

ARTICLE

DOI: 10.1038/s41467-017-00768-1

OPEN

# Priming of lineage-specifying genes by Bcl11b is required for lineage choice in post-selection thymocytes

Satoshi Kojo<sup>1</sup>, Hirokazu Tanaka<sup>1</sup>, Takaho A. Endo<sup>2</sup>, Sawako Muroi<sup>1</sup>, Ye Liu<sup>3</sup>, Wooseok Seo<sup>1</sup>, Mari Tenno<sup>1</sup>, Kiyokazu Kakugawa<sup>1</sup>, Yoshinori Naoe<sup>1</sup>, Krutula Nair<sup>1</sup>, Kazuyo Moro<sup>4</sup>, Yoshinori Katsuragi<sup>5</sup>, Akinori Kanai<sup>6</sup>, Toshiya Inaba<sup>6</sup>, Takeshi Egawa<sup>7</sup>, Byrappa Venkatesh<sup>8</sup>, Aki Minoda<sup>3</sup>, Ryo Kominami<sup>5</sup> & Ichiro Taniuchi<sup>1</sup>

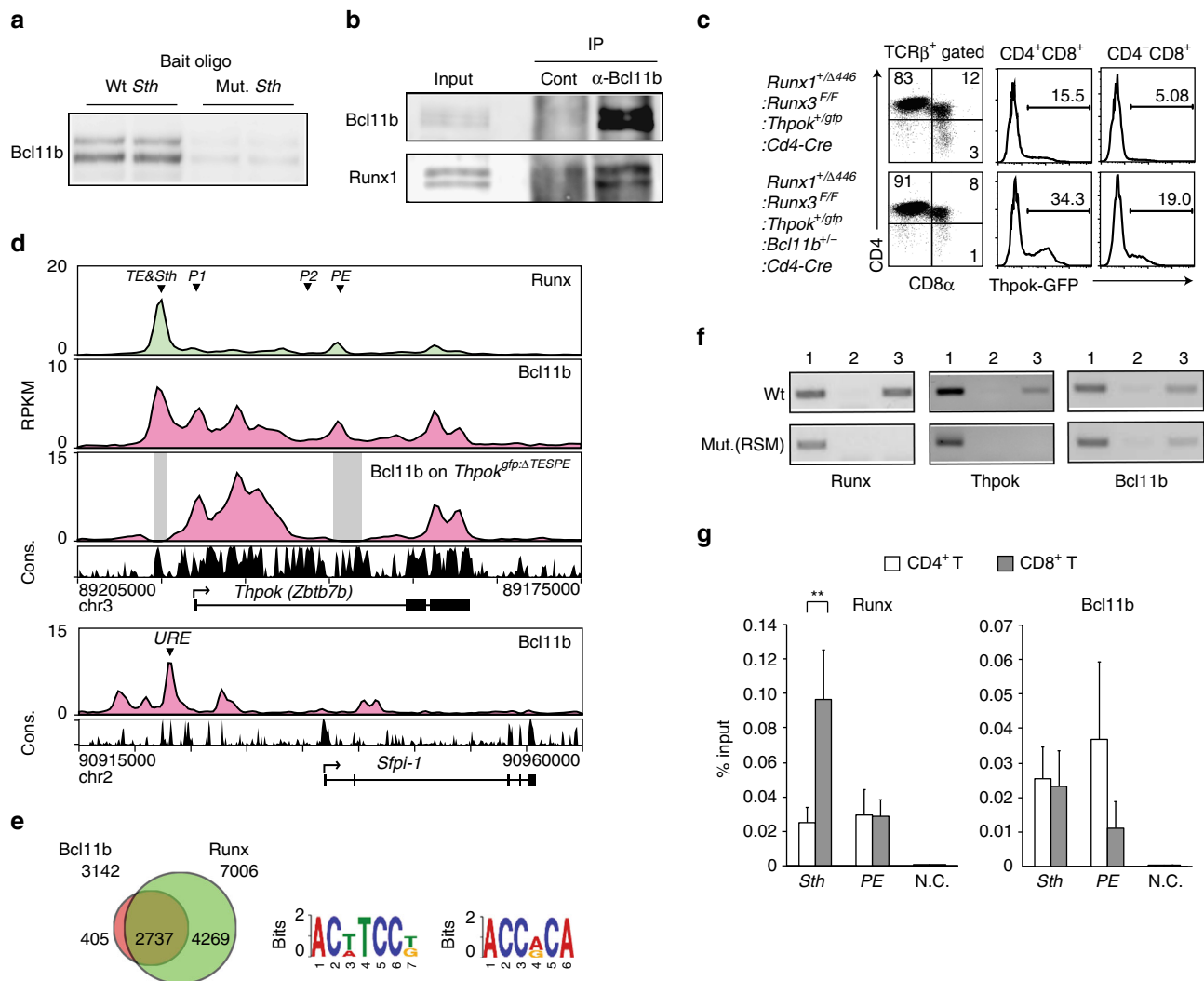
T-lineage committed precursor thymocytes are screened by a fate-determination process mediated via T cell receptor (TCR) signals for differentiation into distinct lineages. However, it remains unclear whether any antecedent event is required to couple TCR signals with the transcriptional program governing lineage decisions. Here we show that Bcl11b, known as a T-lineage commitment factor, is essential for proper expression of ThPOK and Runx3, central regulators for the CD4-helper/CD8-cytotoxic lineage choice. Loss of Bcl11b results in random expression of these factors and, thereby, lineage scrambling that is disconnected from TCR restriction by MHC. Initial *Thpok* repression by Bcl11b prior to the pre-selection stage is independent of a known silencer for *Thpok*, and requires the last zinc-finger motif in Bcl11b protein, which by contrast is dispensable for T-lineage commitment. Collectively, our findings shed new light on the function of Bcl11b in priming lineage-specifying genes to integrate TCR signals into subsequent transcriptional regulatory mechanisms.

