13):3987-3994.

122. Betz AG, Milstein C, Gonzalez-Fernandez A, Pannell R, Larson T, Neuberger MS. Elements regulating somatic hypermutation of an immunoglobulin kappa gene: critical role for the intron enhancer/matrix attachment region. *Cell*. 1994;77(2):239-248.

123. Tumas-Brundage K, Manser T. The transcriptional promoter regulates hypermutation of the antibody heavy chain locus. *J Exp Med*. 1997;185(2):239-250.

124. Nussenzweig A, Nussenzweig MC. Origin of chromosomal translocations in lymphoid cancer. *Cell*. 2010;141(1):27-38.

125. Flemming A. Class-switch recombination revised. *Nat Rev Immunol*. 2019;19(10):596-597.

126. Roco JA, Mesin L, Binder SC, et al. Class-Switch Recombination Occurs Infrequently in Germinal Centers. *Immunity*. 2019;51(2):337-350 e337.

127. Stavnezer J, Guikema JE, Schrader CE. Mechanism and regulation of class switch recombination. *Annu Rev Immunol*. 2008;26:261-292.

128. Kawabe T, Naka T, Yoshida K, et al. The immune responses in CD40-deficient mice: impaired immunoglobulin class switching and germinal center formation. *Immunity*. 1994;1(3):167-178.

129. Snapper CM, Mond JJ. Towards a comprehensive view of immunoglobulin class switching. *Immunol Today*. 1993;14(1):15-17.

130. Lorenz M, Jung S, Radbruch A. Switch transcripts in immunoglobulin class switching. *Science*. 1995;267(5205):1825-1828.

131. Okazaki I, Yoshikawa K, Kinoshita K, Muramatsu M, Nagaoka H, Honjo T. Activation-induced cytidine deaminase links class switch recombination and somatic hypermutation. *Ann N Y Acad Sci*. 2003;987:1-8.

132. Barreto V, Reina-San-Martin B, Ramiro AR, McBride KM, Nussenzweig MC. C-terminal deletion of AID uncouples class switch recombination from somatic hypermutation and gene conversion. *Mol Cell*. 2003;12(2):501-508.

133. Ta VT, Nagaoka H, Catalan N, et al. AID mutant analyses indicate requirement for class-switch-specific cofactors. *Nat Immunol*. 2003;4(9):843-848.

134. Guikema JE, Linehan EK, Tsuchimoto D, et al. APE1- and APE2-dependent DNA breaks in immunoglobulin class switch recombination. *J Exp Med*. 2007;204(12):3017-3026.

135. Rada C, Williams GT, Nilsen H, Barnes DE, Lindahl T, Neuberger MS. Immunoglobulin isotype switching is inhibited and somatic hypermutation perturbed in UNG-deficient mice. *Curr Biol*. 2002;12(20):1748-1755.

136. Robinson MJ, Ding Z, Pitt C, et al. The Amount of BCL6 in B Cells Shortly after Antigen Engagement Determines Their Representation in Subsequent Germinal Centers. *Cell Rep*. 2020;30(5):1530-1541 e1534.

137. Xu J, Husain A, Hu W, Honjo T, Kobayashi M. APE1 is dispensable for S-region cleavage but required for its repair in class switch recombination. *Proc Natl Acad Sci U S A*. 2014;111(48):17242-17247.

138. Xu Z, Zan H, Pone EJ, Mai T, Casali P. Immunoglobulin class-switch DNA recombination: induction, targeting and beyond. *Nat Rev Immunol*. 2012;12(7):517-531.

139. Mesin L, Ersching J, Victora GD. Germinal Center B Cell Dynamics. *Immunity*. 2016;45(3):471-482.

140. Depoil D, Zaru R, Guiraud M, et al. Immunological synapses are versatile structures enabling selective T cell polarization. *Immunity*. 2005;22(2):185-194.

141. Batista FD, Iber D, Neuberger MS. B cells acquire antigen from target cells after synapse formation. *Nature*. 2001;411(6836):489-494.

142. Batista FD, Arana E, Barral P, et al. The role of integrins and coreceptors in refining thresholds for B-cell responses. *Immunol Rev*. 2007;218:197-213.

143. Falkner FG, Zachau HG. Correct transcription of an immunoglobulin kappa gene requires an upstream fragment containing conserved sequence elements. *Nature*. 1984;310(5972):71-74.

144. Parslow TG, Blair DL, Murphy WJ, Granner DK. Structure of the 5' ends of immunoglobulin genes: a novel conserved sequence. *Proc Natl Acad Sci U S A*. 1984;81(9):2650-2654.

145. Mason JO, Williams GT, Neuberger MS. Transcription cell type specificity is conferred by an immunoglobulin VH gene promoter that includes a functional consensus sequence. *Cell*. 1985;41(2):479-487.

146. Gerster T, Matthias P, Thali M, Jiricny J, Schaffner W. Cell type-specificity elements of the immunoglobulin heavy chain gene enhancer. *EMBO J*. 1987;6(5):1323-1330.

147. Lenardo M, Pierce JW, Baltimore D. Protein-binding sites in Ig gene enhancers determine transcriptional activity and inducibility. *Science*. 1987;236(4808):1573-1577.

148. Wirth T, Staudt L, Baltimore D. An octamer oligonucleotide upstream of a TATA motif is sufficient for lymphoid-specific promoter activity. *Nature*. 1987;329(6135):174-178.

149. Mizushima-Sugano J, Roeder RG. Cell-type-specific transcription of an immunoglobulin kappa light chain gene in vitro. *Proc Natl Acad Sci U S A*. 1986;83(22):8511-8515.

150. Dreyfus M, Doyen N, Rougeon F. The conserved decanucleotide from the immunoglobulin heavy chain promoter induces a very high transcriptional activity in B-cells when introduced into an heterologous promoter. *EMBO J*. 1987;6(6):1685-1690.

151. Voliva CF, Jabrane-Ferrat N, Peterlin BM. The function of the octamer-binding site in the DRA promoter. *Immunogenetics*. 1996;43(1-2):20-26.

152. Malone CS, Patrone L, Wall R. An essential octamer motif in the mb-1 (Igalpha) promoter. *Mol Immunol*. 2000;37(6):321-328.

153. Chiefari E, Arcidiacono B, Possidente K, et al. Transcriptional regulation of the HMGA1 gene by octamer-binding proteins Oct-1 and Oct-2. *PLoS One*. 2013;8(12):e83969.

154. Sherman PA, Basta PV, Heguy A, Wloch MK, Roeder RG, Ting JP. The octamer motif is a B-lymphocyte-specific regulatory element of the HLA-DR alpha gene promoter. *Proc Natl Acad Sci U S A*. 1989;86(17):6739-6743.

