Cell type–specific gene expression is driven by the interplay between lineage-specific transcription factors and cis-regulatory elements to which they bind. Adaptive immunity relies on RAG-mediated assembly of T cell receptor (TCR) and immunoglobulin (Ig) genes. Although Rag1 and Rag2 expression is largely restricted to adaptive lymphoid lineage cells, it remains unclear how Rag gene expression is regulated in a cell lineage–specific manner. Here, we identified three distinct cis-regulatory elements, a T cell lineage–specific enhancer (R-TEn) and two B cell–specific elements, R1B and R2B. By generating mice lacking either R-TEn or R1B and R2B, we demonstrate that these distinct sets of regulatory elements drive the expression of Rag genes in developing T and B cells. What these elements have in common is their ability to bind the transcription factor E2A. By generating a mouse strain that carries a mutation within the E2A binding site of R-TEn, we demonstrate that recruitment of E2A to this site is essential for orchestrating changes in chromatin conformation that drive expression of Rag genes in T cells. By mapping cis-regulatory elements and generating multiple mouse strains lacking distinct enhancer elements, we demonstrate expression of Rag genes in developing T and B cells to be driven by distinct sets of E2A-dependent cis-regulatory modules.

INTRODUCTION
Adaptive immunity relies on Rag-mediated assembly of antigen receptor genes [T cell receptor (TCR) and immunoglobulin (Ig) genes] in T and B cells that enables the receptors to recognize diverse antigens. The highly conserved Rag1 and Rag2 initiate V(D)J recombination of TCR and Ig genes during T and B cell development (1–3). Expression of recombination activating gene 1 (Rag1) and Rag2 (Rag1/2) is restricted to T and B progenitors and precursors [CD4-CD8 double-negative (DN) 3 and CD4-CD8 double-positive (DP) thymocytes and pro-B and pre-B cells] (4–6). Recent studies have identified a class of regulatory elements, named super-enhancers (SEs), which are defined as genomic regions with dense clustering of highly active enhancers bound by an array of lineage-specific transcription factors (TFs) (7, 8). An essential feature of SEs is to control genes that have prominent roles in cell type–specific processes, thereby establishing cell identity (7, 8). An SE displays highly cooperative property, which is characterized by the collaborating interactions between numerous TFs, mediators, and RNA polymerases, and the clusters of enhancers within SEs are in close physical contact with one another and with the promoter regions of the gene that they activate, showing elaborating chromatin interactions (SE formation) (9–11). Because SEs can be formed as a consequence of introducing a single TF binding, SEs are exceptionally vulnerable to perturbation of key protein components, which are related to cell type specificities (9). We hypothesized that expression of Rag1/2 during T and B cell development might be regulated by adaptive lymphocyte–specific SEs. Previous studies have identified several cis-regulatory elements in the Rag1/2 gene locus, including the anti-silencing element (ASE) and enhancer of Rag (Erag) (12, 13). Deletion of these elements resulted in the partial loss of Rag1 and Rag2 activity, suggesting the presence of additional regulatory elements for Rag1/2 gene expression (5, 13–15). Besides, what TFs regulate these regulatory elements for Rag gene expression during T and B cell development remains to be determined.

It is well established that a majority of adaptive lymphocyte development trajectories require the regulation by the members of the helix-loop-helix families, such as E proteins and Id proteins (16–18). The mammalian E protein E2A [transcription factor 3 (TCF3)], along with early B cell factor 1 (EBF1), Forkhead box protein O1 (FOXO1), and PAX5, plays particularly important roles in B cell lineage commitment, and they modulate the topological changes of Ig gene loci (19–22). E2A acts in concert with another E protein, HEB [Transcription factor 12 (TCF12)], to establish T cell identity and to determine the cell fate of adaptive versus innate lymphoid lineage (23, 24). On the basis of the commonality of the role of E2A in both T and B cell commitment, we hypothesized that E2A directly might regulate Rag1/2 gene expression in both T and B cell lineages.

Here, we used multiple in vivo approaches to identify T or B cell–specific cis-regulatory elements critical for Rag1/2 expression during

---

1Laboratory of Immunology, Institute for Frontier Medical Sciences, Kyoto University, Kyoto 606-8507, Japan. 2Laboratory of Integrative Biological Sciences, Institute for Frontier Medical Sciences, Kyoto University, Kyoto 606-8507, Japan. 3Institute for Chemical Research, Kyoto University, Uji 611-0011, Japan. 4Baylor Institute for Immunology Research, Baylor Scott & White Research Institute, Dallas, TX, USA. 5Department of Pathology and Tumor Biology, Graduate School of Medicine, Kyoto University, Kyoto 606-8501, Japan. 6Laboratory of Immune Regulation, Institute for Frontier Medical Sciences, Kyoto University, Kyoto 606-8507, Japan. 7Department of Medical Chemistry, Graduate School of Medicine, Kyoto University, Kyoto 606-8501, Japan. 8Institute for the Advanced Study of Human Biology (WPI ASHBI), Graduate School of Medicine, Kyoto University, Kyoto 606-8501, Japan. 9Department of Medicine, Center for Hematology and Regenerative Medicine, Karolinska Institute, Stockholm, Sweden.

*These authors contributed equally to this work.

†Corresponding author. Email: mmiyazaki@infront.kyoto-u.ac.jp
adaptive lymphocyte development, which are essential for V(D)J recombination of antigen recognition receptor genes. Blocking E protein binding to T cell–specific Rag enhancer led to the loss of Rag1/2 gene expression and the disruption of the three-dimensional (3D) genome organization in the Rag gene loci, resulting from the loss of cohesin-mediated chromatin looping. We also found that Rag1 promoter activities are regulated by E2A in vivo. Overall, we demonstrate that E2A activates multiple enhancers to drive expression of Rag1/2 gene during T and B cell development.

RESULTS
Identification of T or B cell–specific regulatory regions in the Rag1/2 gene cluster
As a first approach to explore the regulatory regions for the Rag1/2 expression, we analyzed E2A binding at the Rag1/2 gene cluster in the T progenitors (pro-T; DN3), a DP cell line (A12), and B progenitors (pro-B) (table S1) (21). E2A bound to discrete distal regulatory elements in pro-T and pro-B cells that were associated with assay for transposable-accessible chromatin (ATAC-seq) signals and p300 binding in a cell type–specific manner (Fig. 1A). We designated the putative T cell–specific regulatory region of the Rag genes as R-TEn (Rag–T cell enhancer) and the two B cell–specific regions as R1B and R2B (Rag1-B and Rag2-B), respectively. We also found common E2A binding sites at the Rag1 promoter region (R1pro) in both pro-T and pro-B cells (Fig. 1A). In addition, pro-T and pro-B cells showed the distinct chromatin topologies that resulted from the genomic interactions between individual R-TEn and R1B/R2B regions and Rag1/2 promoters that likely involved E2A and cohesin occupancy at the open chromatin regions (Fig. 1B).