<sup>1</sup>Laboratory for Transcriptional Regulation, RIKEN Center for Integrative Medical Sciences (IMS), 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama 230-0045, Japan. <sup>2</sup>Laboratory for Integrative Genomics, RIKEN Center for Integrative Medical Sciences (IMS), 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama 230-0045, Japan. <sup>3</sup>Division of Genomic Technologies, RIKEN Center for Life Science Technologies (CLST), 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama 230-0045, Japan. <sup>4</sup>Laboratory for Innate Immune Systems, RIKEN Center for Integrative Medical Sciences (IMS), 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama 230-0045, Japan. <sup>5</sup>Division of Molecular Biology, Department of Molecular Genetics, Graduate School of Medical and Dental Sciences, Niigata University, Niigata 951-8510, Japan. <sup>6</sup>Department of Molecular Oncology, Research Institute for Radiation Biology and Medicine, Hiroshima University, 1-2-3, Kasumi, Minami-ku, Hiroshima 734-8553, Japan. <sup>7</sup>Department of Pathology and Immunology, School of Medicine, Washington University School of Medicine, 660 S Euclid, Saint Louis 63110 MO, USA. <sup>8</sup>Institute of Molecular and Cell Biology, Agency for Science, Technology and Research, Biopolis 138673, Singapore. Correspondence and requests for materials should be addressed to I.T. (email: [ichiro.taniuchi@riken.jp](mailto:ichiro.taniuchi@riken.jp))

