

The function of E2A in B-cell development

Masaki Miyazaki¹ and Kazuko Miyazaki¹

¹ Laboratory of Immunology, Institute for Life and Medical Sciences, Kyoto University,
Kyoto 606-8507, Japan

Corresponding author: Masaki Miyazaki, mmiyazaki@infront.kyoto-u.ac.jp ::

Phone+81-75-751-3818;; Fax +81-75-751-3839

Running head: E2A orchestrates enhancer landscape for B-cell development.

Abstract

Helix-loop-helix (bHLH) transcription factors (TFs) plays a key role in various cellular differentiation and function through the regulation of enhancer activity. E2A, a member of the mammalian E-protein family (class I HLH protein), is well known to play an important role in hematopoiesis, especially in adaptive lymphocyte development. E2A instructs B- and T-cell lineage development through the regulation of enhancer activity for B- or T-cell signature gene expression, including *Rag1* and *Rag2* (*Rag1/2*) genes. In

this chapter, we mainly focus on the function of E2A in B-cell development and on the roles of E2A in establishing the enhancer landscape through the recruitment of EP300/KAT3B, chromatin remodeling complex, mediator, cohesion, and TET proteins. Finally, we demonstrate how E2A orchestrates the assembly of the *Rag1/2* gene super-enhancer (SE) formation by changing the chromatin conformation across the *Rag* gene locus.

Keywords: E2A, B-cell development, basic-helix-loop-helix transcription factor, enhancer regulation, adaptive lymphocyte, Rag1 and Rag2 expression

Introduction

Immune responses are mediated by adaptive and innate immune cells, and they cooperatively evoke immune responses aimed at protecting against pathogen invasion.

Innate immune cells (macrophages and neutrophils) are activated by pattern recognition receptors (PRRs) that recognize common microbial components as danger signals. In contrast, adaptive immune cells, such as B-cells and T-cells, are stimulated through antigen-specific diverse receptors, which are generated by somatic recombination involving variable (V), diversity (D), and joining (J) gene segments of immunoglobulin (*Ig*) and T-cell receptor (*TCR*) genes. V(D)J recombination of *Ig* and *TCR* genes is a characteristic feature of adaptive lymphocytes that is mediated by the recombination-activating gene 1 (RAG1) and RAG2 protein complex (*1-3*). The RAG1 and RAG2 complex recognize and cleave the recombination signal sequence (RSS) flanking the V, D, and J gene segments of the *Ig* and *TCR* gene loci. Following B-cell lineage commitment, D-J recombination of the Ig heavy (*Igh*) chain gene is initiated, and then V-DJ recombination occurs in B-cell progenitors (pro-B-cells) (*4*). When the Ig heavy chain is paired with the surrogate light chain ($\lambda 5$ and VpreB), they form a pre-B-cell receptor (pre-BCR). Pre-BCR signaling induces cell expansion accompanied by IL7 signaling and differentiation into B-cell precursors (pre-B-cells) (*5*). If the first Ig heavy chain gene

recombination fails, V-DJ recombination of the second allele of the *Igh* gene is attempted.

In pre-B-cells, V-J rearrangement of the Ig light (*Igl*) chain gene occurs to generate a pair of IgH and IgL for B-cell receptors (IgM-expressing immature B cells) (6). Similarly, TCR β and/or TCR γ/δ V(D)J recombination is initiated in immature CD4⁻CD8⁻ (double negative; DN) thymocytes. Following the success of productive TCR β gene recombination in DN cells, they differentiate into CD4⁺CD8⁺ (double positive; DP) thymocytes and start V-J recombination of *TCR α* genes. DP cells that successfully express the TCR α/β complex undergo positive and negative selection to eliminate harmful T-cells (7). During these *Igh* and *Igl* gene recombination events, RAG1 binds and cleaves DNA [introduction of double-strand DNA breaks (DSBs) at RSSs]. RAG2 contains a plant homeodomain (PHD) finger in its C-terminus and recognizes hypermethylated histone H3K4 modification through this PHD, which enhances RAG1 binding and catalytic activity (3). In both developing T- and B-cells, the RAG1 and RAG2 protein complex binds in a highly focal manner to a small region of D-J gene segments of the *Igh* and *TCR β* genes and J gene segments of *Ig κ* and *TCR α* , referred to as recombination centers. The recombination center coordinates V(D)J recombination by providing specific sites for capturing V-D-J gene segments (8). Because RAG-mediated DNA cleavage can potentially lead to oncogenic translation and since a subset of leukemic cells express *Rag*

genes (**9-11**), the expression of the *Rag1* and *Rag2* genes is limited and strictly controlled in a highly cell type-specific manner (**6**). Only developing B- and T-cells [pro-B and pre-B-cells in the BM, pro-T (DN3) and pre-T (DP) cells in the thymus] express RAG1/2. Hence, *Rag1/2* gene expression typically reflects the cell lineage commitment toward adaptive immune cells, and the regulatory mechanisms of *Rag1/2* gene expression should be controlled by TFs essential for B- and T-cell development (**12**).

During hematopoiesis in adults, hematopoietic stem /progenitor cells (HSPCs) in the BM have the ability to both reconstitute all hematopoietic cells and self-renew to maintain the hematopoietic system (**13**). HSPCs can give rise to multipotent progenitors (MPPs) and then differentiate into lymphoid-primed multipotent progenitors (LMPPs) (**14**), which then differentiate into common lymphoid progenitors (CLPs) or granulocyte-macrophage progenitors (GMPs) (**15, 16**). CLPs give rise to all lymphocytes, including B-cells, T-cells, natural killer (NK) cells, innate lymphoid cells (ILCs) and dendritic cells (DCs), while GMPs are restricted to myeloid cell pathways, such as granulocytes and macrophages. CLPs are defined as lineage marker-negative (Lin^-) $\text{Flt3}^{\text{high}}$ IL7 receptor (IL7R) $^{\text{high}}$ $\text{Sca1}^{\text{low or -}}$ $\text{KIT}^{\text{low or -}}$ cells by flow cytometric analysis. The restriction of lymphoid versus myeloid cell lineage choice during hematopoiesis in vivo was clarified by *IL7R*^{Cre}-mediated fate mapping and further high-resolution endogenous DNA Polylox barcoding

(17-20). As mentioned, CLPs include heterogeneous populations of each lymphoid lineage cell and are subdivided into Ly6D-positive (B-primed lymphoid progenitors: BLPs) and -negative CLPs, and BLPs mainly generate B lineage cells after transplantation, although they do not express the B lineage marker CD19 (21, 22). More recently, among Ly6D-expressing BLPs, GFRA2 and BST1 expression define a developmental trajectory to pro-B-cells in CLPs, and BST1+ BLPs restrict other lymphoid lineages, including T-cell potential (23). During early B-cell development, a gene regulatory network (GRN) of TFs, such as IKAROS, E2A, EBF1, PAX5, and FOXO1, which are essential for B-cell development, orchestrates the B-cell lineage commitment and the expression of lineage-specific genes, including *Rag1/2*, pre-BCR/BCR-associated molecules (*I κ B* and VpreB, *I α* , and *I γ*), and prevents hematopoietic malignancies (24).

E-proteins are basic helix-loop-helix (bHLH) TFs involved in the regulation of multiple hematopoietic development and include E2A (*Tcf3* gene; E12 and E47 protein), HEB (*Tcf12* gene), and E2-2 (*Tcf4* gene) (25). E-proteins form either homodimers or heterodimers with other E-proteins or HLH proteins via the HLH dimerization domain and bind to the E-box motif (CANNTG) within the cis-regulatory element (CRE, enhancer region) of the target genes through the basic DNA binding region. In contrast,

ID-proteins (ID1-ID4) contain an HLH dimerization domain but lack the basic region for DNA binding. Therefore, when ID-proteins form heterodimers with E-proteins, they antagonize the DNA binding of E-proteins and thereby function as negative regulators of E-proteins (**Figure 1A**) (26). It is well established that the transcriptional balance between E- and ID-proteins (the E-ID axis) controls developmental trajectories of lymphoid lineages and their activation, especially T-cells, B-cells, and DCs (**Figure 1B**) (7, 24, 27, 28). The E2A gene encodes the E47 and E12 proteins (splicing isoform of the E2A gene), and the E47 protein is essential for B-cell development (29). E2A also acts in pro-T-cells along with HEB to activate T-cell signature gene expression and simultaneously suppress ILC development in the thymus (7, 30-33). E2A and HEB also regulate the positive selection of DP cells in the thymus (34), and HEB is also needed for Invariant natural killer-T (iNKT) cell development (35, 36). In addition, E2-2 controls type I interferon-producing plasmacytoid DCs (pDCs) and is essential for the maintenance of pDC identity, while ID2 regulates antigen-presenting classical DC (cDC) development by neutralizing E2-2 activity (37-40). ID2 is also well known as a critical regulator of the development of all ILC subsets, including ILC1-3, NK cell, and lymphoid tissue inducer (LTi) cells (41-43). In this chapter, we will focus on the roles of E2A in B-cell development, molecular mechanisms that underpin enhancer activity, and the regulation of *Rag1/Rag2*

gene expression.