Because CpG DNA methylation at enhancers might suppress enhancer function, we analyzed the DNA methylation status at these regions (26–28). Given the mutually exclusive nature of these cell type–specific enhancers of the Rag genes, CpG DNA was hypomethylated in the R-TEn region in DN3 and DP cells but hypermethylated in pro-B and pre-B cells (Fig. 1C). On the contrary, R1B and R2B were hypomethylated in both T and B progenitors/precursors (Fig. 1D). Of note, TFs that are involved in T and B cell commitment bound to R-TEn, R2B, R1B, or R1pro in T and B progenitor/precursor cells, accompanied by chromatin accessibilities (Fig. 1E, fig. S1A, and table S1) (29, 30). According to the analysis of the H3K27Ac signals by the Ranking Ordering of Super-Enhancer (ROSE) algorithm (8), we distinguished between typical enhancers and SEs and found that Rag1/2 gene cluster along with the R-TEn in DP cells and R2B and R1pro in pre-B cells were defined as SEs (Fig. 1E). Some of these cis-regulatory elements were partially overlapped with the previously identified T and B cell type–specific ASE and Erag for Rag1/2 expression (fig. S1, B and C) (12, 13). Collectively, we identified distinct regulatory modules that drive expression of Rag genes in developing T and B cells.

R-TEn, R1B, and R2B are essential for cell type–specific Rag1/2 expression
To validate the biological significance of these T or B cell–specific regulatory regions in vivo, we knocked out these regions individually in mice (R-TEn\textsuperscript{d/d}, R1B\textsuperscript{d/d} and R2B\textsuperscript{d/d}; green bars in Fig. 1, C and D). The deletion of R-TEn resulted in a severe developmental block at the positive selection of DP cells, as seen in ASE deletion mice (13), and we further found the developmental block at the β-selection of DN3 cells. The absence of mature CD4SP, CD8SP, and invariant natural killer T cells, loss of intracellular TCRβ\textsuperscript{d/d} DN3 cells, and partial loss of y8T cells in 4-week-old and fetal thymi were observed in the R-TEn\textsuperscript{d/d} mice without affecting B cell development in the bone marrow (BM) (Fig. 2, A and B, and fig. S2, A to D). As a consequence of Rag1/2 down-regulation in DN3a and DP cells, the incidence of TCRβ V-DJ recombination in DN3a cells and TCRα V-J recombination in CD5− CD69+ DP cells was severely declined, whereas TCRβ D-J gene recombination was less affected (Fig. 2, C to E, and fig. S2E).

As for the mice with deleted R1B and R2B (R1B\textsuperscript{d/d} R2B\textsuperscript{d/d}), significant accumulation of pro-B cells was observed with a few pre-B cells and the absence of IgM+ immature and IgD+ mature B cells (Fig. 2, F and G). In contrast, the deletion of either R1B or R2B showed mild to modest developmental defect at the pro-B stage; comparable phenotypes were observed in the mice with Erag deletion, suggesting the enhancer redundancy in R1B and R2B for Rag1/2 expression (Fig. 2, F and G) (12). As expected, the loss of R1B and R2B did not affect thyymocyte development including γδT cells (fig. S2F). The DNA recombination of IgH and Igκ/Igλ were severely compromised in R1B\textsuperscript{d/d} R2B\textsuperscript{d/d} pro-B and pre-B cells, resulting from the down-regulation of Rag1/Rag2 expression (Fig. 2, H and I, and fig. S2, G and H). Together, whereas T and B cells require Rag1/2 expression for DNA recombination of TCR and Ig genes, the expression of Rag1/2 genes in these two lineages is regulated by distinct regulatory elements.

R-TEn and R1B/R2B orchestrate spatial organizational of the Rag1/2 genomic region
To determine how these cell type–specific regulatory regions modulate chromatin accessibilities and folding, we performed ATAC-seq and in situ Hi-C (31, 32). In wild-type DN3 and DP cells, prominent chromatin accessibilities determined by ATAC-seq at R-TEn, and Rag1 and Rag2 promoter regions were observed (Fig. 3A). Of note, the deletion of R-TEn led to significant loss of the open chromatin regions across the Rag1/2 gene cluster in DP cells; however, in DN3 cells, chromatin accessibilities in Rag1 and Rag2 promoter regions were affected by the loss of R-TEn (Fig. 3A). Similar to this, the loss of R1B and R2B compromised chromatin accessibilities of the Rag1/2 gene cluster in pro-B and pre-B cells (Fig. 3B). These results suggest that the enhancer activities are required to keep the promoter and other regions accessible in the process of differentiation into precursor cells (DP and pre-B cells).