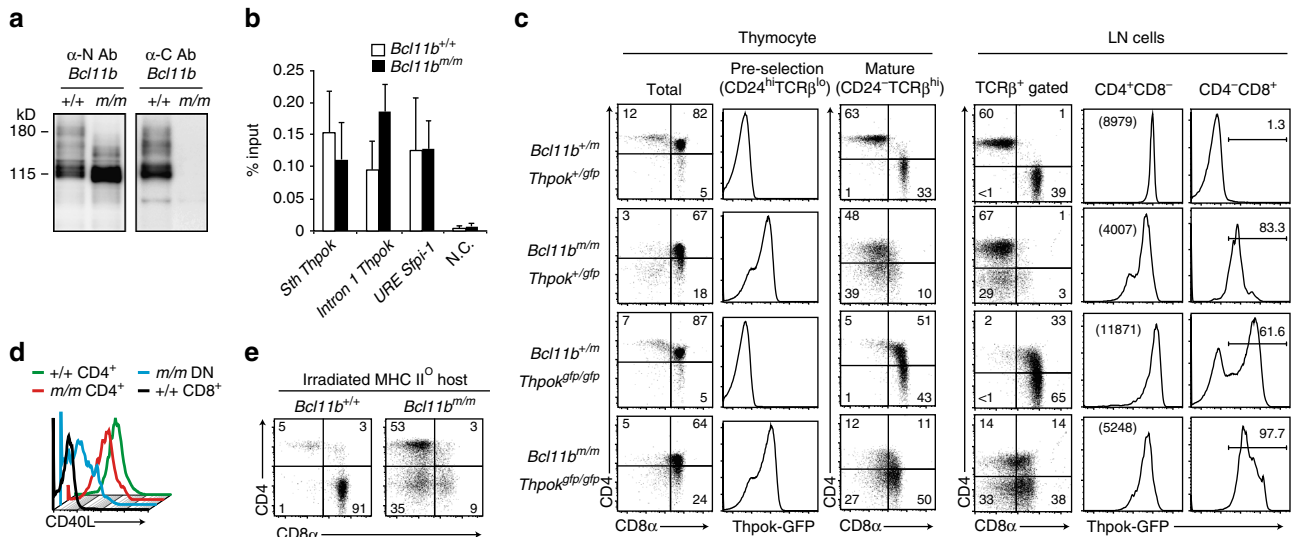
Early thymocyte progenitors (ETPs) retain their developmental potential to become non-T- lymphoid cells, but, upon exposure to the thymic microenvironment, the ETP expression program is dramatically altered to commit them to the T-lymphoid lineage<sup>1,2</sup>. Previous studies demonstrated that a final T-lymphoid lineage commitment occurs at a developmental checkpoint during the transition from the CD4<sup>-</sup>CD8<sup>-</sup> double negative (DN)2a to the DN2b stage, where expression of the transcription factor Bcl11b is induced<sup>3,4</sup>. Bcl11b-deficiency results in a developmental arrest of early T cell progenitors at the DN2a stage along with a concomitant acquisition of myeloid-lineage/natural killer (NK) cell gene signatures<sup>5</sup>. Bcl11b thus serves as a T-lineage commitment factor that eliminates developmental potential to non-T lymphoid cells at the DN2

checkpoint<sup>3-5</sup>. Beside its function during early thymocyte development, Bcl11b continues to be expressed during T lymphocyte differentiation<sup>5,6</sup> and modulates the development of many lymphoid subsets, including natural killer T (NKT)<sup>7</sup>, regulatory T (Treg)<sup>8</sup> and type2 innate lymphoid (ILC2) cells<sup>9-11</sup>.

After commitment to the T-lineage, another checkpoint, known as  $\beta$ -selection, selectively expands DN3 cells that successfully express functional TCR $\beta$  chain after V(D)J recombination of the *Tcrb* locus. Thymocytes that have passed  $\beta$ -selection become  $\alpha\beta$ T-lineage-restricted CD4<sup>+</sup>CD8<sup>+</sup> double positive (DP) precursor thymocytes, which express complete  $\alpha\beta$ TCR complexes for the first time during T cell development. DP precursors are then subjected to additional positive and negative selections that enrich precursors with TCRs recognizing