E2A functions in B-cell lineage commitment and development

Discovery of E2A (E47 and E12) and E2A knockout mouse

E12 and E47 were first identified as specific binding factors to Ig kappa chain enhancer E2 (κ E2) containing an E-box, and they were found to form heterodimers with other HLH proteins, such as MYOD or Achaete-scute T3 (AS-C T3) by electrophoretic mobility shift assay (EMSA) (**44, 45**). E12 was clone number #12 and E47 was clone #47, and further investigation clarified that E12 and E47 were encoded in the E2A gene and generated by differential splicing of exons encoding the DNA binding domain (**46**). Indeed, E12 and E47 were the first E-proteins identified. Interestingly, it was found that the splicing factors HNRNPH1 and HNRNPF (HNRNPH/F) regulate E2A splicing in embryonic stem cells (ESCs), and a high level of HNRNPH/F leads to higher E12 expression, and decreasing HNRNPH/P during differentiation causes switching to E47 expression (alternative splicing switch of E2A) (**47**). Following the discovery of the *E2A* gene, an E2A-deficient mice were generated that failed to generate B-cells and lacked *Ig* DJ rearrangements, accompanied by severe impairments in *Rag1*, *Mb1*, *CD19* and *Pax5* expression (**48, 49**). Interestingly, B-cell development in *E47*-ablated mice was severely blocked at the very early stage of B-cell development (prepro-B-cell) as well as *E2A*-deficient mice (**29**),

while *E12*-deficient mice did not show any defects in B-cell development and maintenance (29, 50). In contrast, E12 is needed to promote receptor editing of *Igl* gene rearrangement upon exposure to self-antigen, and E12 and E47 act to sequentially regulate *Ig* light chain gene rearrangement (50).

E2A expression and function during B-cell development

What is the expression profile of E2A during adaptive lymphocyte (B- and T-cell) and innate lymphoid cell (ILC) development? By analyzing the E2A-GFP fusion protein reporter mouse, it appeared that E2A protein levels were already high in HSCs and increased in LMPPs, CLPs, and BLPs and became highest in pro-B-cells, the committed stage of the B-cell lineage. After pro-B-cells, the E2A level gradually declines in immature B-cell populations (pre-B and IgM⁺ immature B-cells) and is very low in mature IgD⁺ B-cells in the BM (33, 51). Interestingly, during ILC development, the E2A level remains mid-to low in the ILC precursor population ($\alpha 4\beta 7^{\text{hi}}\text{CD}25^{-}$), showing an inverse correlation between E2A and ID2 or IL7R expression (33). Similar to B-cell development, the E2A level is high in early T-progenitors (ETPs), highest in committed DN2 thymocytes, and remains high until the resting DP compartment (30). Consistent with these expression patterns of the E2A protein, E2A plays roles in early hematopoiesis.

E2A/E47 is needed for the maintenance of the HSC pool and the transition from HSC to LMPPs and CLPs (52-55). CLPs are a heterogeneous population, and Ly6D-expressing BLPs are primarily B lineage restricted (21). Although E2A-deficient mice showed fewer CLPs (approximately half), BLPs are severely impaired in E2A null mice (21). Consistent with these findings, the deletion of E2A and HEB after competitive BM transplantation resulted in a complete loss of B-cell development in the BM accompanied by severe defects in BLPs, while impairment in LMPPs and CLPs was mild-to-moderate in E2A/HEB double-deficient cells, indicating cell intrinsic roles of E-proteins in B lineage commitment (33). How about other TFs essential for the B-cell development? The deletion of the *Foxo1* gene in HSPCs (*Vav*^{Cre}) resulted in the accumulation of Ly6D⁺ BLPs and defects in CD19⁺ B cells in the BM, whereas *Foxo1* deletion in BLPs/pro-B-cells (*Mb1*^{Cre}) resulted in the accumulation of pro-B-cells accompanied by lower expression of *E2A* and *Ebf1* (56, 57). Similar to *Foxo1*-deficiency, *Ebf1*-deletion led to a developmental block at the transition from BLPs to pro-B-cells (56). Furthermore, E2A and EBF1 regulate FOXO1 expression, and these three TF regulatory networks regulate the enhancer landscape for B-cell lineage commitment (58). Curiously, not only BLPs and pro-B-cells but also DN2 and DN3 thymocytes were severely affected in the absence of E2A and HEB genes without affecting ETPs in the thymus, indicating that ETPs are

able to migrate into the thymus in the absence of E2A and HEB and that E-proteins are essential for not only B-cell but also T-cell lineage commitment (33). This clearly demonstrates that E2A plays a central role in adaptive lymphocyte lineage commitment. Interestingly, HEB could functionally replace the E2A gene in supporting B-cell development when HEB was driven by the endogenous E2A promoter, indicating that divergence among E-proteins in cell type-specific function is partially defined by their gene expression (59). From these observations, it is suggested that gene duplication from ancestral genes may have occurred in part to permit genes to be placed under novel regulatory elements (46). Indeed, a reconstructed maximum likelihood phylogenetic tree of three E-protein homologs for E2A, E2-2, and HEB shows three clades for these in jawed vertebrates while lamprey E-protein homologs are distinct from those three in jawed vertebrates (60). From this observation, it was suggested that these paralogs were generated through the two widely recognized rounds of whole-genome duplication (WGD) in vertebrates (61). Hence, it is plausible that ancestral jawed vertebrates expressed four E-protein paralogs and that one paralog was lost early in evolution prior to the divergence of jawed vertebrates (60).

The molecular function of E2A in B-cell development

In addition to the bHLH domain, E-proteins contain three highly conserved activation domains (AD1-3) (62-64). AD1 is located at the N-terminus of E2A, AD3 and AD2 are in the middle of the E2A protein, and the bHLH DNA binding and dimerization domain is located at the C-terminus (Figure 1A) (64). AD1 of E2A was first identified to activate target gene transcription through the recruitment of histone acetyltransferases (HATs) EP300/KAT3B and CBP/KAT3A, and AD1 also contributes to binding to the transcriptional repressor ETO protein to suppress target gene expression (62, 65, 66). In addition, AD2 also interacts with P300/KAT3B and CBP/KAT3A or ETO2 cooperatively with AD1 to promote or repress the transcriptional activity of target genes (67-69). It was reported that E-proteins enhance the recruitment of TAF3/TFIID to the promoter through AD3 binding, thereby promoting target gene transcription (70).

A genome-wide analysis of TF occupancies and histone modification has advanced our understanding of transcriptional regulation in B-cell lineage commitment (58). During the transition from prepro-B-cells (*Ebfl*-deficient cells) to pro-B-cells (*Ragl*-deficient cells), E2A occupancies are dynamically increased during developmental progression, and E2A occupancies in pro-B-cells are associated with monomethylation of histone H3K4 (H3K4me1) across putative enhancer regions. In addition, E2A binding is associated with RUNX1 and ETS motifs in prepro-B-cells, while E2A binding across the

putative enhancers specific to pro-B-cells (cluster II and III) is associated with EBF1 motifs. Interestingly, E2A and EBF1 bound to the putative enhancer regions of B-cell-specific genes, such as *Cd19*, *Cd79a* (*Mb1*), *Cd79b*, *Pax5*, *Vpreb1/2/3*, and *Cxcr4*, and they also bound to the *Foxo1* locus to induce its expression. From these genome-wide analyses of TFs it was demonstrated a global network of TFs, BCR signaling and survival factors that orchestrates B-cell specificity (58). Subsequently, they further analyzed the whole-genome chromatin structure upon B-cell lineage commitment by using Hi-C methods (71). By analyzing Hi-C data, it was shown that genomes are organized into two compartments: transcriptionally repressive and permissive chromatin compartments defined by the principal component 1 (PC1) value (72). Notably, the *Ebfl* gene locus repositioned from the lamina to the nuclear interior during the transition from prepro-B-cells to pro-B-cells (71). In this study, E2A bound at the chromatin loop-attachment regions, together with EP300/KAT3B and PU.1, associated with intra- and interdomain interactions that are developmentally regulated. After B lineage commitment, EBF1, E2A, and other TFs switched compartments to establish new chromatin interactions linked to a B-cell signature gene. Another prominent chromatin switched region involved the *Igκ* locus. Specifically, in pro-B-cells the *Igκ* locus showed an intricate and elaborate pattern of interactions of chromatin loops. Interestingly, these long-range chromatin loops

were closely associated with E2A occupancy, suggesting that E2A is involved in tethering V κ segments across vast genomic distance (24, 71). These results strongly support the previous model that E2A directly recruits EP300/KAT3B and CBP/KAT3A HATs to enhancer regions to activate target gene expression. Furthermore, by analyzing whole-genome DNA methylation status, B lineage commitment was accompanied by large-scale changes in not only chromatin structures but also DNA cytosine modification at CpG islands associated with enhancer, promoter, and anchor regions (73). Hypomethylated DNA regions were found exclusively in the permissive chromatin compartment and were associated with active enhancer or anchor regions folding the genome (73). How were the CpG islands in the regulatory regions hypomethylated during B-cell development? Is E2A involved in the process of demethylation? Ten-eleven translocation (TET) enzymes facilitate DNA demethylation, and deletions of TET2 and TET3 in mice at the early B-cell stage using *Mb1*^{Cre} blocked B-cell development at the pro-B stage, accompanied by decreased *Irf4* expression and impaired *Ig κ* recombination (74). TET2/3 double knockout (dKO) pro-B-cells showed increased CpG methylation at the *Ig κ* enhancers (iE κ , 3'E κ , and dE κ) where E2A and PU.1 bind, as well as decreased chromatin accessibility at enhancer regions. Surprisingly, E2A and PU.1 directly interact with TET2, as shown by an immunoprecipitation assay, indicating that lineage-specific TFs such as E2A directly