155. LaBella F, Sive HL, Roeder RG, Heintz N. Cell-cycle regulation of a human histone H2b gene is mediated by the H2b subtype-specific consensus element. *Genes Dev*. 1988;2(1):32-39.

156. Murphy S, Pierani A, Scheidereit C, Melli M, Roeder RG. Purified octamer binding transcription factors stimulate RNA polymerase III--mediated transcription of the 7SK RNA gene. *Cell*. 1989;59(6):1071-1080.

157. Klemm JD, Rould MA, Aurora R, Herr W, Pabo CO. Crystal structure of the Oct-1 POU domain bound to an octamer site: DNA recognition with tethered DNA-binding modules. *Cell*. 1994;77(1):21-32.

158. Herr W, Cleary MA. The POU domain: versatility in transcriptional regulation by a flexible two-in-one DNA-binding domain. *Genes Dev*. 1995;9(14):1679-1693.

159. Staudt LM, Clerc RG, Singh H, LeBowitz JH, Sharp PA, Baltimore D. Cloning of a lymphoid-specific cDNA encoding a protein binding the regulatory octamer DNA motif. *Science*. 1988;241(4865):577-580.

160. Sturm RA, Das G, Herr W. The ubiquitous octamer-binding protein Oct-1 contains a POU domain with a homeo box subdomain. *Genes Dev*. 1988;2(12A):1582-1599.

161. Clerc RG, Corcoran LM, LeBowitz JH, Baltimore D, Sharp PA. The B-cell-specific Oct-2 protein contains POU box- and homeo box-type domains. *Genes Dev*. 1988;2(12A):1570-1581.

162. Muller MM, Ruppert S, Schaffner W, Matthias P. A cloned octamer transcription factor stimulates transcription from lymphoid-specific promoters in non-B cells. *Nature*. 1988;336(6199):544-551.

163. Scheidereit C, Cromlish JA, Gerster T, et al. A human lymphoid-specific transcription factor that activates immunoglobulin genes is a homoeobox protein. *Nature*. 1988;336(6199):551-557.

164. O'Neill EA, Kelly TJ. Purification and characterization of nuclear factor III (origin recognition protein C), a sequence-specific DNA binding protein required for efficient initiation of adenovirus DNA replication. *J Biol Chem*. 1988;263(2):931-937.

165. Scheidereit C, Heguy A, Roeder RG. Identification and purification of a human lymphoid-specific octamer-binding protein (OTF-2) that activates transcription of an immunoglobulin promoter in vitro. *Cell*. 1987;51(5):783-793.

166. Hanke JH, Landolfi NF, Tucker PW, Capra JD. Identification of murine nuclear proteins that bind to the conserved octamer sequence of the immunoglobulin promoter region. *Proc Natl Acad Sci U S A*. 1988;85(10):3560-3564.

167. Botfield MC, Jancso A, Weiss MA. Biochemical characterization of the Oct-2 POU domain with implications for bipartite DNA recognition. *Biochemistry*. 1992;31(25):5841-5848.

168. Aurora R, Herr W. Segments of the POU domain influence one another's DNA-binding specificity. *Mol Cell Biol*. 1992;12(2):455-467.

169. Verrijzer CP, Alkema MJ, van Weperen WW, Van Leeuwen HC, Strating MJ, van der Vliet PC. The DNA binding specificity of the bipartite POU domain and its subdomains. *EMBO J*. 1992;11(13):4993-5003.

170. Wang VE, Schmidt T, Chen J, Sharp PA, Tantin D. Embryonic lethality, decreased erythropoiesis, and defective octamer-dependent promoter activation in Oct-1-deficient mice. *Mol Cell Biol*. 2004;24(3):1022-1032.

171. Wang VE, Tantin D, Chen J, Sharp PA. B cell development and immunoglobulin transcription in Oct-1-deficient mice. *Proc Natl Acad Sci U S A*. 2004;101(7):2005-2010.

172. Schild-Poulter C, Shih A, Tantin D, et al. DNA-PK phosphorylation sites on Oct-1 promote cell survival following DNA damage. *Oncogene*. 2007;26(27):3980-3988.

173. Wang RH, Yu H, Deng CX. A requirement for breast-cancer-associated gene 1 (BRCA1) in the spindle checkpoint. *Proc Natl Acad Sci U S A*. 2004;101(49):17108-17113.

174. Fan W, Jin S, Tong T, et al. BRCA1 regulates GADD45 through its interactions with the OCT-1 and CAAT motifs. *J Biol Chem*. 2002;277(10):8061-8067.

175. Stoykova AS, Sterrer S, Erselius JR, Hatzopoulos AK, Gruss P. Mini-Oct and Oct-2c: two novel, functionally diverse murine Oct-2 gene products are differentially expressed in the CNS. *Neuron*. 1992;8(3):541-558.

176. Corcoran LM, Karvelas M, Nossal GJ, Ye ZS, Jacks T, Baltimore D. Oct-2, although not required for early B-cell development, is critical for later B-cell maturation and for postnatal survival. *Genes Dev*. 1993;7(4):570-582.

177. Corcoran LM, Karvelas M. Oct-2 is required early in T cell-independent B cell activation for G1 progression and for proliferation. *Immunity*. 1994;1(8):635-645.

178. Annweiler A, Muller-Immergluck M, Wirth T. Oct2 transactivation from a remote enhancer position requires a B-cell-restricted activity. *Mol Cell Biol*. 1992;12(7):3107-3116.

179. Corcoran LM, Koentgen F, Dietrich W, Veale M, Humbert PO. All known in vivo functions of the Oct-2 transcription factor require the C-terminal protein domain. *J Immunol*. 2004;172(5):2962-2969.

180. Landolfi NF, Capra JD, Tucker PW. Interaction of cell-type-specific nuclear proteins with immunoglobulin VH promoter region sequences. *Nature*. 1986;323(6088):548-551.

181. Staudt LM, Singh H, Sen R, Wirth T, Sharp PA, Baltimore D. A lymphoid-specific protein binding to the octamer motif of immunoglobulin genes. *Nature*. 1986;323(6089):640-643.

182. Luo Y, Roeder RG. Cloning, functional characterization, and mechanism of action of the B-cell-specific transcriptional coactivator OCA-B. *Mol Cell Biol*. 1995;15(8):4115-4124.

183. Zheng L, Roeder RG, Luo Y. S phase activation of the histone H2B promoter by OCA-S, a coactivator complex that contains GAPDH as a key component. *Cell*. 2003;114(2):255-266.

184. Zwilling S, Konig H, Wirth T. High mobility group protein 2 functionally interacts with the POU domains of octamer transcription factors. *EMBO J*. 1995;14(6):1198-1208.

185. Kakizawa T, Miyamoto T, Ichikawa K, et al. Functional interaction between Oct-1 and retinoid X receptor. *J Biol Chem*. 1999;274(27):19103-19108.