**Fig. 1** Binding of Bcl11b to regulatory regions in the *Thpok* gene. **a** DNA pull-down assay showing in vitro Bcl11b binding to wild-type (Wt) *Thpok* silencer (*Sth*) core sequences, but not efficiently to mutant (Mut.) sequences. One representative of two experiments. **b** Co-immune precipitation assay showing interaction of Bcl11b with Runx1. One representative of two experiments. **c** Flow cytometry showing an increase of CD8<sup>+</sup> T cells de-repressing Thpok-GFP upon reduction of *Bcl11b* dosage in *Runx* mutant mice. One representative of two independent experiments. **d** Bcl11b ChIP-seq tracks at the *Thpok*, *Thpok*<sup>gfp:ΔTESPE</sup>, and *Sfp1* genes in total thymocytes along with Runx ChIP-seq (GSE90794) tracks at the *Thpok* gene (top) for reference. Gene structure, transcriptional orientation and conservations in mammals (Cons.) in the *Thpok* gene are indicated. Positions of *Thpok* silencer (*Sth*), two enhancers (*TE* and *PE*), distal P1-promoter (*P1*) and proximal enhancer (*PE*) in the *Thpok* gene, and upstream regulatory element (*URE*) in the *Sfp1* gene are indicated as arrowheads. **e** Co-occupancy by Runx at Bcl11b-bound genome regions. Runx recognition site (5'-ACCPuCA-3') was listed among the top two common sequences for Bcl11b-bound regions. **f** ChIP-PCR assay for binding of Runx, Bcl11b, and ThPOK to Wt and Runx site mutated (RSM) *Sth* regions in CD4<sup>+</sup> T cells from *Thpok*<sup>+/gfp:241-401RM</sup> mice. Lanes 1, 2, and 3 are input, control IgG and antibody-against the indicated transcription factor, respectively. One representative of two independent experiments. **g** ChIP-qPCR analyses showing binding of Runx and Bcl11b to the *Sth* and *PE* in CD4<sup>+</sup> (white bars) and CD8<sup>+</sup> (gray bars) T cells. Combined data from three independent ChIP experiments is shown. \*\**P* < 0.01 (unpaired *t*-test)



**Fig. 2** CD4-skewed development by impaired *Thpok* regulation from hypomorphic *Bcl11b*<sup>m/m</sup> progenitors. **a** Expression of truncated Bcl11b protein in the neonatal thymus from the hypomorphic *Bcl11b*<sup>m</sup> allele detected by an antibody recognizing the N- (left), but not the C- (right), terminus of Bcl11b. One representative of two experiments. **b** Summary of ChIP assay for binding of Wt (white bars) and hypomorphic (black bars) Bcl11b to indicated regions in total neonatal thymocytes. **c–e** Flow cytometry analyzing the expression of CD4, CD8, and Thpok-GFP during T cell development in Rag1-deficient hosts **c** and irradiated MHC-II deficient hosts **e** reconstituted with fetal liver cells with the indicated genotypes. Data are representative of at least two independent experiments. Numbers in dot plots indicate the percentage of cells in each quadrant. Histograms showing CD40L expression 2 days after in vitro TCR stimulation of peripheral CD4<sup>+</sup>, CD8<sup>+</sup> and CD4<sup>-</sup>CD8<sup>-</sup> DN cells differentiated in Rag1-deficient mice from *Bcl11b*<sup>+/+</sup> and *Bcl11b*<sup>m/m</sup> fetal liver cells **d**. Data are representative of two independent experiments

antigen-MHC complexes with intermediate affinity but eliminate precursors expressing self-reactive TCRs, respectively<sup>12</sup>. Positively selected thymocytes differentiate into distinct T cell subsets with distinct functionalities via the activation of specific developmental programs. For instance, positive selection signaled via TCR engagement by MHC class-II (MHC-II) and class-I (MHC-I) guides precursors to differentiate into CD4<sup>+</sup>CD8<sup>-</sup> helper or CD4<sup>-</sup>CD8<sup>+</sup> cytotoxic T cells through the induction of key transcription factors, ThPOK and Runx3, respectively<sup>13, 14</sup>. Thus, DP precursors must be ready to integrate TCR signals, translating them into the appropriate developmental program. However, an important gap in our understanding of these processes is how TCR signals are coupled to mechanisms that control the expression of lineage-specifying genes, and it remains unclear whether preceding events are required for this coupling.