recruit TET proteins to remove DNA methylation (74). Consistent with this, CpG islands at the *Rag* gene enhancer were hypermethylated when E2A binding was specifically blocked by generating E-box mutations in this enhancer region, indicating that E2A recruits TET protein to the lymphocyte-specific enhancer to eliminate CpG DNA methylation (75). These data suggest that E2A instructs the adaptive lymphoid lineage, especially from multipotent progenitors to the B-cell lineage in the BM. Indeed, it was demonstrated that E2A-deficient BM progenitor cells showed self-renewal activity without losing multipotent differentiation potential *in vitro* (76), and in addition, blocking E2A activity by expressing the E-protein antagonist ID2 in hematopoietic progenitor cells (HPCs) was able to expand these progenitor cells with multipotency and readily differentiate upon withdrawal of ID2 expression into the B-cell lineage (77). Interestingly, a subset of lymphoid and myeloid enhancers was already poised or active enhancer status, indicating multilineage priming of enhancers preceding commitment to the lymphoid or myeloid lineage cells (77). Consistent with this important finding, it was shown that the status of enhancers for lymphoid, myeloid and erythroid lineages were already poised or active in HSCs and MPPs and hematopoiesis progresses through gain and loss of these lineage-specific enhancers, associated with lineage-specific TF regulation (78). Therefore, E2A primarily instructs and specifies the adaptive lymphoid lineage by activating these

poised enhancers for the B-cell lineage and by suppressing myeloid or erythroid enhancers. Indeed, B-cell differentiation in E47-deficient fetal liver progenitor cells can be rescued by forced EBF1 expression but not PAX5 *in vitro*, suggesting that E2A acts upstream of EBF1 (79). How does E2A activate enhancer regions and evolve the enhancer landscape for the adaptive lymphoid lineage? The AD1 helix domain contains a novel four-amino-acid-sequence LDFS motif that is present in class I HLH TF including E2A, HEB and E2-2, and the LDFS motif stimulates transcription by recruiting the SAGA histone acetyltransferase complex (80). Another group called these conserved motifs PCET (EP300/KAT3B, CBP/KAT3A and ETO target) motifs, which are important for interacting with EP300 /KAT3B and ETO (64, 66). As mentioned above, the AD1 domain also recruits acetyltransferases EP300 /KAT3B and CBP/KAT3A (64-66). Histone H3K27 acetylation (H3K27ac) is recognized by bromodomain and extraterminal (BET) histone reader protein, and the bromodomain protein BRG1 complex binds to the E2A-occupied enhancer region to evict nucleosomes to provide chromatin accessibility to other B-cell-specific TFs (81). Recently, a prion-like domain in the C-terminal domain (CTD) of EBF1 has been shown to have liquid-liquid phase separation formation ability, which enables the recruitment of the RNA-binding protein FUS and BRG1 (82). In addition to EBF1, intrinsically disordered protein regions (IDRs) are highly frequently found to be

enriched in E2A, which also contribute to the formation of phase separation. Collectively, E2A and EBF1 cooperatively facilitate chromatin opening through the recruitment of BRG1 and act as pioneer factors to initiate phase-separated SE formation to promote B-cell signature gene expression. To instruct SE formation for B-cell lineage commitment, large-scale chromatin conformation changes are needed. Since blocking E2A binding to the *Rag* enhancer led to a failure of SE formation accompanied by a loss of cohesion complex recruitment across the *Rag* gene locus (75), E2A seems to directly orchestrate chromatin conformation changes (**Figure 2**). According to the integrative analysis of genome-wide Hi-C and ChIP-Seq data in pro-B-cells, E2A occupancy is closely associated with EP300/KAT3B, PU.1, EBF1, CTCF and cohesin complex component (RAD21) at chromatin loop attachment regions (anchor region) (71). In particular, switching nuclear compartments between multipotent progenitors and pro-B-cells, which include the *Ebfl* gene locus, establishes new intra- and interdomain interactions for the B-cell signature gene expression program (46, 71). Notably, while prepro-B-cells mainly exhibited adjacent chromatin looping across the *Igκ* gene locus, pro-B-cells showed an intricate and elaborating chromatin interactions, and many interactions between *Igκ* intronic enhancers and genomic elements were closely associated with E2A occupancy, consistent with the initial discovery of E2A occupancy at the *Igκ* enhancer binding (71).

The regulation of Rag1/2 expression and E2A function.

B- and T-cell-specific Rag gene enhancers.

As mentioned in the introduction, RAG-mediated V(D)J recombination of Ig and TCR genes is a characteristic feature of adaptive lymphocytes, and the regulation of *Rag1/2* genes is a hallmark of adaptive lymphoid lineage commitment (7). *Rag* gene expression is restricted to developing B- and T-cells, and there are two waves of *Rag* expression (6, 83). The first wave of *Rag1/2* expression initiates D-J and V-DJ recombination of the *Igh* gene in pro-B-cells. *Rag1/2* gene expression is transiently downregulated after successful *Igh* gene rearrangement during the developmental transition from pro-B- to pre-B-cells. In pre-B-cells, *Rag1/2* genes are re-expressed, leading to immunoglobulin light chain (IgL) recombination. Following successful recombination of IgL and assembly of Ig, *Rag1/2* expression is repressed. Since persistent *Rag1/2* expression results in aberrant thymic development and profound immunodeficiency (84), *Rag1/2* gene expression is stringently controlled in B- and T-cell development (6). Whereas both developing B- and T-cells express *Rag1/2* genes, they use distinct cell type-specific enhancers for their expression. It was reported that the B-cell-specific enhancer (enhancer of *Rag*; *Erag*) is 23 kb upstream of the *Rag2* gene promoter and that the deletion of *Erag* resulted in a

reduction in *Rag1/2* gene expression and a moderate developmental block at pro-B stages without affecting T-cell development (85). It was shown that E2A bound to the *Erag* enhancer in pro-B-cells and was involved in *Erag* enhancer activity in nonlymphoid cells (85). Following the identification of *Erag*, it has been reported that FOXO1 positively regulates the *Erag* enhancer and that GFI1B and c-MYB negatively regulate the *Erag* region in cell lines (57, 86-88). Regarding the T-cell-specific enhancer, an anti-silencer element (*ASE*), which is located 73 kb upstream of the *Rag2* promoter and 8 kb in length, was identified (89). The deletion of *ASE* in mice showed a reduction in *Rag1/2* expression in DN3 and DP thymocytes but not in developing B-cells in the BM (89). ASE-deficient mice showed a defect in the positive selection of DP cells but not in the b-selection of DN3 cells (89). Since *Erag* deletion resulted in moderate impairment of B-cell development, additional enhancer or other regulatory mechanisms need to be determined (6). In addition, it remains unclear how these enhancers regulate *Rag1* and *Rag2* gene loci in coordination and how their expression is restrained in innate immune cells. Exploring these molecular mechanisms surely shed light on the question of how lymphoid progenitors commit to the adaptive lymphocyte lineage.

By analyzing the E2A occupancies and ATAC-Seq data of DN3 and pro-B-cells in the *Rag1/2* gene loci, two B-cell-specific *Rag* gene enhancers (*Rag* B-cell enhancer 1/2: *R1B*

and *R2B*) and one T-cell-specific *Rag* gene enhancer [*Rag*-T-cell enhancer (*R-TEn*)] were identified. Furthermore, one common E2A binding region adjacent to the *Rag1* gene promoter was also identified (*Rag1* promoter: *R1pro*) (75). Indeed, *R-TEn* is included in the 8 kb *ASE*, *R1B* is located 5 kb upstream of the *Rag1* promoter, and *R2B* partially overlaps with the *Erag* enhancer (**Figure 3A**). *R1B/R2B* and *R-TEn* directly loop uniquely to the *Rag1/2* gene promoter regions and form distinct genome structures between B- and T-cells. B- or T-cell-specific TFs [PAX5, ETS1, IKAROS, and IRF4 (B-cell TFs), and BCL11B, TCF1, RUNX1, IKAROS, and GATA3 (T-cell TFs)] bind to these enhancer regions in a cell type-specific manner. The *R-TEn*, *R2B* and *R1pro* regions are SEs accompanied by elaborate chromatin interactions. While the deletion of either *R1B* or *R2B* in mice showed mild to moderate impairments in B-cell development, *R1B/R2B* double deletion in mice led to a severe developmental block at the pro-B-cell stage, indicating redundancy among these two B-cell enhancers (12, 75). As well as *ASE*-deficient mice, deletion of *R-TEn* in mice resulted in defects in the positive selection of DP cells and the b-selection of DN3 cells. The developmental blocks in *R1B/R2B* or *R-TEn* deletion mice were caused by a failure of *Rag1/2* expression and impaired *TCR* and *IgH* gene recombination (75).