With the in situ Hi-C data, we generated genomic interaction contact maps and applied principal components (PC) analysis to segregate genome into A versus B compartments (33). We observed that the genomic region that spanned the R-TEn and Rag1/2 gene cluster was located in the transcriptionally permissive compartment A in control DP cells (Fig. 3C). The circos plot indicated an intricate and elaborate pattern of prominent interactions involving both small and large loops across the entire genomic region that was consistent with a SE formation (Fig. 3D). Upon the deletion of R-TEn, the transcriptionally repressive compartment B extended to the R-TEn region, as reflected by the APCR (Fig. 3C). Despite the loss of interactions, the Rag1/2 region remained transcriptionally permisssible as evident by the compartment A assignment. Furthermore, the losses of sub–topologically associating domain (TAD) formation and local compaction were in agreement with the decrease in TADs and loops scores and changes in the distal-to-local ratio (DLR) as...
**Fig. 1.** T or B cell–specific regulatory elements associated with E2A occupancy in Rag1/2 gene cluster. (A) E2A and p300 occupancy in Rag1-deficient DN3 cells (pro-T), A12 cells constitutive E47 expression (DP), and pro-B cells along with normalized ATAC-seq reads. (B) Circos diagrams illustrate Hi-C, ChIP-seq, and ATAC-seq signals in pro-T and pro-B cells. Black lines represent chromatin interactions. P values are correlated with the thickness of the connecting lines (i.e., the lower the P values, the thicker the lines). (C and D) Cpg DNA methylation in (C) R-TEn, (D) R1B, and R2B in DN3 and DP, pro-B, and pre-B cells. Top panels show whole-genome bisulfite sequencing (WGBS) analysis along with E2A occupancy and ATAC-seq reads within R-TEn (C), R1B, and R2B (D). Bottom panels show the Cpg DNA methylation status in regulatory regions. Methylated status (black) and unmethylated status (white) are shown. (E) Genome browser shots of ATAC-seq and ChIP-seq of H3K27Ac, H3K4me1, and indicated TF bindings in DP and pre-B cells. Red bars on top of the panels indicate SE regions defined by the ROSE algorithm. Obtained data (Fig. 1, A to D) were described in table S1. See also fig. S1.
Fig. 2. R-TEn, R1B, and R2B are essential for T or B cell type–specific Rag1/2 expression. (A) Flow cytometric analysis of CD4 versus CD8 expression, CD69 versus CD5 (DP), and CD28 versus forward scatter (FSC) (DN3) from wild-type control (Ctrl) and R-TEn d/d mice. (B) Cell numbers of thymocytes and indicated cell populations; single-positive (SP), double-positive (DP), and CD25 (CD4−CD8−Linneg), and CD28 versus forward scatter (FSC) (DN3) from wild-type control (Ctrl) and R-TEn d/d mice.

(C) PCR analysis evaluating TCR β rearrangements in DN3a cells. Arrow and arrowheads indicate germline (GL) and rearranged bands, respectively. (D) qPCR analysis of the rearrangement of Vα-Jα segments in genomic DNA extracted from CD69−DP cells. (E) qPCR analysis of Rag1/Rag2 transcript levels. (F) Flow cytometric analysis of B cell populations. (G) Cell numbers of indicated cell populations. (H) PCR analysis involving IgH rearrangements in the genomic DNA from pro-B cells. (I) qPCR analysis of Rag1/Rag2 transcript levels. All mice were analyzed at 4 weeks old. Data represent the means ± SD. *P < 0.05, **P < 0.01, and ***P < 0.001 (Student’s test). Data are representative of four independent experiments [A and B; n = 5 and 7, biological replicates], two independent experiments produced similar results (C, D, H, and I), one experiment [E; n = 3, technical replicates], and four experiments [F and G; n = 7, 5, 5, and 6, biological replicates]. See also fig. S2.
Fig. 3. Impact of R-TEn on chromatin accessibilities and spatial chromatin organization in local compaction and sub-TAD formation across the Rag1/2 gene cluster. (A and B) Browser shots of normalized ATAC-seq reads across the Rag1/2 gene cluster and Traf6 gene in (A) DN3a and CD69− DP cells and (B) pro-B and pre-B cells in 4-week-old wild-type control, R-TEn d/d, and R1B d/d R2B d/d mice. Blue arrowheads indicate significantly reduced peaks in R-TEn d/d and R1B d/d R2B d/d cells. (C) In situ Hi-C contact maps. Left: Positive PC1 values (compartment A) in black, negative PC1 values (compartment B) in gray, and changes in PC1 and distal-to-local ratio (DLR) are shown. Right: Contact maps, TADs and Loops score, and PC1 values across the indicated genomic region are shown. Color scales in the Hi-C contact maps indicate the ratios of observed versus expected interaction frequencies: blue, lower than expected; and red, higher than expected. Contact maps, wild-type control (top) and R-TEn d/d (bottom) DP cells. (D and E) Circos diagrams representing genomic interactions (black lines) across the indicated genomic region in (D) DP and (E) pro-B cells in 4-week-old wild-type control, R-TEn d/d, and R1B d/d R2B d/d mice. Thickness of connecting lines reflects P values associated with the indicated interactions. One experiment [(A and B); n = 2, biological replicates] and one experiment (C to E). See also fig. S3.
compared with those of control DP cells (Fig. 3, C and D) (34). Likewise, similar results were observed in R1B<sup>−/−</sup>R2B<sup>−/−</sup> pro-B cells but to a lesser extent (Fig. 3E and fig. S3, A and B). These results suggest that T or B cell–specific enhancers are required for the establishment of 3D chromatin organization of Rag1/2 gene SE during T and B cell development.

E protein binding is indispensible for both the enhancer and promoter activities of Rag genes

Because a number of TFs involved in T cell commitment were recruited to R-TEn, we examined which TFs and chromatin modulators regulate the R-TEn activity and Rag gene SE formation (table S1) (28, 35–38). Rag1/2 expression and chromatin accessibility of R-TEn were not affected in Tcf7<sup>−/−</sup>, Bcl11b<sup>−/−</sup>, or ThymoD-deficient thymocytes (fig. S4, A and B). Upon the deletion of Satb1 or Rad21, H3K27Ac levels at R-TEn and Rag1 promoter regions remained high, albeit to a lesser extent (fig. S4C). The deletion of E2A (Tcf3) and HEB (Tcf12) in DP cells led to severe impairment of positive selection of DP cells and significant down-regulation of Rag1/2 expression in CD5<sup>−</sup>CD69<sup>−</sup> DP cells, which were accompanied by the significant reduction of ATAC-seq signals in R-TEn and Rag1/Rag2 promoter regions (Fig. 4K). These results strongly suggest that the Rag1/2 enhancer and Rag1 promoter activities critically require the binding of E proteins. Because TCRβ D-J recombination and γδT cell development were severely impaired, E2A binding to the Rag1 promoter is essential for initiating Rag1 expression upon adaptive lymphocyte lineage commitment and independent from enhancer activity (a stepwise Rag expression) (fig. S4L).

E proteins orchestrate cohesin-mediated looping to activate Rag gene expression

To understand how E proteins regulate enhancer and the genome conformation, we performed chromatin immunoprecipitation sequencing (ChIP-seq), ATAC-seq, and chromosome conformation capture (3C) assay (15, 39). Consistent with the SEs, we found a stitched pattern of H3K27Ac and H3K4me1 peaks, associated with the chromatin interactions of the R-TEn and promoter regions across the Rag1/2 gene cluster in wild-type DP cells (Fig. 5, A and B). In R-TEn-p1-Ebox<sup>mut/mut</sup> DP cells, the chromatin interactions across the entire gene cluster were disrupted, which coincided with the loss of enhancer activity, despite no impact on the CD4 gene SE formation (Fig. 5, A and B, and fig. S5A). These results indicate that the E protein binding to R-TEn is essential for the Rag gene SE formation. The Cpg islands at peak1 and peak2 of R-TEn were hypermethylated when E protein binding was absent, suggesting a failure of ten-eleven translocation (TET) family protein recruitment to R-TEn by E2A/E proteins (Fig. 5C) (27).