One such lineage-specifying transcription factor, Zbtb7b, also known as T-helper-inducing POZ/Krüppel-like factor (ThPOK), is a member of the BTB-POZ zinc-finger transcription factor family<sup>15</sup> and is encoded by *Zbtb7b*, hereafter referred to as the *Thpok* gene. Previous genetic studies for ‘gain and loss’ of ThPOK function demonstrated that its presence or absence in post-selection thymocytes is a major determinant of the CD4-helper (ThPOK<sup>+</sup>) versus CD8-cytotoxic (ThPOK<sup>-</sup>) lineage dichotomy<sup>16–18</sup>. Consistent with these findings, expression of the *Thpok* gene is restricted to MHC-II selected thymocytes<sup>16</sup> in positively selected thymocytes. Accordingly, *Thpok* regulation has been recognized as an ideal model to study how TCR signals are coupled with the transcriptional program that establishes the identity of CD4<sup>+</sup> helper T cells. Such studies identified a transcriptional silencer in *Thpok*, hereafter denoted as *Thpok* silencer (*Sth*), which is an essential cis-acting element restricting *Thpok* expression to post-selection thymocytes in the helper lineage<sup>19, 20</sup>. In addition to *Sth*, there are at least three enhancers in the *Thpok* locus<sup>20</sup>. Among them, the thymic enhancer (*TE*), located upstream of the *Sth*, acts first to initiate *Thpok* expression<sup>21</sup>, which is subsequently upregulated through the activity of a

proximal enhancer (*PE*) locating 1.8 kb downstream of the proximal P2-promoter<sup>22</sup>. While factors that regulate *Sth* activity, such as Runx family proteins<sup>19</sup> and MAZR<sup>23, 24</sup>, have been identified, the factors involved in the activation of *TE* and *PE* remain poorly characterized. *Gata3*<sup>25</sup> and *Tcf1/Lef1*<sup>26</sup> regulate *Thpok* expression, but primarily do so by targeting other regulatory regions such as general T-lymphoid enhancer, the known third enhancer. In contrast to *Thpok* regulation, very little is known about transcription factors that orchestrate CD8<sup>+</sup> T cell-specific expression of *Runx3* from its distal P1-promoter<sup>27</sup>. Signals emanating from the IL-7 cytokine receptor have been shown to activate *Runx3*<sup>28, 29</sup>; however, the intermediaries and their cis-regulatory targets in *Runx3* remain to be established<sup>30</sup>.