186. Wang JM, Prefontaine GG, Lemieux ME, Pope L, Akimenko MA, Hache RJ. Developmental effects of ectopic expression of the glucocorticoid receptor DNA binding domain are alleviated by an amino acid substitution that interferes with homeodomain binding. *Mol Cell Biol*. 1999;19(10):7106-7122.

187. Prefontaine GG, Walther R, Giffin W, Lemieux ME, Pope L, Hache RJ. Selective binding of steroid hormone receptors to octamer transcription factors determines transcriptional synergism at the mouse mammary tumor virus promoter. *J Biol Chem*. 1999;274(38):26713-26719.

188. Inamoto S, Segil N, Pan ZQ, Kimura M, Roeder RG. The cyclin-dependent kinase-activating kinase (CAK) assembly factor, MAT1, targets and enhances CAK activity on the POU domains of octamer transcription factors. *J Biol Chem*. 1997;272(47):29852-29858.

189. Zwilling S, Annweiler A, Wirth T. The POU domains of the Oct1 and Oct2 transcription factors mediate specific interaction with TBP. *Nucleic Acids Res*. 1994;22(9):1655-1662.

190. Stern S, Tanaka M, Herr W. The Oct-1 homoeodomain directs formation of a multiprotein-DNA complex with the HSV transactivator VP16. *Nature*. 1989;341(6243):624-630.

191. Cleary MA, Stern S, Tanaka M, Herr W. Differential positive control by Oct-1 and Oct-2: activation of a transcriptionally silent motif through Oct-1 and VP16 corecruitment. *Genes Dev*. 1993;7(1):72-83.

192. Teitell MA. OCA-B regulation of B-cell development and function. *Trends Immunol*. 2003;24(10):546-553.

193. Luo Y, Fujii H, Gerster T, Roeder RG. A novel B cell-derived coactivator potentiates the activation of immunoglobulin promoters by octamer-binding transcription factors. *Cell*. 1992;71(2):231-241.

194. Feldhaus AL, Klug CA, Arvin KL, Singh H. Targeted disruption of the Oct-2 locus in a B cell provides genetic evidence for two distinct cell type-specific pathways of octamer element-mediated gene activation. *EMBO J*. 1993;12(7):2763-2772.

195. Johnson DG, Carayannopoulos L, Capra JD, Tucker PW, Hanke JH. The ubiquitous octamer-binding protein(s) is sufficient for transcription of immunoglobulin genes. *Mol Cell Biol*. 1990;10(3):982-990.

196. Strubin M, Newell JW, Matthias P. OBF-1, a novel B cell-specific coactivator that stimulates immunoglobulin promoter activity through association with octamer-binding proteins. *Cell*. 1995;80(3):497-506.

197. Pierani A, Heguy A, Fujii H, Roeder RG. Activation of octamer-containing promoters by either octamer-binding transcription factor 1 (OTF-1) or OTF-2 and requirement of an additional B-cell-specific component for optimal transcription of immunoglobulin promoters. *Mol Cell Biol*. 1990;10(12):6204-6215.

198. Gstaiger M, Knoepfel L, Georgiev O, Schaffner W, Hovens CM. A B-cell coactivator of octamer-binding transcription factors. *Nature*. 1995;373(6512):360-362.

199. Schubart K, Massa S, Schubart D, Corcoran LM, Rolink AG, Matthias P. B cell development and immunoglobulin gene transcription in the absence of Oct-2 and OBF-1. *Nat Immunol*. 2001;2(1):69-74.

200. Schubart DB, Rolink A, Kosco-Vilbois MH, Botteri F, Matthias P. B-cell-specific coactivator OBF-1/OCA-B/Bob1 required for immune response and germinal centre formation. *Nature*. 1996;383(6600):538-542.

201. Kim U, Qin XF, Gong S, et al. The B-cell-specific transcription coactivator OCA-B/OBF-1/Bob-1 is essential for normal production of immunoglobulin isotypes. *Nature*. 1996;383(6600):542-547.

202. Nielsen PJ, Georgiev O, Lorenz B, Schaffner W. B lymphocytes are impaired in mice lacking the transcriptional co-activator Bob1/OCA-B/OBF1. *Eur J Immunol*. 1996;26(12):3214-3218.

203. Gstaiger M, Georgiev O, van Leeuwen H, van der Vliet P, Schaffner W. The B cell coactivator Bob1 shows DNA sequence-dependent complex formation with Oct-1/Oct-2 factors, leading to differential promoter activation. *EMBO J*. 1996;15(11):2781-2790.

204. Cepek KL, Chasman DI, Sharp PA. Sequence-specific DNA binding of the B-cell-specific coactivator OCA-B. *Genes Dev*. 1996;10(16):2079-2088.

205. Kemler I, Bucher E, Seipel K, Muller-Immergluck MM, Schaffner W. Promoters with the octamer DNA motif (ATGCAAAT) can be ubiquitous or cell type-specific depending on binding affinity of the octamer site and Oct-factor concentration. *Nucleic Acids Res*. 1991;19(2):237-242.

206. Kamps MP, Corcoran L, LeBowitz JH, Baltimore D. The promoter of the human interleukin-2 gene contains two octamer-binding sites and is partially activated by the expression of Oct-2. *Mol Cell Biol*. 1990;10(10):5464-5472.

207. Schirm S, Jiricny J, Schaffner W. The SV40 enhancer can be dissected into multiple segments, each with a different cell type specificity. *Genes Dev*. 1987;1(1):65-74.

208. Chasman D, Cepek K, Sharp PA, Pabo CO. Crystal structure of an OCA-B peptide bound to an Oct-1 POU domain/octamer DNA complex: specific recognition of a protein-DNA interface. *Genes Dev*. 1999;13(20):2650-2657.

209. Sauter P, Matthias P. Coactivator OBF-1 makes selective contacts with both the POU-specific domain and the POU homeodomain and acts as a molecular clamp on DNA. *Mol Cell Biol*. 1998;18(12):7397-7409.

210. Tomilin A, Remenyi A, Lins K, et al. Synergism with the coactivator OBF-1 (OCA-B, BOB-1) is mediated by a specific POU dimer configuration. *Cell*. 2000;103(6):853-864.

211. Yu X, Wang L, Luo Y, Roeder RG. Identification and characterization of a novel OCA-B isoform. implications for a role in B cell signaling pathways. *Immunity*. 2001;14(2):157-167.

212. Tiedt R, Bartholdy BA, Matthias G, Newell JW, Matthias P. The RING finger protein Siah-1 regulates the level of the transcriptional coactivator OBF-1. *EMBO J*. 2001;20(15):4143-4152.

213. Boehm J, He Y, Greiner A, Staudt L, Wirth T. Regulation of BOB.1/OBF.1 stability by SIAH. *EMBO J*. 2001;20(15):4153-4162.