Recent studies demonstrated that, in differentiated neutrophils, SEs act, at least in part, by recruitment of the cohesin machinery (40). To explore how E protein binding at R-TEn mediates the chromatin interactions across the Rag1/2 gene cluster, we examined for possible CCCTC-binding factor (CTCF) and cohesin (Smc3) occupancy. In wild-type DP cells, most of the CTCF binding sites were co-occupied with the cohesin upstream of the R-TEn and around the Rag1 gene (Fig. 5D). The number of CTCF and cohesin co-occupied sites was significantly reduced in the R-TEn-p1-Ebox<sup>mut/mut</sup> DP cells, as evident by the loss of cohesin binding across the Rag1/2 gene cluster (Fig. 5D). Because the deletion of R-TEn resulted in the loss of sub-TAD formation at the Rag1/2 gene cluster (Fig. 3C), we investigated how CTCF and cohesin occupancy might influence the sub-TAD boundaries. As expected, CTCF and cohesin co-occupancies were observed at the sub-TAD boundaries (TAD-Bd1, TAD-Bd1, TAD-Bd2, and TAD-Bd3, in wild-type DP cells; notably, TAD-Bd2 includes R1B enhancer region (Fig. 5E). In the R-TEn-p1-Ebox<sup>mut/mut</sup> DP cells, the sub-TAD1 could no longer form because of the significant loss of cohesin occupancies across the sub-TAD1 and at the TAD-Bd1 and TAD-Bd2 (Fig. 5E). Together, we conclude that E protein binding to cis-regulatory elements facilitates the recruitment of the cohesin machinery for Rag gene SE formation (Fig. 5F).

Chromatin state dynamics of Rag1/2 regulatory elements and Rag gene cluster in lymphocytes and macrophages

To investigate how those cell type–specific Rag1/2 enhancers develop during T or B cell differentiation from hematopoietic stem cell (HSC), we analyzed our data together with ImmGen ATAC-seq data (table S1). ATAC-seq signals in R2B and Rag1/2-promoter regions emerged at the common lymphoid progenitor (CLP) stage.
Fig. 4. E protein bindings to R1-TEn and R1pro are indispensable for Rag1/2 expression and DNA recombination of TCR and Ig genes. (A) CD4 versus CD8 expression and CD69 versus CD5 (DP) in control and Tc6fmut/Tc2fmut CD4+CD8− mice. (B) qPCR analysis of Rag1/2 expression in sorted CD5−CD69− DP cells. (C) Browser shots of ATAC-seq reads across the indicated genomic region in DP cells. Blue arrowheads indicate significantly reduced peaks in Tc6fmut/Tc2fmut CD4+CD8− cells. (D) CD4 versus CD8 in total thymocytes and CD9 versus CD5 (DP), derived from (D) R-TEn-p2mut, R-TEn-p2mut, and (E) R1pro-Eboxmutmut mice. (F) CD4 versus CD8 expression in total thymocytes and CD44 versus CD25 expression in CD4+ CD5−CD69− DP cells. (G) The ratio of DN3b to DN3a and cell number of γδ T cells in the thymus. (H) Flow cytometric analysis in control and R1pro-Eboxmutmut mice. (I) PCR analysis involving TCRβ-Jδ and Vδ-Jδ, and IgH rearrangements in the genomic DNA. Arrow and arrowheads indicate germline and rearranged bands, respectively. (J) qPCR analysis of Rag1/2 expression. (K) Browser shots of ATAC-seq reads across the Rag1/2 gene cluster (top) and Rag1 promoter region (bottom). Blue arrowheads indicate significantly reduced peaks. All mice were analyzed at 4 weeks old. Data represent the means ± SD. **p < 0.01 and ****p < 0.001 (Student's t test). Data are representative of two experiments [(A and B); n = 3; and (C), three (D), four (E), three [(F and G); n = 7, 5, 4, and 4; and (H)], one [(I and J)]; n = 3 technical replicates], and two experiments (K). See also fig. S4.
Fig. 5. E protein binding to R-TEn directs 3D genome structure for SE formation through CTCF and cohesin complex-mediated chromatin looping. (A) ATAC-seq reads and ChIP-seq for H3K27Ac, H3K4me1, and H3K27me3 in the Rag1/2 gene cluster from wild-type control and R-TEn-p1-Ebox\textsuperscript{mut/mut} mice. Green bars and arrowheads indicate significantly reduced peaks in R-TEn-p1-Ebox\textsuperscript{mut/mut} DP cells. (B) 3C analysis with R-TEn bait (top) or Rag1 promoter bait (bottom) in wild-type and R-TEn-p1-Ebox\textsuperscript{mut/mut} DP cells. (C) E protein binding to R-TEn-peak1 is required for CpG DNA demethylation in R-TEn region in thymocytes. (D) ChIP-seq of CTCF and Smc3 (cohesin) and ATAC-seq reads in wild-type and R-TEn-p1-Ebox\textsuperscript{mut/mut} DP cells. Black (CTCF) and blue (Smc3) arrowheads indicate significantly reduced peaks. (E) E protein binding to the regulatory regions directs active sub-TAD formation through the CTCF and cohesin recruitment to the sub-TAD boundaries. Contact maps (Fig. 3C), ChIP-seq of CTCF and Smc3, and ATAC-seq reads are shown. Sub-TAD boundaries (TAD-Bd1, TAD-Bd2, and TAD-Bd3) are indicated by blue boxes. Black (CTCF) and blue (Smc3) arrowheads indicate significantly reduced peaks. (F) Model depicting the stepwise establishment of adaptive lymphocyte–specific SE for Rag1/2 gene expression mediated by E2A. Data are representative of two experiments with similar results (A to D). Data represent the means ± SD (IB), n = 3 technical, *P < 0.05, **P < 0.01, and ****P < 0.0001 (Student's t test). See also fig. S5.
and persisted until the pre-B stage (Fig. 6A, left). In contrast, R-TEn became accessible in DN2a cells and remained open until it closed in mature CD45P and CD85P cells (Fig. 6A, right). Of note, R2B was accessible at the noncommitted DN1 stage but became inaccessible at the committing DN2a stage, indicating a mutual exclusivity of enhancer engagement between R2B and R-TEn after lineage commitment. On the other hand, RIB showed a low level of ATAC-seq signals during thymocyte development (Fig. 6A). We also observed chromatin accessibility at NWC (Nad Wyraz Ciekawy), which is the third evolutionarily conserved gene with unknown function in Rag gene locus and is located within Rag2 intron, from long-term HSC to lymphoid-primed multipotent progenitor (Fig. 6A) (41). By using available Hi-C data, we observed an increasing number of loops across the Rag1/2 gene cluster during T and B cell developmental progression (Fig. 6S and table S1). These loops were absent when Rag1/2 genes were not expressed in mature T and B cells (Fig. 6A). The timing of the loop formation correlated strongly with the changes in the PC1 compartment and TADs and Loops scores between T and B cell development (Fig. S6, B and C).