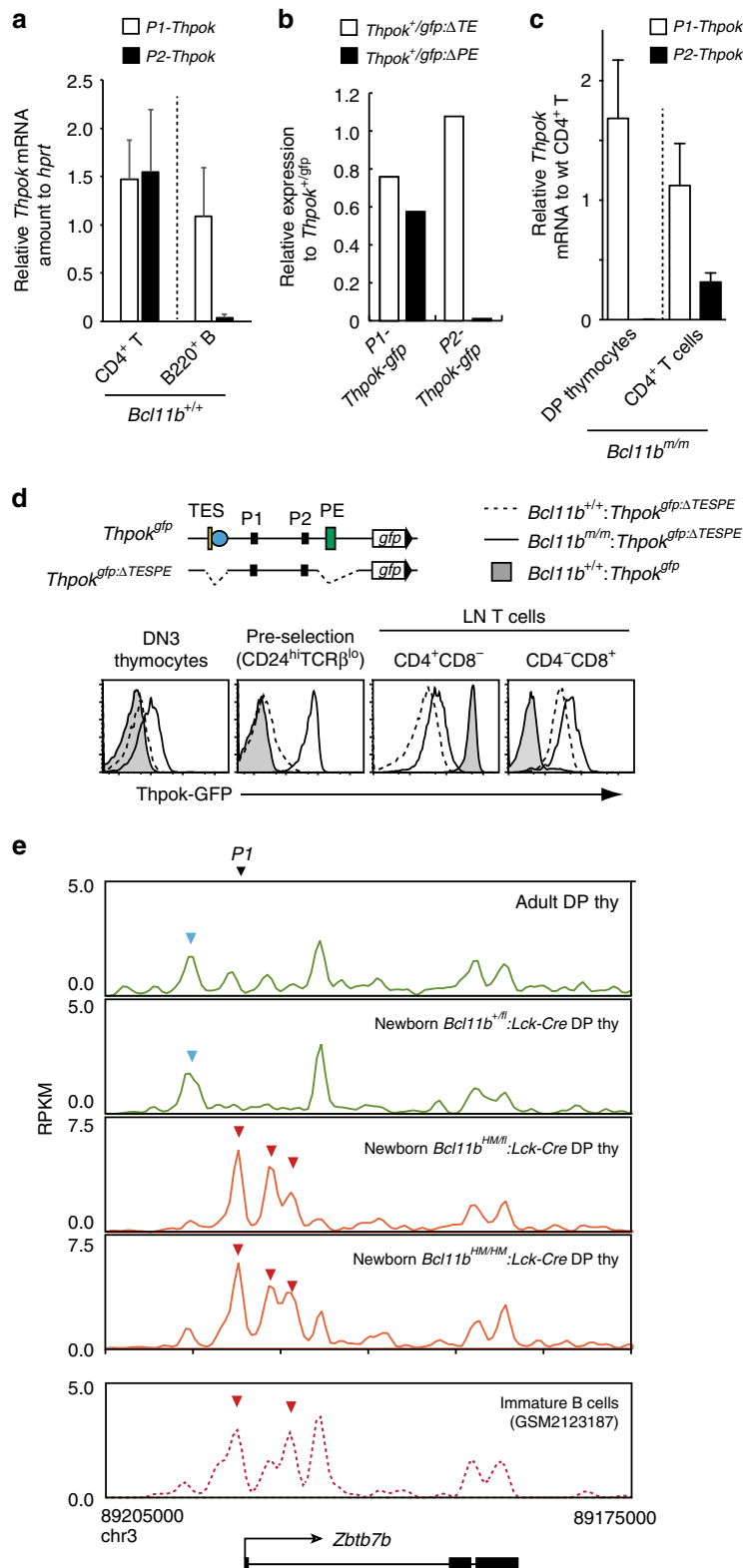
Here we report two mechanisms by which Bcl11b governs the transcriptional program dissecting helper versus cytotoxic lineage commitment: *Sth*-independent repression of *Thpok* in DN thymocytes and enhancer-dependent *Runx3* repression in CD4-lineage cells. Deletion of Bcl11b in thymocytes at post-β-selection stage causes chaotic *Thpok* and *Runx3* expression, inducing random differentiation of both MHC-I and MHC-II selected cells into the helper and cytotoxic subsets. Along with earlier requirement for Bcl11b prior to DP stage in later *Foxp3* activation, we conclude that lineage-specifying genes are ‘primed’ by Bcl11b before or during transition to the DP stage to represent an essential event for coupling TCR signals to expression programs for differentiating into the appropriate T-effector subsets.

## Results

**Bcl11b binds to the *Thpok* locus.** Based on our assumption that proteins bound to *Thpok* silencer (*Sth*) should mediate coupling of TCR signals during positive selection with release of *Thpok* expression, we attempted to purify protein complexes by in vitro capture with an oligo-nucleotide harboring core *Sth* sequences. Consistent with prior chromatin immunoprecipitation (ChIP) assays<sup>31</sup>, Bcl11b protein was efficiently enriched by affinity

purification with core *Sth* sequences, but to lesser extent with mutant *Sth* sequences (Fig. 1a). Bcl11b also associated with other known *Sth* binding proteins, Runx1 (Fig. 1b). Furthermore, reduction of *Bcl11b* dosage to the half (*Bcl11b*<sup>+/-</sup>) in the *Runx* mutant mice, in which combined mutations of *Runx1* and *Runx3* genes attenuated *Sth*-mediated *Thpok* repression<sup>32</sup>, resulted in an increase of CD8<sup>+</sup> subset de-repressing a *Thpok*<sup>gfp</sup> reporter<sup>22</sup> (Fig. 1c), indicating genetic interaction between two factors in the