214. Kemler I, Schreiber E, Muller MM, Matthias P, Schaffner W. Octamer transcription factors bind to two different sequence motifs of the immunoglobulin heavy chain promoter. *EMBO J*. 1989;8(7):2001-2008.

215. Poellinger L, Roeder RG. Octamer transcription factors 1 and 2 each bind to two different functional elements in the immunoglobulin heavy-chain promoter. *Mol Cell Biol*. 1989;9(2):747-756.

216. Kissinger CR, Liu BS, Martin-Blanco E, Kornberg TB, Pabo CO. Crystal structure of an engrailed homeodomain-DNA complex at 2.8 A resolution: a framework for understanding homeodomain-DNA interactions. *Cell*. 1990;63(3):579-590.

217. Otting G, Qian YQ, Billeter M, et al. Protein--DNA contacts in the structure of a homeodomain--DNA complex determined by nuclear magnetic resonance spectroscopy in solution. *EMBO J*. 1990;9(10):3085-3092.

218. Wolberger C, Vershon AK, Liu B, Johnson AD, Pabo CO. Crystal structure of a MAT alpha 2 homeodomain-operator complex suggests a general model for homeodomain-DNA interactions. *Cell*. 1991;67(3):517-528.

219. Pfisterer P, Zwilling S, Hess J, Wirth T. Functional characterization of the murine homolog of the B cell-specific coactivator BOB.1/OBF.1. *J Biol Chem*. 1995;270(50):29870-29880.

220. LeBowitz JH, Clerc RG, Brenowitz M, Sharp PA. The Oct-2 protein binds cooperatively to adjacent octamer sites. *Genes Dev*. 1989;3(10):1625-1638.

221. Botquin V, Hess H, Fuhrmann G, et al. New POU dimer configuration mediates antagonistic control of an osteopontin preimplantation enhancer by Oct-4 and Sox-2. *Genes Dev*. 1998;12(13):2073-2090.

222. Remenyi A, Tomilin A, Pohl E, et al. Differential dimer activities of the transcription factor Oct-1 by DNA-induced interface swapping. *Mol Cell*. 2001;8(3):569-580.

223. Prabhu A, O'Brien DP, Weisner GL, Fulton R, Van Ness B. Octamer independent activation of transcription from the kappa immunoglobulin germline promoter. *Nucleic Acids Res*. 1996;24(23):4805-4811.

224. Junker S, Pedersen S, Schreiber E, Matthias P. Extinction of an immunoglobulin kappa promoter in cell hybrids is mediated by the octamer motif and correlates with suppression of Oct-2 expression. *Cell*. 1990;61(3):467-474.

225. Corcoran L, Emslie D, Kratina T, et al. Oct2 and Obf1 as Facilitators of B:T Cell Collaboration during a Humoral Immune Response. *Front Immunol*. 2014;5:108.

226. Hardy RR, Hayakawa K. Development and physiology of Ly-1 B and its human homolog, Leu-1 B. *Immunol Rev*. 1986;93:53-79.

227. Herzenberg LA, Stall AM, Lalor PA, et al. The Ly-1 B cell lineage. *Immunol Rev*. 1986;93:81-102.

228. Hardy RR, Hayakawa K. CD5 B cells, a fetal B cell lineage. *Adv Immunol*. 1994;55:297-339.

229. Humbert PO, Corcoran LM. oct-2 gene disruption eliminates the peritoneal B-1 lymphocyte lineage and attenuates B-2 cell maturation and function. *J Immunol*. 1997;159(11):5273-5284.

230. Chevrier S, Kratina T, Emslie D, Karnowski A, Corcoran LM. Germinal center-independent, IgM-mediated autoimmunity in sanroque mice lacking Obf1. *Immunol Cell Biol*. 2014;92(1):12-19.

231. Emslie D, D'Costa K, Hasbold J, et al. Oct2 enhances antibody-secreting cell differentiation through regulation of IL-5 receptor alpha chain expression on activated B cells. *J Exp Med*. 2008;205(2):409-421.

232. Fehr T, Lopez-Macias C, Odermatt B, et al. Correlation of anti-viral B cell responses and splenic morphology with expression of B cell-specific molecules. *Int Immunol*. 2000;12(9):1275-1284.

233. Hasbold J, Corcoran LM, Tarlinton DM, Tangye SG, Hodgkin PD. Evidence from the generation of immunoglobulin G-secreting cells that stochastic mechanisms regulate lymphocyte differentiation. *Nat Immunol*. 2004;5(1):55-63.

234. Paul WE, Ohara J. B-cell stimulatory factor-1/interleukin 4. *Annu Rev Immunol*. 1987;5:429-459.

235. Karnowski A, Chevrier S, Belz GT, et al. B and T cells collaborate in antiviral responses via IL-6, IL-21, and transcriptional activator and coactivator, Oct2 and OBF-1. *J Exp Med*. 2012;209(11):2049-2064.

236. Hodson DJ, Shaffer AL, Xiao W, et al. Regulation of normal B-cell differentiation and malignant B-cell survival by OCT2. *Proc Natl Acad Sci U S A*. 2016;113(14):E2039-2046.

237. Qin XF, Reichlin A, Luo Y, Roeder RG, Nussenzweig MC. OCA-B integrates B cell antigen receptor-, CD40L- and IL 4-mediated signals for the germinal center pathway of B cell development. *EMBO J*. 1998;17(17):5066-5075.

238. Stauss D, Brunner C, Berberich-Siebelt F, Hopken UE, Lipp M, Muller G. The transcriptional coactivator Bob1 promotes the development of follicular T helper cells via Bcl6. *EMBO J*. 2016;35(8):881-898.

239. Corcoran LM, Hasbold J, Dietrich W, et al. Differential requirement for OBF-1 during antibody-secreting cell differentiation. *J Exp Med*. 2005;201(9):1385-1396.

240. Ren X, Siegel R, Kim U, Roeder RG. Direct interactions of OCA-B and TFII-I regulate immunoglobulin heavy-chain gene transcription by facilitating enhancer-promoter communication. *Mol Cell*. 2011;42(3):342-355.

241. Peperzak V, Vikstrom I, Walker J, et al. Mcl-1 is essential for the survival of plasma cells. *Nat Immunol*. 2013;14(3):290-297.

242. Levels MJ, Fehres CM, van Baarsen LGM, et al. BOB.1 controls memory B-cell fate in the germinal center reaction. *J Autoimmun*. 2019;101:131-144.

243. Casellas R, Jankovic M, Meyer G, et al. OcaB is required for normal transcription and V(D)J recombination of a subset of immunoglobulin kappa genes. *Cell*. 2002;110(5):575-585.