E2A/E-proteins regulate Rag gene enhancer and Rag1 promoter activities

What TF regulates the *Rag* enhancer? While *Rag1/2* gene expression and chromatin accessibility of *R-TEn* were not affected in *Tcf1*- or *Bcl11b*-deficient thymocytes, the deletion of E2A and HEB in DP cells led to a significant downregulation of *Rag1/2* expression, accompanied by a loss of chromatin accessibility in the *R-TEn* and *Rag1* promoter regions, indicating that E2A regulates and maintains *Rag* gene enhancer activity.

In addition, it was reported that the chromatin organizer SATB1 bound to the *ASE* and *Rag1* promoter to facilitate inclusion of the *Rag1* and *Rag2* gene loci in the active chromatin hub (90). They also reported GATA3 and E2A as critical regulators of *ASE* (91). Finally, blocking the E2A/E-protein binding to this enhancer by generating mice with E-box mutations in the *R-TEn* region (*R-TEn-p1-E-box^{mut/mut}*) resulted in impaired *Rag1/2* expression and developmental block at the positive selection of DP thymocytes, as seen in the entire enhancer deletion mice (*R-TEn^{del/del}*), indicating that the E2A/E-protein regulates this cell type-specific *Rag1/2* enhancer (75). Furthermore, blocking E2A binding to this enhancer led to the complete loss of chromatin accessibility and chromatin interactions across the entire *Rag1/2* gene locus accompanied by a loss of cohesin complex recruitment. This suggests that the E2A orchestrates cohesin-mediated looping and SE formation, collaborating with SATB1 (chromatin hub) to activate *Rag* gene

expression (12, 90). In addition to the *Rag* enhancer, E2A commonly binds to the *R1pro* region, which is located in close proximity from the *Rag1* promoter. In early studies it was demonstrated that RUNX1 and E2A regulate *Rag1*-promoter activity by using a DP cell line (91). E-box mutation in *R1pro* (*R1pro-E-box^{mut/mut}*) resulted in a severe developmental block at both b-selection of thymocytes and pre-BCR selection of pro-B-cells, caused by defects in V(D)J recombination of *TCR β* and *IgH* genes, similar to *Rag1*- or *Rag2*-deficient mice. Surprisingly, *Rag2* expression in DN3 and pro-B-cells from *R1pro-E-box^{mut/mut}* mice is normal, and the chromatin accessibilities of the *R1B/R2B* and *R-TEn* regions are not affected by a loss of *Rag1* promoter activity, indicating that cell type-specific enhancers and *Rag2* expression are independently regulated by *Rag1* promoter activity (75). This outcome clearly demonstrates that E-proteins regulate *Rag1* promoter activity in the adaptive lymphoid lineage. However, *Rag2* promoter activity seems to be more complicatedly regulated.

The *Rag2* promoter has two CREs (D3 and Ep) to which C/EBP and NFkB bind to activate the *Rag2* promoter (92, 93). In addition, GATA3, PAX5, c-MYB, and LEF1 bind to the *Rag2* promoter region (94, 95). However, the regulation of *Rag2* promoter activity and responsible TFs in vivo remains to be determined. Finally, how is *Rag1/2* expression suppressed at positive selection? It was reported that signal-dependent suppression of

Rag1/2 gene expression and the disassembly of *Rag1* enhancer-promoter interactions depend on IKAROS. Blocking IKAROS binding to the *Rag1* promoter by mutation (*Rag1p* mIkaros) or *Ikaros* deletion (*Ikzf1* KO) in VL3-3M2 DP cell lines caused upregulation of *Rag1* gene expression, and the suppression of *Rag1/2* expression mediated by strong signals (PMA and ionomycin) was substantially impaired in *Rag1p* mIkaros and *Ikzf1* KO cells, accompanied by a loss of disassembly of enhancer-promoter interactions (91). However, the suppression of *Rag1/2* gene expression upon mature BCR signaling remains undetermined.

Taken together, upon lineage commitment to the adaptive lymphoid lineage from multipotent progenitors, E2A binds to cell type-specific *Rag* gene enhancers (*R1B/R2B* and *R-TEn*) and the *R1pro* region independently and then recruits CBP/KAT3A and EP300/KAT3B, which induces a high level of histone H3K27ac across enhancer regions. The chromatin remodeling complex BAF, including BRG1 and BRM, recognizes acetylated histones through the bromodomain (BRD) and evicts nucleosomes across these enhancer regions (96). E2A directly recruits TET proteins to remove CpG DNA methylation across the enhancer regions (74). In fact, hypomethylation of CpG islands across the *R1B*, *R2B*, and *R-TEn* enhancers was observed in pro-B/pre-B and DN3/DP cells, respectively, and blocking E2A binding to *R-TEn* resulted in a failure to remove

CpG island methylation across the *R-TEn* region, indicating that E2A is needed for the recruitment of TET proteins to the cell type-specific enhancer regions (75). Since blocking E2A binding to *R-TEn* led to a loss of chromatin interactions across the entire *Rag* gene locus accompanied by a failure of cohesin complex recruitment, E2A instructs SE formation by recruiting the cohesin complex (12, 75). In contrast, in innate immune cells such as macrophages, there is no topologically-associated domain (TAD) formation across the *Rag* gene locus, and *R1B* seems to function as an insulator to sequester the *Rag1/2* gene away from the active neighboring *Traf6* gene to prevent aberrant *Rag1/2* gene expression (Figure 3B) (7, 12).

In summary, E2A instructs adaptive immunity through orchestrating B- or T-cell-specific enhancer landscapes cooperatively with other E-proteins, B- or T-cell-specific TFs, transcription mediators, chromatin remodeling factors, and TET proteins, by which E-protein/E2A simultaneously suppresses the innate immune cell lineage.

Acknowledgment

We deeply thank C. Murre for reviewing the manuscript and for kind suggestions.

Funding: This work was supported by the KAKENHI (23H02734 to M.M.) and Japan

Agency for Medical Research and Development (AMED) (22gm6110030h0004 to M.M.)

and Japan Science and Technology Agency (JST) (JPMJPR2388 to M.M.).

Figure legends

Figure 1. (A) Mammalian E-protein family and ID-proteins.

(B) The role of E-proteins and ID-proteins in hematopoiesis.

Figure 2. E2A organizes the chromatin conformation and super-enhancer (SE) formation to induce the B-cell signature gene expression.

Figure 3. The *Rag* gene locus and its cis-regulatory elements. (A) A schematic diagram of the *Rag1* and *Rag2* gene locus and B-cell-specific (*R1B* and *R2B*) and T-cell-specific (*R-TEn*) enhancer regions. (B) Regulation of B- and T-cell-specific *Rag* gene enhancers and *Rag* gene locus mediated by E2A/E-protein.

References

1. Schatz DG, Oettinger MA, Baltimore D. The V(D)J recombination activating gene, RAG-1. *Cell*. 1989;59(6):1035-48.
2. 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):1517-23.
3. Schatz DG, Ji Y. Recombination centers and the orchestration of V(D)J recombination. *Nat Rev Immunol*. 2011;11(4):251-63.
4. Vettermann C, Schlissel MS. Allelic exclusion of immunoglobulin genes: models and mechanisms. *Immunol Rev*. 2010;237(1):22-42. doi: 10.1111/j.1600-065X.2010.00935.x.
5. Mandel EM, Grosschedl R. Transcription control of early B cell differentiation. *Curr Opin Immunol*. 2010;22(2):161-7. doi: 10.1016/j.coi.2010.01.010.
6. Kuo TC, Schlissel MS. Mechanisms controlling expression of the RAG locus during lymphocyte development. *Curr Opin Immunol*. 2009;21(2):173-8. doi: 10.1016/j.coi.2009.03.008.
7. Hidaka R, Miyazaki K, Miyazaki M. The E-Id Axis Instructs Adaptive Versus Innate Lineage Cell Fate Choice and Instructs Regulatory T Cell Differentiation.

Front Immunol. 2022;13:890056. doi: 10.3389/fimmu.2022.890056.

8. Ji Y, Resch W, Corbett E, Yamane A, Casellas R, Schatz DG. The in vivo pattern of binding of RAG1 and RAG2 to antigen receptor loci. *Cell*. 2010;141(3):419-31. doi: 10.1016/j.cell.2010.03.010.

9. Umiel T, Pattengale P, Weinberg K. Recombination activating gene-1 (RAG-1) expression in all differentiation stages of B-lineage precursor acute lymphoblastic leukemia. *Leukemia*. 1993;7(3):435-40.

10. Han Q, Ma J, Gu Y, Song H, Kapadia M, Kawasaki YI, et al. RAG1 high expression associated with IKZF1 dysfunction in adult B-cell acute lymphoblastic leukemia. *J Cancer*. 2019;10(16):3842-50. doi: 10.7150/jca.33989.