regulation of *Sth* function. Our ChIP sequencing (ChIP-seq) assay of total thymocytes also detected that Bcl11b associated with *Sth* and proximal enhancer (PE) as well as intronic regions downstream from distal P1-promoter in the *Thpok* locus (Fig. 1d). Bcl11b also bound to an upstream regulatory element (URE)<sup>33</sup> in the *Sfpi-1* locus, which encodes a myeloid transcription factor, PU.1, and is a putative target of Bcl11b to eliminate myeloid potential during T-lineage commitment (Fig. 1d). A more



global analysis revealed the Runx recognition motif (5'-ACCPuCA-3') as the second-ranked sequence enriched at Bcl11b-bound regions in thymocytes (Fig. 1e) and Runx bindings coincided with ~85 % of Bcl11b-bound regions (Fig. 1e), which is consistent with previous observation using neuronal cells<sup>34</sup>. However, Bcl11b binding to *Sth* was Runx-independent, as determined using primary CD4<sup>+</sup> T cells in which two Runx sites within the *Sth* were mutated by a knock-in on the *Thpok*<sup>gfp</sup> reporter allele<sup>35</sup>, whereas ThPOK binding to the *Thpok* silencer requires Runx-binding (Fig. 1f). Our analytical ChIP assay in CD4<sup>+</sup> and CD8<sup>+</sup> T cell subset detected that Bcl11b binding at *PE* was greater in *Thpok*-expressing CD4<sup>+</sup> T cells, while fivefold more Runx was bound to *Sth* in ThPOK-negative cells (Fig. 1g). These results suggest that Bcl11b directly regulates activity of *Thpok* locus through binding to the silencer and enhancer elements.

**Impaired *Thpok* regulation in *Bcl11b* mutant mice.** Because there is a developmental arrest at the DN2a stage when Bcl11b function is completely lacking, it is necessary to utilize a *Bcl11b*<sup>fl</sup> allele to examine the function of Bcl11b during T cell development beyond the DN stage. We therefore attempted to generate our own *Bcl11b*<sup>fl</sup> allele. However, our first trial unexpectedly generated a mutant *Bcl11b* allele, hereafter referred as to a *Bcl11b*<sup>m</sup>, which produced a truncated Bcl11b protein due to a frameshift mutation incorporated into the targeting vector during its construction (Fig. 2a and Supplementary Fig. 1a–c). Despite its loss of the final zinc-finger motif, truncated mutant Bcl11b protein could bind to the *Sth* and an intronic region in the *Thpok*, as well as to the URE in *Sfpi-1* (Fig. 2b). Similar to *Bcl11b* null animals (*Bcl11b*<sup>-/-</sup>), homozygous *Bcl11b*<sup>m/m</sup> mice died as neonates (Supplementary Fig. 1d). However, in contrast to the null animals, which completely lack CD4<sup>+</sup>CD8<sup>+</sup> DP thymocytes and thereby have small thymus<sup>36</sup>, we noticed that total thymocyte number is only slightly reduced in *Bcl11b*<sup>m/m</sup> neonates (Supplementary Fig. 1e) and, more strikingly, the DP population was present in those mice (Supplementary Fig. 1f). In addition, *Bcl11b*<sup>m/m</sup> progenitors gave rise to ILC2 cells (Supplementary Fig. 1g), a subset that does not develop from *Bcl11b*<sup>-/-</sup> progenitors<sup>9, 10</sup>. These results indicated that the mutant *Bcl11b*<sup>m</sup> is a hypomorphic allele, thereby hereafter its product is referred to as hypomorphic Bcl11b protein (Bcl11b<sup>HM</sup>).