244. Saito M, Novak U, Piovan E, et al. BCL6 suppression of BCL2 via Miz1 and its disruption in diffuse large B cell lymphoma. *Proc Natl Acad Sci U S A*. 2009;106(27):11294-11299.

245. Cerchietti LC, Polo JM, Da Silva GF, et al. Sequential transcription factor targeting for diffuse large B-cell lymphomas. *Cancer Res*. 2008;68(9):3361-3369.

246. Garaud JC, Schickel JN, Blaison G, et al. B cell signature during inactive systemic lupus is heterogeneous: toward a biological dissection of lupus. *PLoS One*. 2011;6(8):e23900.

247. Reininger L, Winkler TH, Kalberer CP, Jourdan M, Melchers F, Rolink AG. Intrinsic B cell defects in NZB and NZW mice contribute to systemic lupus erythematosus in (NZB x NZW)F1 mice. *J Exp Med*. 1996;184(3):853-861.

248. Wang JH, Avitahl N, Cariappa A, et al. Aiolos regulates B cell activation and maturation to effector state. *Immunity*. 1998;9(4):543-553.

249. Sun J, Matthias G, Mihatsch MJ, Georgopoulos K, Matthias P. Lack of the transcriptional coactivator OBF-1 prevents the development of systemic lupus erythematosus-like phenotypes in Aiolos mutant mice. *J Immunol*. 2003;170(4):1699-1706.

250. Vinuesa CG, Cook MC, Angelucci C, et al. A RING-type ubiquitin ligase family member required to repress follicular helper T cells and autoimmunity. *Nature*. 2005;435(7041):452-458.

251. Chapuy B, McKeown MR, Lin CY, et al. Discovery and characterization of super-enhancer-associated dependencies in diffuse large B cell lymphoma. *Cancer Cell*. 2013;24(6):777-790.

252. Wang T, Birsoy K, Hughes NW, et al. Identification and characterization of essential genes in the human genome. *Science*. 2015;350(6264):1096-1101.

253. Willis SN, Tellier J, Liao Y, et al. Environmental sensing by mature B cells is controlled by the transcription factors PU.1 and SpiB. *Nat Commun*. 2017;8(1):1426.

254. Oracki SA, Walker JA, Hibbs ML, Corcoran LM, Tarlinton DM. Plasma cell development and survival. *Immunol Rev*. 2010;237(1):140-159.

255. Lins K, Remenyi A, Tomilin A, et al. OBF1 enhances transcriptional potential of Oct1. *EMBO J*. 2003;22(9):2188-2198.

256. Babb R, Cleary MA, Herr W. OCA-B is a functional analog of VP16 but targets a separate surface of the Oct-1 POU domain. *Mol Cell Biol*. 1997;17(12):7295-7305.

257. Ye BH, Cattoretti G, Shen Q, et al. The BCL-6 proto-oncogene controls germinal-centre formation and Th2-type inflammation. *Nat Genet*. 1997;16(2):161-170.

258. Duan H, Xiang H, Ma L, Boxer LM. Functional long-range interactions of the IgH 3' enhancers with the bcl-2 promoter region in t(14;18) lymphoma cells. *Oncogene*. 2008;27(53):6720-6728.

259. Han T, Oh S, Kang K. ETS family protein GABP is a novel co-factor strongly associated with genomic YY1 binding sites in various cell lines. *Genes & Genomics*. 2016;38(2):119-125.

260. Heinz S, Benner C, Spann N, et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell*. 2010;38(4):576-589.

261. Garrett-Sinha LA, Dahl R, Rao S, Barton KP, Simon MC. PU.1 exhibits partial functional redundancy with Spi-B, but not with Ets-1 or Elf-1. *Blood*. 2001;97(9):2908-2912.

262. Su GH, Chen HM, Muthusamy N, et al. Defective B cell receptor-mediated responses in mice lacking the Ets protein, Spi-B. *EMBO J*. 1997;16(23):7118-7129.

263. Ye M, Ermakova O, Graf T. PU.1 is not strictly required for B cell development and its absence induces a B-2 to B-1 cell switch. *J Exp Med*. 2005;202(10):1411-1422.

264. Saelee P, Kearly A, Nutt SL, Garrett-Sinha LA. Genome-Wide Identification of Target Genes for the Key B Cell Transcription Factor Ets1. *Front Immunol*. 2017;8:383.

265. Chan WY, Follows GA, Lacaud G, et al. The paralogous hematopoietic regulators Lyl1 and Scl are coregulated by Ets and GATA factors, but Lyl1 cannot rescue the early Scl-/- phenotype. *Blood*. 2007;109(5):1908-1916.

266. Gottgens B, Nastos A, Kinston S, et al. Establishing the transcriptional programme for blood: the SCL stem cell enhancer is regulated by a multiprotein complex containing Ets and GATA factors. *EMBO J*. 2002;21(12):3039-3050.

267. Gottgens B, Broccardo C, Sanchez MJ, et al. The scl +18/19 stem cell enhancer is not required for hematopoiesis: identification of a 5' bifunctional hematopoietic-endothelial enhancer bound by Fli-1 and Elf-1. *Mol Cell Biol*. 2004;24(5):1870-1883.

268. Rainis L, Toki T, Pimanda JE, et al. The proto-oncogene ERG in megakaryoblastic leukemias. *Cancer Res*. 2005;65(17):7596-7602.

269. Shapiro-Shelef M, Lin KI, McHeyzer-Williams LJ, Liao J, McHeyzer-Williams MG, Calame K. Blimp-1 is required for the formation of immunoglobulin secreting plasma cells and pre-plasma memory B cells. *Immunity*. 2003;19(4):607-620.

270. Kallies A, Hasbold J, Tarlinton DM, et al. Plasma cell ontogeny defined by quantitative changes in blimp-1 expression. *J Exp Med*. 2004;200(8):967-977.

271. Bordon A, Bosco N, Du Roure C, et al. Enforced expression of the transcriptional coactivator OBF1 impairs B cell differentiation at the earliest stage of development. *PLoS One*. 2008;3(12):e4007.

272. Heckman CA, Duan H, Garcia PB, Boxer LM. Oct transcription factors mediate t(14;18) lymphoma cell survival by directly regulating bcl-2 expression. *Oncogene*. 2006;25(6):888-898.

273. Hori N, Nakano H, Takeuchi T, et al. A dyad oct-binding sequence functions as a maintenance sequence for the unmethylated state within the H19/Igf2-imprinted control region. *J Biol Chem*. 2002;277(31):27960-27967.

274. Dong B, Zhao FQ. Involvement of the ubiquitous Oct-1 transcription factor in hormonal induction of beta-casein gene expression. *Biochem J*. 2007;401(1):57-64.