11. Aref S, Khaled N, Menshaw NE, Sabry M, Agder MA. Clinical value of RAG1 expression and IKZF1 deletions in Philadelphia negative pediatric B cell precursor acute lymphoblastic leukemia. *Pediatr Hematol Oncol*. 2020;37(5):380-9. doi: 10.1080/08880018.2020.1767739.

12. Miyazaki K, Miyazaki M. The Interplay Between Chromatin Architecture and Lineage-Specific Transcription Factors and the Regulation of Rag Gene Expression. *Front Immunol*. 2021;12:659761. doi: 10.3389/fimmu.2021.659761.

13. Spangrude GJ, Heimfeld S, Weissman IL. Purification and characterization of

mouse hematopoietic stem cells. *Science*. 1988;241(4861):58-62. doi: 10.1126/science.2898810.

14. Adolfsson J, Månsson R, Buza-Vidas N, Hultquist A, Liuba K, Jensen CT, et al. Identification of Flt3+ lympho-myeloid stem cells lacking erythro-megakaryocytic potential a revised road map for adult blood lineage commitment. *Cell*. 2005;121(2):295-306. doi: 10.1016/j.cell.2005.02.013.

15. Kondo M, Weissman IL, Akashi K. Identification of clonogenic common lymphoid progenitors in mouse bone marrow. *Cell*. 1997;91(5):661-72. doi: 10.1016/s0092-8674(00)80453-5.

16. Akashi K, Traver D, Miyamoto T, Weissman IL. A clonogenic common myeloid progenitor that gives rise to all myeloid lineages. *Nature*. 2000;404(6774):193-7. doi: 10.1038/35004599.

17. Schlenner SM, Madan V, Busch K, Tietz A, Laufle C, Costa C, et al. Fate mapping reveals separate origins of T cells and myeloid lineages in the thymus. *Immunity*. 2010;32(3):426-36. doi: 10.1016/j.immuni.2010.03.005.

18. Busch K, Klapproth K, Barile M, Flossdorf M, Holland-Letz T, Schlenner SM, et al. Fundamental properties of unperturbed haematopoiesis from stem cells in vivo. *Nature*. 2015;518(7540):542-6. doi: 10.1038/nature14242.

19. Pei W, Feyerabend TB, Rössler J, Wang X, Postrach D, Busch K, et al. Polylox barcoding reveals haematopoietic stem cell fates realized in vivo. *Nature*. 2017;548(7668):456-60. doi: 10.1038/nature23653.
20. Shang F, Rodewald HR. Toward the dissection of hematopoietic stem cell fates and their determinants. *Curr Opin Genet Dev*. 2022;75:101945. doi: 10.1016/j.gde.2022.101945.
21. Inlay MA, Bhattacharya D, Sahoo D, Serwold T, Seita J, Karsunky H, et al. Ly6d marks the earliest stage of B-cell specification and identifies the branchpoint between B-cell and T-cell development. *Genes Dev*. 2009;23(20):2376-81. doi: 10.1101/gad.1836009.
22. Mansson R, Zandi S, Welinder E, Tsapogas P, Sakaguchi N, Bryder D, et al. Single-cell analysis of the common lymphoid progenitor compartment reveals functional and molecular heterogeneity. *Blood*. 2010;115(13):2601-9. doi: 10.1182/blood-2009-08-236398.
23. Jensen CT, Åhsberg J, Sommarin MNE, Strid T, Somasundaram R, Okuyama K, et al. Dissection of progenitor compartments resolves developmental trajectories in B-lymphopoiesis. *J Exp Med*. 2018;215(7):1947-63. doi: 10.1084/jem.20171384.

24. Aubrey M, Warburg ZJ, Murre C. Helix-Loop-Helix Proteins in Adaptive Immune Development. *Front Immunol.* 2022;13:881656. doi: 10.3389/fimmu.2022.881656.
25. Kee BL. E and ID proteins branch out. *Nat Rev Immunol.* 2009;9(3):175-84. doi: 10.1038/nri2507.
26. Murre C. Helix-loop-helix proteins and lymphocyte development. *Nat Immunol.* 2005;6(11):1079-86. doi: 10.1038/ni1260.
27. Lazorchak A, Jones ME, Zhuang Y. New insights into E-protein function in lymphocyte development. *Trends Immunol.* 2005;26(6):334-8. doi: 10.1016/j.it.2005.03.011.
28. Miyazaki K, Miyazaki M, Murre C. The establishment of B versus T cell identity. *Trends in immunology.* 2014;35(5):205-10.
29. Bain G, Robanus Maandag EC, te Riele HP, Feeney AJ, Sheehy A, Schlissel M, et al. Both E12 and E47 allow commitment to the B cell lineage. *Immunity.* 1997;6(2):145-54. doi: 10.1016/s1074-7613(00)80421-5.
30. Miyazaki M, Rivera RR, Miyazaki K, Lin YC, Agata Y, Murre C. The opposing roles of the transcription factor E2A and its antagonist Id3 that orchestrate and enforce the naive fate of T cells. *Nat Immunol.* 2011;12(10):992-1001. doi:

10.1038/ni.2086.

31. Braunstein M, Anderson MK. HEB-deficient T-cell precursors lose T-cell potential and adopt an alternative pathway of differentiation. *Mol Cell Biol.* 2011;31(5):971-82. doi: 10.1128/mcb.01034-10.

32. Braunstein M, Anderson MK. HEB in the spotlight: Transcriptional regulation of T-cell specification, commitment, and developmental plasticity. *Clin Dev Immunol.* 2012;678705(10):22.

33. Miyazaki M, Miyazaki K, Chen K, Jin Y, Turner J, Moore AJ, et al. The E-Id protein axis specifies adaptive lymphoid cell identity and suppresses thymic innate lymphoid cell development. *Immunity.* 2017;46(5):818-34. e4.

34. Jones ME, Zhuang Y. Acquisition of a functional T cell receptor during T lymphocyte development is enforced by HEB and E2A transcription factors. *Immunity.* 2007;27(6):860-70. doi: 10.1016/j.immuni.2007.10.014.

35. D'Cruz LM, Knell J, Fujimoto JK, Goldrath AW. An essential role for the transcription factor HEB in thymocyte survival, Tcra rearrangement and the development of natural killer T cells. *Nat Immunol.* 2010;11(3):240-9. doi: 10.1038/ni.1845.

36. Roy S, Zhuang Y. Orchestration of invariant natural killer T cell development

by E and Id proteins. *Crit Rev Immunol.* 2015;35(1):33-48. doi: 10.1615/critrevimmunol.2015012207.

37. Cisse B, Caton ML, Lehner M, Maeda T, Scheu S, Locksley R, et al. Transcription factor E2-2 is an essential and specific regulator of plasmacytoid dendritic cell development. *Cell.* 2008;135(1):37-48. doi: 10.1016/j.cell.2008.09.016.

38. Ghosh HS, Cisse B, Bunin A, Lewis KL, Reizis B. Continuous expression of the transcription factor e2-2 maintains the cell fate of mature plasmacytoid dendritic cells. *Immunity.* 2010;33(6):905-16. doi: 10.1016/j.immuni.2010.11.023.

39. Grajkowska LT, Ceribelli M, Lau CM, Warren ME, Tiniakou I, Nakandakari Higa S, et al. Isoform-Specific Expression and Feedback Regulation of E Protein TCF4 Control Dendritic Cell Lineage Specification. *Immunity.* 2017;46(1):65-77. doi: 10.1016/j.immuni.2016.11.006.

40. Verykokakis M, Zook EC, Kee BL. ID'ing innate and innate-like lymphoid cells. *Immunol Rev.* 2014;261(1):177-97. doi: 10.1111/imr.12203.

41. Yokota Y, Mansouri A, Mori S, Sugawara S, Adachi S, Nishikawa S, et al. Development of peripheral lymphoid organs and natural killer cells depends on

the helix-loop-helix inhibitor Id2. *Nature*. 1999;397(6721):702-6. doi: 10.1038/17812.

42. Serafini N, Vosshenrich CA, Di Santo JP. Transcriptional regulation of innate lymphoid cell fate. *Nat Rev Immunol*. 2015;15(7):415-28. doi: 10.1038/nri3855.

43. Cherrier DE, Serafini N, Di Santo JP. Innate Lymphoid Cell Development: A T Cell Perspective. *Immunity*. 2018;48(6):1091-103. doi: 10.1016/j.immuni.2018.05.010.

44. Murre C, McCaw PS, Baltimore D. A new DNA binding and dimerization motif in immunoglobulin enhancer binding, daughterless, MyoD, and myc proteins. *Cell*. 1989;56(5):777-83. doi: 10.1016/0092-8674(89)90682-x.

45. Murre C, McCaw PS, Vaessin H, Caudy M, Jan LY, Jan YN, et al. Interactions between heterologous helix-loop-helix proteins generate complexes that bind specifically to a common DNA sequence. *Cell*. 1989;58(3):537-44. doi: 10.1016/0092-8674(89)90434-0.

46. Murre C. Helix-loop-helix proteins and the advent of cellular diversity: 30 years of discovery. *Genes Dev*. 2019;33(1-2):6-25. doi: 10.1101/gad.320663.118.