Rag1/2 exclusively express in T and B progenitor and precursor cells, which raises the question on how Rag1/2 are repressed in innate immune cells. To address this, we compared the Hi-C data of BM-derived macrophage (BMDM) with those of DP and pre-B cells. Although TADs are known to be mainly conserved among different cell types (42), a distinct TAD boundary was formed between Rag1/2 and Tra6 genes in BMDMs, which is consistent with TAD-Bd2 region, whereas the TAD boundaries in DP and pre-B cells were formed upstream of R-TEn (Fig. 6B and fig. S6, D and E). The sub-TAD boundaries were described in Fig. 5F (TAD-Bd1, TAD-Bd2, and TAD-Bd3). We noted the similarities in the compartment and TADs and Loops score between the R-TEn<sup>Ad</sup> DP cells and BMDMs, implying the importance of the R-TEn activity for the local genome organization (Fig. 6B). To examine the involvement of CTCF and cohesin complex in insulator formation and chromatin interaction in BMDMs, we compared CTCF and cohesin binding and interaction across this locus and at the TAD-Bd2 region between DP, pre-B, and BMDM cells. We found three regions of CTCF and cohesin binding in TAD-Bd2 region (TAD-Bd2a, R1B enhancer, and TAD-Bd2b) (Fig. 6C). In BMDMs, the loss of CTCF and cohesin binding across the Rag1/2 gene cluster locus between TAD-Bd1 and TAD-Bd2 appeared to be consistent with the decrease in chromatin interactions (Fig. 6C). The CTCF and cohesin co-occupancy at TAD-Bd2a was associated with the sub-TAD boundary region of TAD-Bd1. Also, RIB did not interact with the Rag enhancers or TAD-Bd1 but instead interacted with the neighboring genes in BMDMs (Fig. 6C). In contrast, TAD-Bd2a and RIB interacted extensively with TAD-Bd1 and the Rag enhancers as well as within the domains in DP and pre-B cells (Fig. 6C). These differences in chromatin interactions may explain the mechanism of cell type–specific Rag gene expression during adaptive lymphocyte development (fig. S6F).

In addition, compared with HSC, ΔPC1 and ΔDLR in DP and pre-B cells revealed a compartment switch from B to A and an increase in local compartment during their developmental process from HSC, whereas these changes were absent in BMDM (fig. S6G). These results suggest that the topological insulator is developed during innate immune cell differentiation and that the boundary formation represses the Rag1/2 gene cluster by sequestering it in compartment B (Fig. 6D).

**Enhancers of the Rag1/2 genes and their E-boxes are evolutionarily conserved**

Here, we investigated and showed the critical functions and importance of R-TEn, R1B, and R2B enhancers in regulating Rag1/2 expression and establishing 3D chromatin organization in vivo, especially E-box DNA sequences in R-TEn-peak1 enhance region. Although Rag1/2 genes are known to be conserved among jawed vertebrates, the conservation of their regulatory elements has yet to be investigated. Therefore, we analyzed the phylogenetical conservation of the R-TEn, R1B, and R2B sequences, including E-box sequences in R-TEn-peak1. Regarding peak1 of R-TEn and peak2 of R2B, sequence similarities were readily observed among mammals as well as most birds and reptiles (Fig. 7A and fig. S7, A to C). On the other hand, for peak2 of R-TEn and R1B or peak1 of R2B, sequence similarities were detected in mammals (Fig. 7A and fig. S7, B and C). However, sequence similarities of these enhancers were not noticeable in the corresponding genomic regions of amphibians and fish (Fig. 7A and fig. S7C). Furthermore, the conserved R-TEn, R1B, and R2B regions were found to harbor conserved E-boxes (Fig. 7B). Among the seven E-boxes identified in the mouse R-TEn enhancer, which were essential for Rag1/2 expression in thymocyte development (Fig. 4), E-box2 (E2) and E-box3 (E3) were highly conserved among most mammals, birds, and reptiles (Fig. 7C and fig. S7D).

**DISCUSSION**

The establishment of cell type–specific gene expression programs is driven by lineage-specific TFs, cis-regulatory elements, and the 3D organization of the genome (43, 44). Depending on the circumstances, this interplay can function as a barrier, primer, or drive to control cell fate. Our study underscores the importance of lineage-specific TF bindings to regulatory elements in the regulation of 3D genome organization upon lineage commitment.

The assembly of TCR and Ig genes relies upon Rag1/2 expression and changes in chromatin conformation and accessibility across the TCR and Ig genes. E2A has been shown to regulate the chromatin conformation and accessibility of these gene loci (22, 45–48). Although previous studies reported the down-regulation of Rag1/2 expression in E2A-deficient pro-T, pro-B, and CLP cells, it still remained to be clarified whether E2A could directly regulate the Rag gene expression, because E2A deficiencies affected many genes required for T and B cell development (19, 23, 49, 50). Here, by mutating E-boxes within enhancers, we show that E protein binding to regulatory elements is required for Rag1/2 expression in vivo. Similarly, the generation of E-box mutants in TCR and Ig locus should be required to distinguish the role of E2A in Ig/TCR topology without affecting Rag gene expression.

E protein activity is required for T and B cell development and suppression of innate lymphoid lineage (23). However, it remains unclear how E proteins function differently during T and B cell development. We speculate that E2A acts in synergy with Notch signaling upon T cell lineage commitment; in contrast, without Notch signaling, E2A orchestrates B cell fate along with EBF1, FOXO1, and other TFs (21). We found the co-occupancy of intracellular domain of Notch (ICN) and E2A at R-TEn (fig. S1A) and the accessible R-TEn in DN2a cells, which received Notch signaling to commit into T cell lineage (Fig. 6A).