275. Poellinger L, Yoza BK, Roeder RG. Functional cooperativity between protein molecules bound at two distinct sequence elements of the immunoglobulin heavy-chain promoter. *Nature*. 1989;337(6207):573-576.

276. Soufi A, Garcia MF, Jaroszewicz A, Osman N, Pellegrini M, Zaret KS. Pioneer transcription factors target partial DNA motifs on nucleosomes to initiate reprogramming. *Cell*. 2015;161(3):555-568.

277. Kang J, Gemberling M, Nakamura M, et al. A general mechanism for transcription regulation by Oct1 and Oct4 in response to genotoxic and oxidative stress. *Genes Dev*. 2009;23(2):208-222.

278. Champhekar A, Damle SS, Freedman G, Carotta S, Nutt SL, Rothenberg EV. Regulation of early T-lineage gene expression and developmental progression by the progenitor cell transcription factor PU.1. *Genes Dev*. 2015;29(8):832-848.

279. Hatzi K, Melnick A. Breaking bad in the germinal center: how deregulation of BCL6 contributes to lymphomagenesis. *Trends Mol Med*. 2014;20(6):343-352.

280. Bunting KL, Soong TD, Singh R, et al. Multi-tiered Reorganization of the Genome during B Cell Affinity Maturation Anchored by a Germinal Center-Specific Locus Control Region. *Immunity*. 2016;45(3):497-512.

281. Greiner A, Muller KB, Hess J, Pfeffer K, Muller-Hermelink HK, Wirth T. Up-regulation of BOB.1/OBF.1 expression in normal germinal center B cells and germinal center-derived lymphomas. *Am J Pathol*. 2000;156(2):501-507.

282. Basso K, Klein U, Niu H, et al. Tracking CD40 signaling during germinal center development. *Blood*. 2004;104(13):4088-4096.

283. Saito M, Gao J, Basso K, et al. A signaling pathway mediating downregulation of BCL6 in germinal center B cells is blocked by BCL6 gene alterations in B cell lymphoma. *Cancer Cell*. 2007;12(3):280-292.

284. Ochiai K, Maienschein-Cline M, Simonetti G, et al. Transcriptional regulation of germinal center B and plasma cell fates by dynamical control of IRF4. *Immunity*. 2013;38(5):918-929.

285. Klein U, Casola S, Cattoretti G, et al. Transcription factor IRF4 controls plasma cell differentiation and class-switch recombination. *Nat Immunol*. 2006;7(7):773-782.

286. Shen Y, Hendershot LM. Identification of ERdj3 and OBF-1/BOB-1/OCA-B as direct targets of XBP-1 during plasma cell differentiation. *J Immunol*. 2007;179(5):2969-2978.

287. Filippakopoulos P, Qi J, Picaud S, et al. Selective inhibition of BET bromodomains. *Nature*. 2010;468(7327):1067-1073.

288. Teng Y, Takahashi Y, Yamada M, et al. IRF4 negatively regulates proliferation of germinal center B cell-derived Burkitt's lymphoma cell lines and induces differentiation toward plasma cells. *Eur J Cell Biol*. 2007;86(10):581-589.

289. Kuppers R, Dalla-Favera R. Mechanisms of chromosomal translocations in B cell lymphomas. *Oncogene*. 2001;20(40):5580-5594.

290. Taub R, Kirsch I, Morton C, et al. Translocation of the c-myc gene into the immunoglobulin heavy chain locus in human Burkitt lymphoma and murine plasmacytoma cells. *Proc Natl Acad Sci U S A*. 1982;79(24):7837-7841.

291. Kramer MH, Hermans J, Wijburg E, et al. Clinical relevance of BCL2, BCL6, and MYC rearrangements in diffuse large B-cell lymphoma. *Blood*. 1998;92(9):3152-3162.

292. Nishikura K, Erikson J, ar-Rushdi A, Huebner K, Croce CM. The translocated c-myc oncogene of Raji Burkitt lymphoma cells is not expressed in human lymphoblastoid cells. *Proc Natl Acad Sci U S A*. 1985;82(9):2900-2904.

293. Cleary ML, Smith SD, Sklar J. Cloning and structural analysis of cDNAs for bcl-2 and a hybrid bcl-2/immunoglobulin transcript resulting from the t(14;18) translocation. *Cell*. 1986;47(1):19-28.

294. Haluska FG, Tsujimoto Y, Croce CM. The t(8;14) chromosome translocation of the Burkitt lymphoma cell line Daudi occurred during immunoglobulin gene rearrangement and involved the heavy chain diversity region. *Proc Natl Acad Sci U S A*. 1987;84(19):6835-6839.

295. Li W, Gupta SK, Han W, et al. Targeting MYC activity in double-hit lymphoma with MYC and BCL2 and/or BCL6 rearrangements with epigenetic bromodomain inhibitors. *J Hematol Oncol*. 2019;12(1):73.

296. Ye BH, Chaganti S, Chang CC, et al. Chromosomal translocations cause deregulated BCL6 expression by promoter substitution in B cell lymphoma. *EMBO J*. 1995;14(24):6209-6217.

297. Lossos IS, Akasaka T, Martinez-Climent JA, Siebert R, Levy R. The BCL6 gene in B-cell lymphomas with 3q27 translocations is expressed mainly from the rearranged allele irrespective of the partner gene. *Leukemia*. 2003;17(7):1390-1397.

298. Wang Y, Shi J, Yan J, et al. Germinal-center development of memory B cells driven by IL-9 from follicular helper T cells. *Nat Immunol*. 2017;18(8):921-930.

299. Hagenbuchner J, Ausserlechner MJ. Targeting transcription factors by small compounds--Current strategies and future implications. *Biochem Pharmacol*. 2016;107:1-13.

300. Caboni L, Lloyd DG. Beyond the ligand-binding pocket: targeting alternate sites in nuclear receptors. *Med Res Rev*. 2013;33(5):1081-1118.

301. Leung CH, Chan DS, Ma VP, Ma DL. DNA-binding small molecules as inhibitors of transcription factors. *Med Res Rev*. 2013;33(4):823-846.

302. Adhikari B, Bozilovic J, Diebold M, et al. PROTAC-mediated degradation reveals a non-catalytic function of AURORA-A kinase. *Nat Chem Biol*. 2020;16(11):1179-1188.

303. Burslem GM, Crews CM. Proteolysis-Targeting Chimeras as Therapeutics and Tools for Biological Discovery. *Cell*. 2020;181(1):102-114.

304. Bondeson DP, Mares A, Smith IE, et al. Catalytic in vivo protein knockdown by small-molecule PROTACs. *Nat Chem Biol*. 2015;11(8):611-617.

305. Winter GE, Buckley DL, Paulk J, et al. DRUG DEVELOPMENT. Phthalimide conjugation as a strategy for in vivo target protein degradation. *Science*. 2015;348(6241):1376-1381.