47. Yamazaki T, Liu L, Lazarev D, Al-Zain A, Fomin V, Yeung PL, et al. TCF3 alternative splicing controlled by hnRNP H/F regulates E-cadherin expression

and hESC pluripotency. *Genes Dev.* 2018;32(17-18):1161-74. doi: 10.1101/gad.316984.118.

48. Bain G, Maandag EC, Izon DJ, Amsen D, Kruisbeek AM, Weintraub BC, et al. E2A proteins are required for proper B cell development and initiation of immunoglobulin gene rearrangements. *Cell.* 1994;79(5):885-92.

49. Zhuang Y, Soriano P, Weintraub H. The helix-loop-helix gene E2A is required for B cell formation. *Cell.* 1994;79(5):875-84.

50. Beck K, Peak MM, Ota T, Nemazee D, Murre C. Distinct roles for E12 and E47 in B cell specification and the sequential rearrangement of immunoglobulin light chain loci. *J Exp Med.* 2009;206(10):2271-84. doi: 10.1084/jem.20090756.

51. Zhuang Y, Jackson A, Pan L, Shen K, Dai M. Regulation of E2A gene expression in B-lymphocyte development. *Mol Immunol.* 2004;40(16):1165-77.

52. Yang Q, Kardava L, St Leger A, Martincic K, Varnum-Finney B, Bernstein ID, et al. E47 controls the developmental integrity and cell cycle quiescence of multipotential hematopoietic progenitors. *J Immunol.* 2008;181(9):5885-94. doi: 10.4049/jimmunol.181.9.5885.

53. Dias S, Månsson R, Gurbuxani S, Sigvardsson M, Kee BL. E2A proteins promote development of lymphoid-primed multipotent progenitors. *Immunity.*

2008;29(2):217-27. doi: 10.1016/j.immuni.2008.05.015.

54. Semerad CL, Mercer EM, Inlay MA, Weissman IL, Murre C. E2A proteins maintain the hematopoietic stem cell pool and promote the maturation of myelolymphoid and myeloerythroid progenitors. *Proc Natl Acad Sci U S A*. 2009;106(6):1930-5. doi: 10.1073/pnas.0808866106.

55. Borghesi L, Aites J, Nelson S, Lefterov P, James P, Gerstein R. E47 is required for V(D)J recombinase activity in common lymphoid progenitors. *J Exp Med*. 2005;202(12):1669-77. doi: 10.1084/jem.20051190.

56. Mansson R, Welinder E, Åhsberg J, Lin YC, Benner C, Glass CK, et al. Positive intergenic feedback circuitry, involving EBF1 and FOXO1, orchestrates B-cell fate. *Proc Natl Acad Sci U S A*. 2012;109(51):21028-33. doi: 10.1073/pnas.1211427109.

57. Dengler HS, Baracho GV, Omori SA, Bruckner S, Arden KC, Castrillon DH, et al. Distinct functions for the transcription factor Foxo1 at various stages of B cell differentiation. *Nat Immunol*. 2008;9(12):1388-98. doi: 10.1038/ni.1667.

58. Lin YC, Jhunjhunwala S, Benner C, Heinz S, Welinder E, Mansson R, et al. A global network of transcription factors, involving E2A, EBF1 and Foxo1, that orchestrates B cell fate. *Nat Immunol*. 2010;11(7):635-43. doi: 10.1038/ni.1891.

59. Zhuang Y, Barndt RJ, Pan L, Kelley R, Dai M. Functional replacement of the mouse E2A gene with a human HEB cDNA. *Mol Cell Biol.* 1998;18(6):3340-9. doi: 10.1128/mcb.18.6.3340.
60. Yoshikawa G, Miyazaki K, Ogata H, Miyazaki M. The Evolution of Rag Gene Enhancers and Transcription Factor E and Id Proteins in the Adaptive Immune System. *Int J Mol Sci.* 2021;22(11). doi: 10.3390/ijms22115888.
61. Simakov O, Marlétaz F, Yue JX, O'Connell B, Jenkins J, Brandt A, et al. Deeply conserved synteny resolves early events in vertebrate evolution. *Nat Ecol Evol.* 2020;4(6):820-30. doi: 10.1038/s41559-020-1156-z.
62. Aronheim A, Shiran R, Rosen A, Walker MD. The E2A gene product contains two separable and functionally distinct transcription activation domains. *Proc Natl Acad Sci U S A.* 1993;90(17):8063-7. doi: 10.1073/pnas.90.17.8063.
63. Massari ME, Jennings PA, Murre C. The AD1 transactivation domain of E2A contains a highly conserved helix which is required for its activity in both *Saccharomyces cerevisiae* and mammalian cells. *Mol Cell Biol.* 1996;16(1):121-9. doi: 10.1128/mcb.16.1.121.
64. Liang JJ, Peng H, Wang JJ, Liu XH, Ma L, Ni YR, et al. Relationship between the structure and function of the transcriptional regulator E2A. *J Biol Res*

(Thessalon). 2021;28(1):15. doi: 10.1186/s40709-021-00146-5.

65. Bradney C, Hjelmeland M, Komatsu Y, Yoshida M, Yao TP, Zhuang Y. Regulation of E2A activities by histone acetyltransferases in B lymphocyte development. *J Biol Chem*. 2003;278(4):2370-6. doi: 10.1074/jbc.M211464200.

66. Zhang J, Kalkum M, Yamamura S, Chait BT, Roeder RG. E protein silencing by the leukemogenic AML1-ETO fusion protein. *Science*. 2004;305(5688):1286-9. doi: 10.1126/science.1097937.

67. Quong MW, Massari ME, Zwart R, Murre C. A new transcriptional-activation motif restricted to a class of helix-loop-helix proteins is functionally conserved in both yeast and mammalian cells. *Mol Cell Biol*. 1993;13(2):792-800. doi: 10.1128/mcb.13.2.792-800.1993.

68. Denis CM, Langelaan DN, Kirilin AC, Chitayat S, Munro K, Spencer HL, et al. Functional redundancy between the transcriptional activation domains of E2A is mediated by binding to the KIX domain of CBP/p300. *Nucleic Acids Res*. 2014;42(11):7370-82. doi: 10.1093/nar/gku206.

69. Bradney C, Hjelmeland M, Komatsu Y, Yoshida M, Yao TP, Zhuang Y. Regulation of E2A activities by histone acetyltransferases in B lymphocyte development. *J Biol Chem*. 2003;278(4):2370-6.

70. Chen WY, Zhang J, Geng H, Du Z, Nakadai T, Roeder RG. A TAF4 coactivator function for E proteins that involves enhanced TFIID binding. *Genes Dev.* 2013;27(14):1596-609. doi: 10.1101/gad.216192.113.
71. Lin YC, Benner C, Mansson R, Heinz S, Miyazaki K, Miyazaki M, et al. Global changes in the nuclear positioning of genes and intra- and interdomain genomic interactions that orchestrate B cell fate. *Nat Immunol.* 2012;13(12):1196-204. doi: 10.1038/ni.2432.
72. Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science.* 2009;326(5950):289-93. doi: 10.1126/science.1181369.
73. Benner C, Isoda T, Murre C. New roles for DNA cytosine modification, eRNA, anchors, and superanchors in developing B cell progenitors. *Proc Natl Acad Sci U S A.* 2015;112(41):12776-81. doi: 10.1073/pnas.1512995112.
74. Lio CW, Zhang J, González-Avalos E, Hogan PG, Chang X, Rao A. Tet2 and Tet3 cooperate with B-lineage transcription factors to regulate DNA modification and chromatin accessibility. *Elife.* 2016;5. doi: 10.7554/eLife.18290.
75. Miyazaki K, Watanabe H, Yoshikawa G, Chen K, Hidaka R, Aitani Y, et al. The

transcription factor E2A activates multiple enhancers that drive Rag expression in developing T and B cells. *Sci Immunol.* 2020;5(51). doi: 10.1126/sciimmunol.abb1455.

76. Ikawa T, Kawamoto H, Wright LY, Murre C. Long-term cultured E2A-deficient hematopoietic progenitor cells are pluripotent. *Immunity.* 2004;20(3):349-60. doi: 10.1016/s1074-7613(04)00049-4.

77. Mercer EM, Lin YC, Benner C, Jhunjhunwala S, Dutkowski J, Flores M, et al. Multilineage priming of enhancer repertoires precedes commitment to the B and myeloid cell lineages in hematopoietic progenitors. *Immunity.* 2011;35(3):413-25. doi: 10.1016/j.immuni.2011.06.013.

78. Lara-Astiaso D, Weiner A, Lorenzo-Vivas E, Zaretzky I, Jaitin DA, David E, et al. Immunogenetics. Chromatin state dynamics during blood formation. *Science.* 2014;345(6199):943-9. doi: 10.1126/science.1256271.

79. Seet CS, Brumbaugh RL, Kee BL. Early B cell factor promotes B lymphopoiesis with reduced interleukin 7 responsiveness in the absence of E2A. *J Exp Med.* 2004;199(12):1689-700. doi: 10.1084/jem.20032202.

80. Massari ME, Grant PA, Pray-Grant MG, Berger SL, Workman JL, Murre C. A conserved motif present in a class of helix-loop-helix proteins activates

transcription by direct recruitment of the SAGA complex. *Mol Cell*. 1999;4(1):63-73. doi: 10.1016/s1097-2765(00)80188-4.