Conspicuously, the cell type–specific enhancers and promoter regions of the Rag1/2 genes were accessible at the progenitor stage
Fig. 6. Chromatin dynamics of R-TEn, R1B, and R2B during T and B cell differentiation and insulator formation in macrophages. (A) ATAC-seq reads across the Rag1/2 gene cluster during B (left) and T (right) cell development from HSC (Immgen). Blue arrow indicates NWC. (B) Contact maps from Hi-C data of wild-type DP, pre-B (64), and BMDM (65) are shown. Black dotted lines indicate TAD boundaries, and yellow lines indicate sub-TADs. PC1 values and TADs and Loops score in DP, pre-B, BMDM, and R-TEn d/d DP cells are shown. (C) Circos diagrams representing genomic interactions and ChIP-seq of CTCF and cohesin (Smc1a and Rad21) are shown across the Rag1/2 gene cluster in DP, pre-B, BMDM, B1, B2, and B3 were as seen in Fig. 5E. The degree of the thickness of the connecting lines reflects P values associated with the indicated interactions. Sub-TAD boundaries [TAD-Bd1, TAD-Bd2 (TAD-Bd2a, R1B, and TAD-Bd2b), and TAD-Bd3] are shown. Blue and red arrowheads indicate preferred interaction direction of CTCF motifs. See also fig. S5.
Fig. 7. Comparative analysis of R-TEn, R1B, and R2B enhancers among vertebrates. (A) Summary of the conservation of R-TEn (peak1 and peak2), R1B, and R2B (peak1 and peak2) among vertebrates, as shown in fig. S1 (B and C). Red, blue, and aqua dot lines indicate the border between Placentalia and Marsupialia (red), mammal and bird (blue), and reptile and amphibia (aqua). (B) Sequence logos showing conserved motifs in R1B, R2B, and R-TEn regions. Red bars indicate the E-box and Ikaros binding motifs. (C) Top panel shows genome browser view showing the sequence conservation of the R-TEn-peak1 regions among indicated species and summary of E-box motifs conservation among vertebrates, as seen in fig. S7D. Red boxes indicate seven E-box sequences in R-TEn-peak1 region in mouse (E1 to E7). Black regions indicate genomic regions similar to human sequences. Middle and bottom panels show the DNA sequences in E-box2 to E-box5 (E2 to E5) among indicated species. Conservation indicates the DNA sequence conservations among 100 vertebrates. See also fig. S7.
(DN3 and pro-B cells) without enhancer or promoter activity (E-box mutations in R-TEn or R1pro), respectively. However, the accessibility at the Rag1/2 promoter regions depends on the enhancer activity at the precursor stage (DP and pre-B cells), suggesting that the enhancer activity maintains the open status of the promoter regions through rounds of cell division after β-selection and pre-B cell receptor selection. Consistent with this, we noticed that the level of E2A occupancy at R-TEn did not decline after pre-TCR signaling to likely maintain R-TEn activity during DN-to-DP transition, whereas other E2A binding sites disappeared after β-selection (25).

Of note, the observed differences in TCRβ D-J recombination and γδT cell development between the R-TEnmut and R1pro-E-boxmut DN3 cells suggest that the low level of Rag1 expression induced by E2A-mediated R1pro activity in T progenitors is sufficient to permit TCRβ D-J and TCRγ/TCRδ recombination, while the high level of Rag expression induced by enhancer is required for long-range V-DJ recombination (i.e., a stepwise Rag regulation; fig. S4L). This is in line with the previous report that recombination of TCRγ/δ occurs concurrently with TCRβ D-J recombination at the DN2 stage, preceding TCRβ V-to-DJ recombination (51).

We observed that blocking E2A binding to these cis-regulatory elements was sufficient to disrupt the 3D genome organization of the Rag1/2 SE regions. This is in line with the disappearance of chromatin interactions in Rag gene loci in mature T and B cells, which express a low level of E2A and a high level of E protein antagonist Id3 or Id2 (fig. S6A) (52). This vulnerability of SE structure also seems to be related to the specificity of Rag1/2 genes. The Rag1/2 genes must be exclusively regulated during the adaptive lymphocyte development because the endonuclease activity of Rag proteins poses a threat to genome integrity.

Upon T lineage commitment, regulome and chromatin architectures are reorganized by T cell–specific TFs and noncoding transcripts at enhancer regions (28, 37, 53). Here, we show that the E protein binding at R-TEn is critically required for CTCF and cohesin binding across the Rag1/2 gene cluster and for the cohesin binding at the sub-TAD1 boundary regions (TAD-Bd1 and TAD-Bd2). A “super” or “stripe” anchor is defined as a loop anchor that interacts with an entire domain at a high frequency to drive loop extrusion (26, 54). Because TAD-Bd1 interacts with the whole sub-TAD1 region and contains many CTCF and cohesin binding sites, it likely functions as a stripe anchor to tether R-TEn to the Rag1 and Rag2 promoters. Because stripe anchors are closely associated with hypomethylated regulatory regions (54), indeed, R-TEn and other regulatory regions are hypomethylated at these progenitor/precursor stages.

First, we propose the following two steps for the Rag SE formation. First, upon adaptive lymphocyte lineage commitment, E2A binds to and activates enhancers and promoters by recruiting CBP/P300 and TET proteins to initiate Rag1/2 gene expression. Second, these regions interact with each other (enhancer-promoter communication) and with the anchor regions to facilitate local compaction through the recruitment of CTCF and cohesin complex. This, in turn, leads to the elaborate SE formation to induce a sufficiently high level of Rag1/2 expression, which is likely maintained by other factors such as Satb1 (Fig. 5F) (15). Of note, some of the observed chromatin interactions do not appear to be involved in the communication between the regulatory elements.

MATERIALS AND METHODS

Study design

Cell type–specific gene expression results from the interplay between lineage-specific TFs and 3D genome architecture. The aim of this study was to clarify how adaptive lymphocyte–specific Rag1/2 gene expression is initiated and regulated by TFs. The regulatory mechanism of Rag1/2 gene expression typically reflects the cell fate decision of adaptive lymphocyte lineage and is an ideal candidate for the analysis of cell type–specific gene expression. However, the depletion of a TF or a chromatin regulator may compromise genome-wide gene expression, making it difficult to distinguish between direct and indirect effects. To address this, we investigated the chromatin biology of a gene locus that is regulated by cell type–specific elements during differentiation and the functional consequences of enhancer knockout and TF binding motif mutations in mice. We analyzed and collected cells from 4-week-old all mutant mice. When possible, we used littermate control mice as a control. Most of studies were repeated at least twice or using more than two biological replicates. We performed in situ Hi-C using DP and pro-B cells one time with several preliminary experiments.