306. Burslem GM, Smith BE, Lai AC, et al. The Advantages of Targeted Protein Degradation Over Inhibition: An RTK Case Study. *Cell Chem Biol*. 2018;25(1):67-77 e63.

307. Challa-Malladi M, Lieu YK, Califano O, et al. Combined genetic inactivation of beta2-Microglobulin and CD58 reveals frequent escape from immune recognition in diffuse large B cell lymphoma. *Cancer Cell*. 2011;20(6):728-740.

308. Nijland M, Veenstra RN, Visser L, et al. HLA dependent immune escape mechanisms in B-cell lymphomas: Implications for immune checkpoint inhibitor therapy? *Oncoimmunology*. 2017;6(4):e1295202.

309. Tada K, Maeshima AM, Hiraoka N, et al. Prognostic significance of HLA class I and II expression in patients with diffuse large B cell lymphoma treated with standard chemoimmunotherapy. *Cancer Immunol Immunother*. 2016;65(10):1213-1222.

310. Rimsza LM, Roberts RA, Miller TP, et al. Loss of MHC class II gene and protein expression in diffuse large B-cell lymphoma is related to decreased tumor immunosurveillance and poor patient survival regardless of other prognostic factors: a follow-up study from the Leukemia and Lymphoma Molecular Profiling Project. *Blood*. 2004;103(11):4251-4258.

311. Rosenwald A, Wright G, Chan WC, et al. The use of molecular profiling to predict survival after chemotherapy for diffuse large-B-cell lymphoma. *N Engl J Med*. 2002;346(25):1937-1947.

312. Cycon KA, Rimsza LM, Murphy SP. Alterations in CIITA constitute a common mechanism accounting for downregulation of MHC class II expression in diffuse large B-cell lymphoma (DLBCL). *Exp Hematol*. 2009;37(2):184-194.

313. Jiang Y, Ortega-Molina A, Geng H, et al. CREBBP Inactivation Promotes the Development of HDAC3-Dependent Lymphomas. *Cancer Discov*. 2017;7(1):38-53.

314. Mondello P, Tadros S, Teater M, et al. Selective Inhibition of HDAC3 Targets Synthetic Vulnerabilities and Activates Immune Surveillance in Lymphoma. *Cancer Discov*. 2020;10(3):440-459.

315. Greaves P, Gribben JG. The role of B7 family molecules in hematologic malignancy. *Blood*. 2013;121(5):734-744.

316. Stopeck AT, Gessner A, Miller TP, et al. Loss of B7.2 (CD86) and intracellular adhesion molecule 1 (CD54) expression is associated with decreased tumor-infiltrating T lymphocytes in diffuse B-cell large-cell lymphoma. *Clin Cancer Res*. 2000;6(10):3904-3909.

317. Dustin ML. The immunological synapse. *Cancer Immunol Res*. 2014;2(11):1023-1033.

318. Ren S, Tian Q, Amar N, et al. The immune checkpoint, HVEM may contribute to immune escape in non-small cell lung cancer lacking PD-L1 expression. *Lung Cancer*. 2018;125:115-120.

319. Cardenas MG, Oswald E, Yu W, Xue F, MacKerell AD, Jr., Melnick AM. The Expanding Role of the BCL6 Oncoprotein as a Cancer Therapeutic Target. *Clin Cancer Res*. 2017;23(4):885-893.

320. Muris JJ, Meijer CJ, Ossenkoppele GJ, Vos W, Oudejans JJ. Apoptosis resistance and response to chemotherapy in primary nodal diffuse large B-cell lymphoma. *Hematol Oncol*. 2006;24(3):97-104.

321. Driegen S, Ferreira R, van Zon A, et al. A generic tool for biotinylation of tagged proteins in transgenic mice. *Transgenic Res*. 2005;14(4):477-482.

322. Yang H, Wang H, Jaenisch R. Generating genetically modified mice using CRISPR/Cas-mediated genome engineering. *Nat Protoc*. 2014;9(8):1956-1968.

323. Cato MH, Yau IW, Rickert RC. Magnetic-based purification of untouched mouse germinal center B cells for ex vivo manipulation and biochemical analysis. *Nat Protoc*. 2011;6(7):953-960.

324. Choukrallah MA, Song S, Rolink AG, Burger L, Matthias P. Enhancer repertoires are reshaped independently of early priming and heterochromatin dynamics during B cell differentiation. *Nat Commun*. 2015;6:8324.

325. Baubec T, Colombo DF, Wirbelauer C, et al. Genomic profiling of DNA methyltransferases reveals a role for DNMT3B in genic methylation. *Nature*. 2015;520(7546):243-247.

326. Shaffer AL, Wright G, Yang L, et al. A library of gene expression signatures to illuminate normal and pathological lymphoid biology. *Immunol Rev*. 2006;210:67-85.

327. Skene PJ, Henikoff S. A simple method for generating high-resolution maps of genome-wide protein binding. *Elife*. 2015;4:e09225.

328. Skene PJ, Henikoff JG, Henikoff S. Targeted in situ genome-wide profiling with high efficiency for low cell numbers. *Nat Protoc*. 2018;13(5):1006-1019.

329. Liu N, Hargreaves VV, Zhu Q, et al. Direct Promoter Repression by BCL11A Controls the Fetal to Adult Hemoglobin Switch. *Cell*. 2018;173(2):430-442 e417.

330. Meers MP, Bryson TD, Henikoff JG, Henikoff S. Improved CUT&RUN chromatin profiling tools. *Elife*. 2019;8.

Abbreviations

GC: germinal center

KO: knockout

WT: wild type

CSR: class switch recombination

PC: plasma cell

Bmem: memory B cell

HSC: haematopoietic stem cells

GSEA: gene set enrichment analysis

Ig: immunoglobulin

IgH: immunoglobulin heavy chain

IgL: immunoglobulin light chain

pre-BCR: pre-B cell receptor

BCR: B cell receptor

PORE: Palindromic Oct factor Recognition Element

MORE: More palindromic Oct factor Recognition Element

TF: transcription factor

MPP: multipotent progenitor

CMP: common myeloid progenitor

CLP: common lymphoid progenitor

DLBCL: diffuse large B cell lymphoma

DZ: dark zone

LZ: light zone

MZ: marginal zone

SHM: somatic hypermutation

T_{FH}: T follicular helper cell

PWM: positional weight matrices

HTH: helix-turn-helix

CTD: carboxy-terminal domain

EMSA: electrophoretic mobility shift assays

Acknowledgements

Acknowledgements

First and foremost, I would like to acknowledge my family for their abiding support of me during my doctoral study and my dissertation.

To my supervisor, Dr. Patrick Matthias, I would like to express my deep gratitude to him for giving me this chance to perform my thesis at FMI. During my study, his attentive supports and guidance helped push research forward and facilitated my personal and scientific growth. I hope that he might be as proud to be my mentor as I am to be his mentee.