81. Bossen C, Murre CS, Chang AN, Mansson R, Rodewald HR, Murre C. The chromatin remodeler Brg1 activates enhancer repertoires to establish B cell identity and modulate cell growth. *Nat Immunol*. 2015;16(7):775-84. doi: 10.1038/ni.3170.

82. Wang Y, Zolotarev N, Yang CY, Rambold A, Mittler G, Grosschedl R. A Prion-like Domain in Transcription Factor EBF1 Promotes Phase Separation and Enables B Cell Programming of Progenitor Chromatin. *Immunity*. 2020. doi: 10.1016/j.immuni.2020.10.009.

83. Wilson A, Held W, MacDonald HR. Two waves of recombinase gene expression in developing thymocytes. *J Exp Med*. 1994;179(4):1355-60. doi: 10.1084/jem.179.4.1355.

84. Wayne J, Suh H, Sokol KA, Petrie HT, Witmer-Pack M, Edelhoff S, et al. TCR selection and allelic exclusion in RAG transgenic mice that exhibit abnormal T cell localization in lymph nodes and lymphatics. *J Immunol*. 1994;153(12):5491-502.

85. Hsu LY, Loring J, Liang HE, Greenbaum S, Cado D, Zhuang Y, et al. A conserved transcriptional enhancer regulates RAG gene expression in developing

B cells. *Immunity*. 2003;19(1):105-17. doi: 10.1016/s1074-7613(03)00181-x.

86. Amin RH, Schlissel MS. Foxo1 directly regulates the transcription of recombination-activating genes during B cell development. *Nat Immunol*. 2008;9(6):613-22. doi: 10.1038/ni.1612.

87. Schulz D, Vassen L, Chow KT, McWhirter SM, Amin RH, Möröy T, et al. Gfi1b negatively regulates Rag expression directly and via the repression of FoxO1. *J Exp Med*. 2012;209(1):187-99. doi: 10.1084/jem.20110645.

88. Timblin GA, Xie L, Tjian R, Schlissel MS. Dual Mechanism of Rag Gene Repression by c-Myb during Pre-B Cell Proliferation. *Mol Cell Biol*. 2017;37(12). doi: 10.1128/mcb.00437-16.

89. Yannoutsos N, Barreto V, Misulovin Z, Gazumyan A, Yu W, Rajewsky N, et al. A cis element in the recombination activating gene locus regulates gene expression by counteracting a distant silencer. *Nat Immunol*. 2004;5(4):443-50. doi: 10.1038/ni1053.

90. Hao B, Naik AK, Watanabe A, Tanaka H, Chen L, Richards HW, et al. An anti-silencer- and SATB1-dependent chromatin hub regulates Rag1 and Rag2 gene expression during thymocyte development. *J Exp Med*. 2015;212(5):809-24. doi: 10.1084/jem.20142207.

91. Naik AK, Byrd AT, Lucander ACK, Krangel MS. Hierarchical assembly and disassembly of a transcriptionally active RAG locus in CD4(+)CD8(+) thymocytes. *J Exp Med*. 2019;216(1):231-43. doi: 10.1084/jem.20181402.
92. Yu W, Misulovin Z, Suh H, Hardy RR, Jankovic M, Yannoutsos N, et al. Coordinate regulation of RAG1 and RAG2 by cell type-specific DNA elements 5' of RAG2. *Science*. 1999;285(5430):1080-4. doi: 10.1126/science.285.5430.1080.
93. Monroe RJ, Chen F, Ferrini R, Davidson L, Alt FW. RAG2 is regulated differentially in B and T cells by elements 5' of the promoter. *Proc Natl Acad Sci U S A*. 1999;96(22):12713-8. doi: 10.1073/pnas.96.22.12713.
94. Kishi H, Wei XC, Jin ZX, Fujishiro Y, Nagata T, Matsuda T, et al. Lineage-specific regulation of the murine RAG-2 promoter: GATA-3 in T cells and Pax-5 in B cells. *Blood*. 2000;95(12):3845-52.
95. Wu CX, Zhao WP, Kishi H, Dokan J, Jin ZX, Wei XC, et al. Activation of mouse RAG-2 promoter by Myc-associated zinc finger protein. *Biochem Biophys Res Commun*. 2004;317(4):1096-102. doi: 10.1016/j.bbrc.2004.03.159.
96. Nguyen UT, Bittova L, Müller MM, Fierz B, David Y, Houck-Loomis B, et al. Accelerated chromatin biochemistry using DNA-barcoded nucleosome libraries. *Nat Methods*. 2014;11(8):834-40. doi: 10.1038/nmeth.3022.

List of abbreviations

Helix-loop-helix; HLH, transcription factor; TF, recombination-activating gene 1; Rag1, *Rag1* and *Rag2*; *Rag1/2*, pattern recognition receptors; PRRs, variable; V, diversity; D, joining; J, immunoglobulin; *Ig*, T-cell receptor; *TCR*, recombination signal sequence; RSS, Ig heavy chain; *Igh*, B-cell progenitors; pro-B-cells, pre-B-cell receptor; pre-BCR, B-cell precursors; pre-B-cells, Ig light chain; *Igl*, CD4⁻CD8⁻ double negative; DN, CD4⁺CD8⁺ double positive; DP, plant homeodomain; PHD, hematopoietic stem/progenitor cells; HSPCs, bone marrow; BM, multipotent progenitors; MPPs, Lymphoid-primed multipotent progenitors; LMPPs, common lymphoid progenitors; CLPs, granulocyte/macrophage progenitors; GMPs, innate lymphoid cells; ILCs, dendritic cells; DCs, lineage marker-negative; Lin⁻, IL7 receptor; IL7R, B-primed lymphoid progenitors; BLPs, GDNF Family Receptor 2; GFRA2, BM stroma cell antigen 1; BST1, gene regulatory network; GRN, basic helix-loop-helix; bHLH, cis-regulatory element; CRE, plasmacytoid DCs; pDCs, classical DC; cDC, lymphoid tissue inducer; LTi, Ig kappa chain enhancer E2; κ E2, Achaete-scute T3; AS-C T3, electrophoretic mobility shift assay; EMSA, embryonic stem cells; ESCs, very early stage of B-cell development; prepro-B-cell, whole-genome duplication; WGD, activation domains; AD,

histone acetyltransferase; HAT, Ten-eleven translocation; TET, PCET; EP300/KAT3B-CBP/KAT3A and ETO target, Myoglobin 1; Mb1, Paired box 5; Pax5, bromodomain and extraterminal; BET, intrinsically disordered protein regions; IDRs, enhancer of *Rag*; *Erag*, anti-silencer element; *ASE*, *Rag*-T-cell enhancer; *R-TEn*, *Rag* B-cell enhancer 1/2; *R1B* and *R2B*, *Rag1* promoter; *R1pro*, E-box mutations in the *R-TEn* region; *R-TEn-p1-E-box^{mut/mut}*, super-enhancer; SE, E-box mutation in *R1pro*; *R1pro-E-box^{mut/mut}*, bromodomain; BRD, heterogeneous nuclear ribonucleoproteins H1; HNRNPH1, heterogeneous nuclear ribonucleoproteins F; HNRNPF, TATA-box binding protein Associated Factor 3/Transcription initiation factor TFIID subunit 3; TAF3/TFIID, V-Set Pre-B Cell Surrogate Light Chain 1; Vpreb, C-X-C Motif Chemokine Receptor 4; Cxcr4, Hi-C; Analysis for genome-wide chromatin interactions, FUS; Fused in Sarcoma, Special AT-Rich Sequence Binding Protein 1; SATB1, phorbol myristate acetate; PMA, BRG1/BRM associated factor; BAF, TNF Receptor Associated Factor 6; Traf6.