Mice

All CRISPR-Cas9–mediated knockout or mutation knock-in mice (R-TEnmut, R1Bmut, R2Bmut, R-TEn-p1mut, R-TEn-p2mut, R-TEn-p1-Eboxmut, and R1pro-Eboxmut) were generated in the laboratory of Integrative Biological Science, Institute for Frontier Life and Medical Sciences, Kyoto University. Tcf3fl/fl and Tcf1fl/fl were previously described (55). CD4Cre mice were obtained from the Jackson laboratory. All mice were C57BL/6 background and bred and housed in specific pathogen–free conditions under the Institutional Animal Care and Use Guidelines of Kyoto University. There was no difference in T and B cell development between males and females of all mutant mice. In most of the mouse experiments, we used 4-week-old mice, except for the analysis of fetal thymocytes at 18.5 day post-coitum (d.p.c.).

Generation of mouse lines by CRISPR-Cas9–mediated gene editing

Guide RNAs were designed using the CRISPRdirect website (https://crispr.dbcls.jp/) (56). The synthesized CRISPR RNAs (crRNAs), trans-activating crRNA (tracrRNA), and Cas9 protein were microinjected into the cytoplasm or pronucleus of fertilized eggs obtained from C57BL/6 mice to generate R-TEnmut, R1Bmut, R2Bmut, R-TEn-p1mut, R-TEn-p2mut, and R-TEn-p1-Eboxmut mice. The following genome sequences in mm9 coordinates were deleted: R-TEnmut, chr2:101391187–101393046; R1Bmut, chr2:101505041–101506422; R2Bmut, chr2:101437824–101441436; R-TEn-p1mut, R1Bmut, R2Bmut, R-TEn-p2mut, and R-TEn-p1-Eboxmut mice. The following genome sequences in mm9 coordinates were deleted: R-TEnmut, chr2:101391187–101393046; R1Bmut, chr2:101505041–101506422; R2Bmut, chr2:101437824–101441436 and chr2:101505041–101506422; R-TEn-p1mut, chr2:101391187–101393210; and R-TEn-p2mut, chr2:101392103–101393046. To generate E-box mutant mice (R-TEn-p1-Eboxmut and R1pro-Eboxmut), repair templates including the mutations of seven individual E-boxes in the R-TEn region or R1pro region, crRNAs, tracrRNA, and Cas9 protein were microinjected into C57BL/6 zygotes. Polymerase chain reaction (PCR) of genomic DNA detected alleles carrying the deletion or mutations. Mouse lines were established from founder mice carrying deletions or mutations by mating with C57BL/6 syngeneic mice.
CpG methylation analysis by bisulfite sequencing
Genomic DNA was prepared with a PureLink Genomic DNA Mini Kit (Invitrogen) and then treated with sodium bisulfite (EZ DNA Methylation-Gold Kit, Zymo Research). PCR amplified the modified DNA. The PCR products were subcloned into pTA2 vectors (TOYOBO). Purified DNA from the colonies (10 to 12 colonies per region) were sequenced.

Assay for transposase-accessible chromatin (ATAC-seq)
ATAC-seq was performed as described (23, 31). Briefly, sorted cells were lysed, and the nuclear pellets were incubated in 25 μl of reaction buffer containing 2 μl of Tn5 transposase and 12.5 μl of Tagment DNA buffer (Nextera Sample preparation kit from Illumina) at 37°C for 1 hour. After the incubation, the tagmented DNA was purified with a DNA Clean and Concentrator kit (Zymo Research). For preparing libraries, the purified DNA were amplified with determined cycles of PCR based on the amplification curve, and then the PCR products were size-selected using AMPure XP beads to remove large DNA fragments. The libraries were sequenced on a HiSeq 2500 sequencer (Illumina) with paired-end mode. Reads were aligned to reference genome mm9 using Bowtie2 (57). Tags that mapped to unique DNA sequences were analyzed further using HOMER (58). Tag directories were generated with makeTagDirectory command. The makeUCSCfile program generated sequencing depth–normalized bedGraph files to visualize the ATAC-seq signals on the University of California, Santa Cruz (UCSC) Genome Browser (59). ATAC-seq peaks were called using findPeaks program using default parameters. The replicates were analyzed separately, and the overlapped ATAC-seq peaks extracted using mergePeaks.pl command were used for statistical analysis. Differential ATAC-seq peaks between control and mutant mice at the Rag gene loci were statistically calculated using the HOMER tool getDifferentialPeaks program to identify regions with greater than twofold normalized read coverage and P < 1 × 10−7 (Poisson).
To compare abundance of ATAC reads at the Rag enhancer peaks across the different experiments, normalized tag densities were calculated by HOMER’s annotatePeaks.pl program with option -size 200.

Hi-C
In situ Hi-C was performed using DP and pro-B cells from control and deletion mutant mice as described (32). Hi-C libraries were size selected to 400 to 600 base pairs using ProNex (Promega) and sequenced paired-ended on a HiSeq 2500 (Illumina). Hi-C reads were trimmed from the 3’ ends of sequences to GATC (Mbo I restriction enzyme site) and aligned to the mouse reference genome mm9 using bowtie2. Only read pairs that were mapped uniquely were used for further analysis. All PCR duplicates were also removed. Subsequent Hi-C analysis was performed using HOMER as described (22, 34). Hi-C matrixes were visualized using Java TreeView (60). Hi-C compartment analysis was performed with HOMER’s runHiCpca.pl program. Circos diagrams were generated by Circos software (61) using HOMER’s analyzeHiC program with -circos option. In chromatin compaction analysis, distal-to-local log2 ratio (DLR) of Hi-C interactions at a locus was calculated in 5-kb resolution at 15-kb window by HOMER as described (34). To find TADs and chromatin loops, findTADsAndLoops.pl program was run in 3-kb resolution at 15-kb window by HOMER. The program generated a bedGraph file that describes the directionality index called TADs and Loops score in this paper.