I would like to thank my thesis committee members, Dr. med. Daniela Finke and Dr. Dirk Schübeler, for their help and scientific advices during our committee meetings.

Thanks also to Michael Stadler for his supervision and supports on the bioinformatics part of this thesis. His professional insights, stylish programming skills and creative solutions tremendously stimulated my interest and expedited my ability in doing computational analysis independently.

I want to thank my former colleague in the lab, Mohamed-Amin Choukrallah, particularly for showing me the way to start my training in bioinformatics.

I want to thank Chun Cao, who kindly taught me practical mouse handling techniques and helped me maintain mouse lines.

I also want to thank the current lab members, including Longlong Wang, Gabriele Matthias, Nothjunge Stephan, Wairkar Mrunali, Sanchez Jacint, Langousis Gerasimos, Aguiar Moreira Etori, and Ghosh Sucheta for stimulating discussions and a good atmosphere in the lab.

I also would like to thank Dr.Meinrad Busslinger for sharing with us the transgenic *Rosa26*$^{BirA/BirA}$ mice, Dr. Stephen Nutt, Simon N. Willis and Jacob T. Jackson for sharing with us the *Spi1*$^{-/-}$ mice and culturing *Spi1*$^{-/-}$ B cells for us, Dr. Steffen Junker for sharing with us Ramos and Daudi cell lines, Dr, Jeffrey Chao for sharing with us lentiviral expression plasmids, Dr. Barna Fodor, Mathias Frederiksen and Fabian Wu for sharing their professional experience in CUT&RUN assays with us, and Dr. Gaël Aurey for helping us sort human GC B cells.

Thanks also go to the Novartis Research Foundation for financial support, Hubertus Kohler for FACS analysis, Michael Stadler and Hans-Rudolf Hotz for helping us perform computational analysis, Sebastien Smallwood and the entire functional genomics facility for performing

sequencing services, Jean-François Spetz and the entire transgenic mice facility for generation of CRISPR/Cas9 edited mice and embryo mouse work.

Curriculum Vitae

Shuang Song

Friedrich Miescher Institute for Biomedical Research (www.fmi.ch)

Maulbeerstrasse 66, CH-4058 Basel, Switzerland

Email: shuang.song@fmi.ch

PERSONAL

Citizenship: P. R. China
Language Chinese native, English fluent, German basic
Civil status: Married
Home address: Bleichestrasse 3, 4058 Basel, Switzerland

EDUCATION

Friedrich Miescher Institute for Biomedical Research (FMI), Basel, Switzerland

2014 - present — Ph.D. *summa cum laude* — *Genetics*
Supervisor: Prof. Patrick Matthias

Biopharmaceutical R&D Center
Peking Union Medical College, Tsinghua University, Suzhou, China

2010 - 2013 withdrew — Master - Ph. D. program — *Biomedical Engineering*
Supervisor: Prof. Yongping Jiang

Nanjing Agricultural University, Nanjing, China

2006 - 2010 — B.Sc. — *Biotechnology*

EMPLOYMENT

Friedrich Miescher Institute for Biomedical Research (FMI), Basel, Switzerland

2014 - present — Ph.D. thesis research (Supervisor: Prof. Patrick Matthias)

Matthias Lab: **transcriptional control of OCT factors and OBF1 in germinal center formation**

The OCT transcription factors and coactivator OBF1 are critical for germinal center (GC) formation. I performed genomic analyses of these factors in primary murine mature B cells, murine and human GC B cells, as well as B lymphoma cells. I demonstrated that OBF1 binds to genes important for GCs, maintains proliferation of GC-derived B lymphoma cells, and is a master regulator controlling the GC program.

Biopharmaceutical R&D Center
Peking Union Medical College, Tsinghua University, Suzhou, China

2010 - 2013 withdrew — Research associate (Supervisor: Prof. Yongping Jiang)

Jiang Lab: **large-scale fermentation, refolding and purification of endostatin**

Endostatin is a broad-spectrum angiogenesis inhibitor that enhances the efficacy of conventional chemotherapy for several types of tumors. However, the production of this protein is difficult. I tried various parameters for large-scale production of endostatin by *E.coli* using bioreactor, and optimized the refolding conditions of endostatin inclusion body. Eventually, I achieved over 40% refolding rate for this protein.

Nanjing Agricultural University, Nanjing, China

2007 - 2010 Research assistant (Supervisor: Prof. Jiandong Jiang)

Jiang Lab: **biodegradation of halogenated pesticides and herbicides**

I was invovled in the projects to isolate pesticide-transforming bacteria and understand the enzymatical mechanisms of biogradation of DDT and chlorothalonil. I was trained to perform routine biochemistry, molecular cloning techniques, as well as bacteria isolation and metabolites detection techniques.

Awards

2010	The Honorary Title of Outstanding Student of Nanjing Agricultural University
2009	Third Prize for Band C in 2009 National English Contest for College Students
2008-2009	Scholarships for oustanding individuals
2007-2008	English Excellence Award
2006-2008	Scholarships for National Excellent Student

PUBLICATIONS

1. **Song S**, Cao C, Choukrallah M-A, Tang FY, Christofori G, Kohler H, Wu F, Fodor BD, Frederiksen M, Willis SN, Jackson JT, Nutt SL, Dirnhofer S, Stadler MB, Matthias P (2020). OBF1 and Oct factors control the germinal center transcriptional program. *Blood* doi: 10.1182/blood.2020010175
Commentary by R. Küppers

2. Tang FY, Gao RZ, Jeevan-Raj B, Wyss CB, Kalathur RKR, Piscuoglio S, Ng CKY, Hindupur SK, Nuciforo S, Dazert E, Bock T, **Song S**, Buechel D, Morini MF, Hergovich A, Matthias P, Lim D-S, Terracciano LM, Heim MH, Hall MN, Christofori G (2019). LATS1 but not LATS2 represses autophagy by a kinase-independent scaffold function. *Nature Communications* doi: 10.1038/s41467-019-13591-7

3. **S Song**, P Matthias (2018). The Transcriptional Regulation of Germinal Center Formation. *Frontiers in Immunology* doi: 10.3389/fimmu.2018.02026

4. Choukrallah M-A, **Song S**, Rolink AG, Burger L, Matthias P (2015). Enhancer repertoires are reshaped independently of early priming and heterochromatin dynamics during B cell differentiation. *Nature Communications* doi: 10.1038/ncomms9324

5. **Song S**, Jiang WH, Jiang YP (2012). Large-Scale production,refolding and purification of a novel recombinant human granulocyte colony stimulating factor. *Chinese Journal of Biomedical Engineering* doi: CNKI:SUN:ZSWY.0.2012-04-013