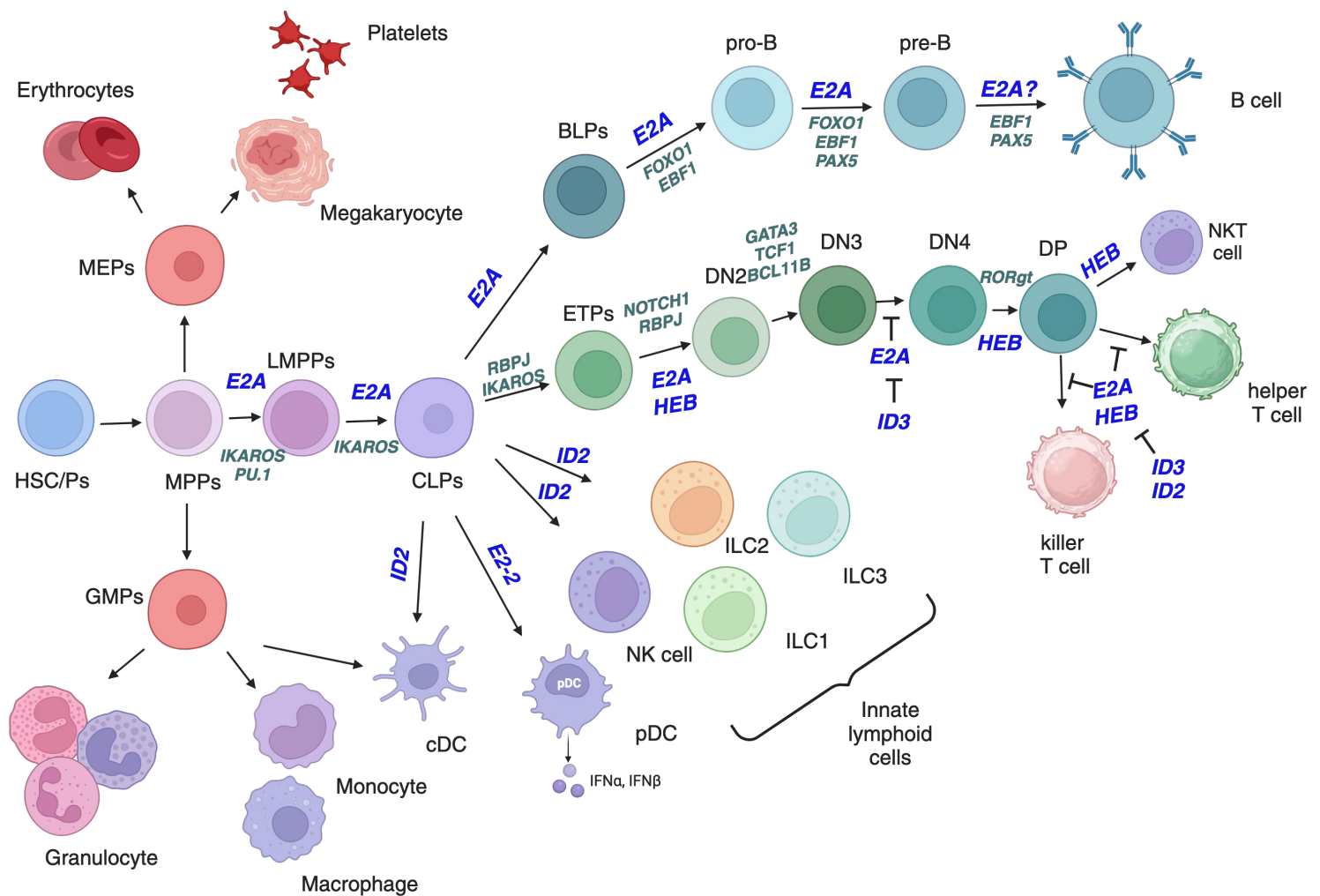
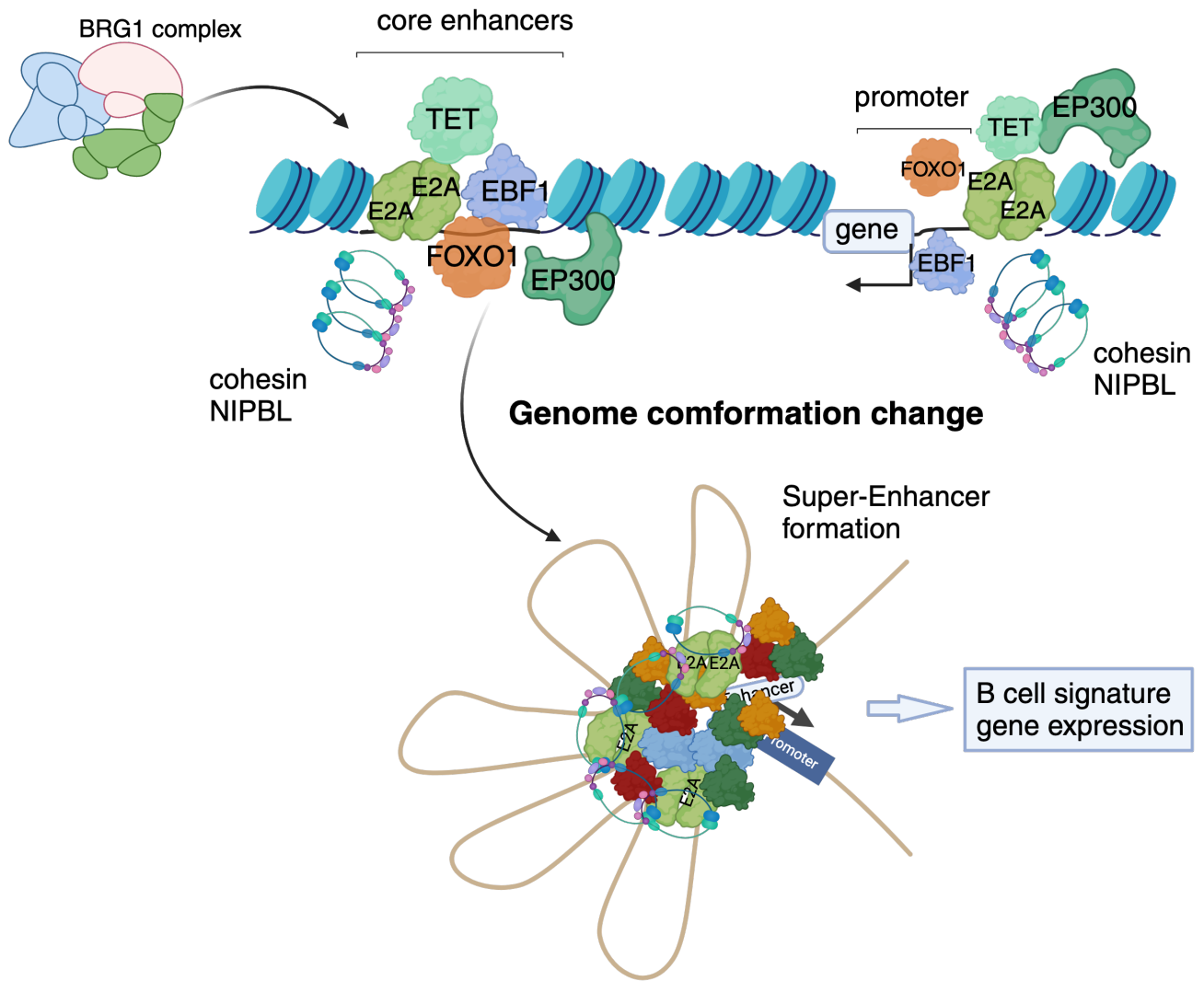
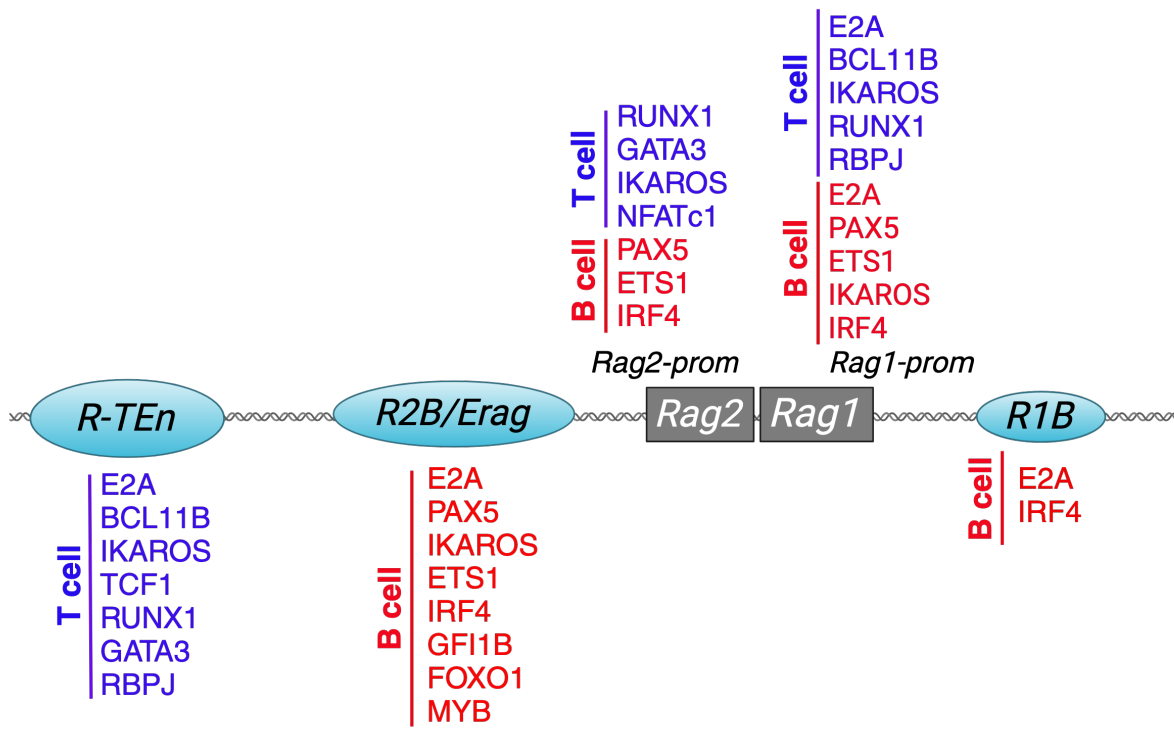


Figure 2



A



B