TCRβ, IgH, and IgL TCRα recombination
Genomic DNA was purified from sorted DN3a, CD69- DP, pro-B, and pre-B cells. Semi-quantitative analysis of IgH and TCRβ was performed by KOD FX Neo (TOYOBO). Quantitative real-time PCR of IgL and TCRα was performed by TB Green Premix Ex Taq (Takara).

Sequence analysis
Genomic sequence data for the following species were downloaded from the National Center for Biotechnology Information FTP site (ftp://ftp.ncbi.nlm.nih.gov/genomes): Homo sapiens (human), Mus musculus (house mouse), Dasyus novemcinctus (nine-banded armadillo), Sarcopilus harrisii (Tasmanian devil), Gallus gallus (chicken), Chrysemys picta (western painted turtle), Xenopus tropicalis (tropical clawed frog), Danio rerio (zebrafish), and Callorhinchus milii (elephant shark). The region containing Rag1 and its upstream sequence up to the start of Traf6 (i.e., Rag1 upstream region) and the 100-kb region containing Rag2 and its upstream sequence (i.e., Rag2 upstream region) were extracted from each of the genomic sequences. The dot matrix for the Rag1 and Rag2 upstream regions were generated by an in-house script using BLASTN (62). MEME Suite (63) was used to identify the conserved motifs in the Rag1 and Rag2 upstream regions and to generate sequence logos.

Quantification and statistical analysis
Microsoft Excel software and Prism packages calculated P values for two-tailed Student’s test in two-group comparison. The statistical significance level was 0.05.

SUPPLEMENTARY MATERIALS
immunology.sciencemag.org/cgi/content/full/5/51/eabb1455/DC1
Methods
Fig. S1. Characterization of R-TEn and R1B/R2B.
Fig. S2. R-TEn and R1B/R2B are essential for cell type–specific Rag1/Rag2 expression.
Fig. S3. E-box mutations in Rag1/Rag2 and deletion mutant mice as described (5). Hi-C libraries were size selected to 400 to 600 base pairs using ProNex (Promega) and sequenced paired-ended on a HiSeq 2500 (Illumina). Hi-C reads were trimmed from the 3’ ends of sequences to GATC (Mbo I restriction enzyme site) and aligned to the mouse reference genome mm9 using bowtie2. Only read pairs that were mapped uniquely were used for further analysis. All PCR duplicates were also removed. Subsequent Hi-C analysis was performed using HOMER as described (22, 34). Hi-C matrixes were visualized using Java TreeView (60). Hi-C compartment analysis was performed with HOMER’s runHiCpca.pl program. Circos diagrams were generated by Circos software (61) using HOMER’s analyzeHiC program with -circos option. In chromatin compaction analysis, distal-to-local log2 ratio (DLR) of Hi-C interactions at a locus was calculated in 5-kb resolution at 15-kb window by HOMER as described (34). To find TADs and chromatin loops, findTADsAndLoops.pl program was run in 3-kb resolution at 15-kb window by HOMER. The program generated a bedGraph file that describes the directionality index called TADs and Loops score in this paper.

REFERENCES AND NOTES


Acknowledgments: We thank C. Murre for insightful suggestions and for reviewing the manuscript; H. Kawamoto and D. Schatz for helpful discussion; and M. Nakamura, N. Kirihata, T. Kondo, and Y. Sando (Kyoto University) for technical support. We thank Y. Zhuang (Duke University) for Tcf7L2 and Tcf12/ mice. Funding: This work was funded by the KARENHI (Grants-in-Aid for Scientific Research) from the MEXT of Japan (16H05205 and 19H03487 for M.M., 17K08629 for K.M., and 19H05656 for S.O.), the High Performance Computing Infrastructure System Research Project (hp190158 for S.O.), the Japan Agency for Medical Research and Development (AMED) (JP19cm0106501h0004 and JP19gm1110010h0001 for S.O.), the Mochida Memorial Foundation (M.M.), and the Takeda Science Foundation (M.M.). Computational time was in part provided by the Supercomputer System at the institute for Chemical Research, Kyoto University and by the Human Genome Center at the Institute of Medical Science of the University of Tokyo. Author contributions: M.M. and K.M. conceived the project and performed the majority of experiments and analyses. H.W., R.T., and G.K. generated deletion and mutant mouse lines. G.Y. and H.O. performed genomic sequencing. Y.O., S.T., and T.U. contributed to experiments and provided technical advice. O.T., Y.A., and K.O. contributed to gene expressions and 3C and methylation analyses. Y.O. and S.O. contributed to the DNA analysis. K.M. and K.C. performed bioinformatic analysis. R.H., Y.A., and K.O. contributed to the manuscript; H.W., R.T., and G.K. wrote the manuscript. Competing interests: The authors declare that they have no competing interests. Data and materials availability: Data for ChiP-seq, ATAC-seq, and in situ Hi-C have been deposited to the National Center for Biotechnology Information Gene Expression Omnibus (GSE141223). All (other) data needed to evaluate the conclusions in the paper are present in the paper or the Supplementary Materials. All mice are either commercially available or available under a material transfer agreement (MTA).

Submitted 9 February 2020
Accepted 21 July 2020
Published 4 September 2020
10.1126/sciimmunol.abb1455

Citation: Miyazaki et al., Sci. Immunol. 5, eabb1455 (2020).
The transcription factor E2A activates multiple enhancers that drive Rag expression in developing T and B cells

Kazuko Miyazaki, Hitomi Watanabe, Genki Yoshikawa, Kenian Chen, Reiko Hidaka, Yuki Aitani, Kai Osawa, Rie Takeda, Yotaro Ochi, Shizue Tani-ichi, Takuya Uehata, Osamu Takeuchi, Koichi Ikuta, Seishi Ogawa, Gen Kondoh, Yin C. Lin, Hiroyuki Ogata and Masaki Miyazaki

DOI: 10.1126/sciimmunol.abb1455

To each their own
Recombination activating genes (RAGs) RAG1 and RAG2 play central roles in assembling functional T and B cell receptors in developing lymphocytes. Expression of Rag1 and Rag2 in hematopoiesis is restricted to these two lymphoid lineages, but precisely how this is accomplished has remained a mystery. Here, Miyazaki et al. have identified three key enhancer elements that recruit the transcription factor E2A to promote the expression of Rag1 and Rag2 genes during lymphocyte development. By generating mouse strains lacking one or more of these enhancer elements, they report that T and B cells use distinct enhancer modules to activate and maintain expression of Rag1 and Rag2 genes.