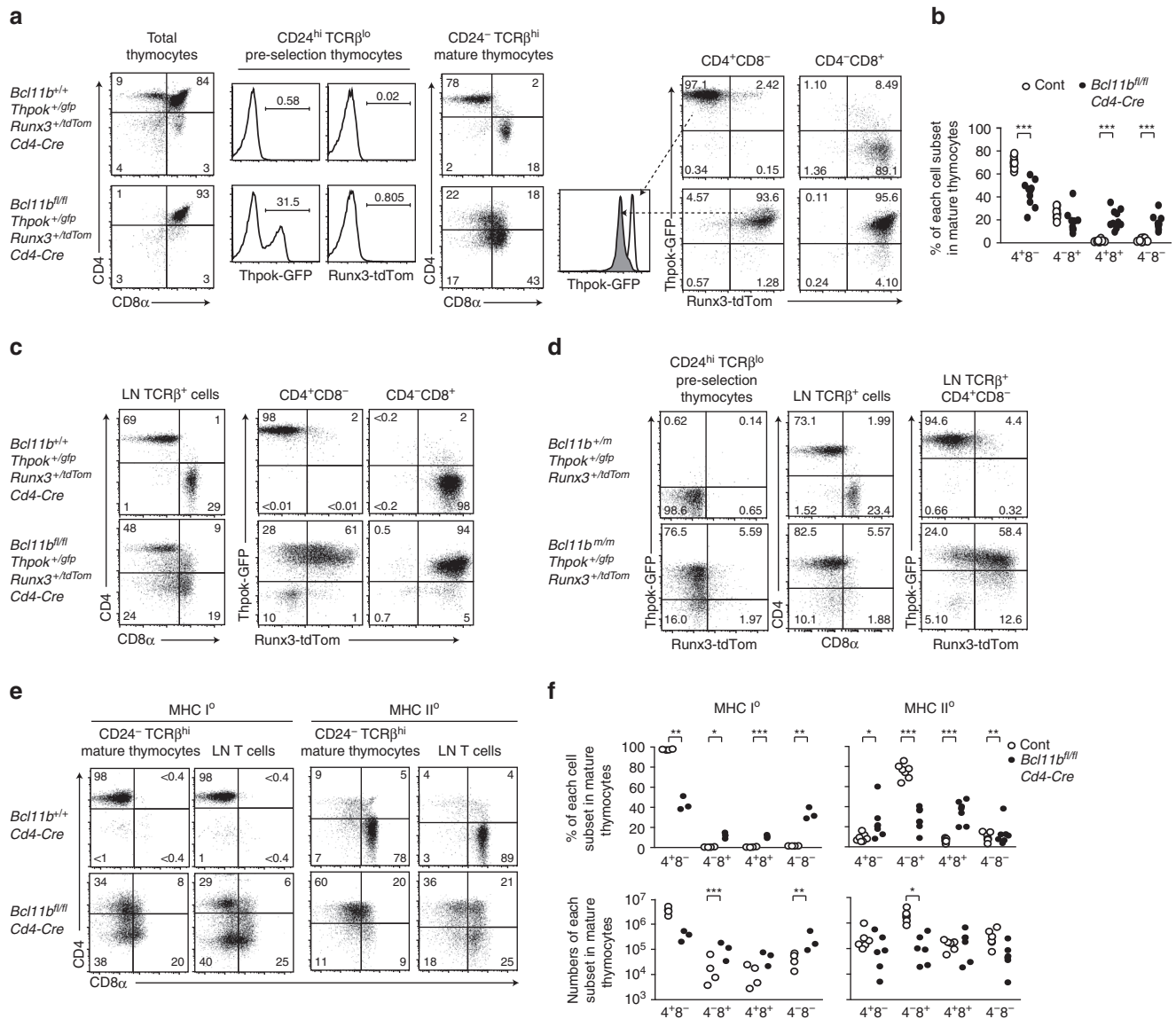
We next investigated the T cell development of *Bcl11b*<sup>m/m</sup> progenitors that also harbor a *Thpok*<sup>gfp</sup> reporter allele. For this purpose, we transferred *Bcl11b*<sup>m/m</sup>:*Thpok*<sup>+/gfp</sup> or control *Thpok*<sup>+/gfp</sup> fetal liver progenitors into T cell-deficient *Rag1*<sup>-/-</sup> hosts, allowing differentiation of the progenitors to proceed for two months. As shown in Fig. 2c, *Thpok*-GFP expression was absent in pre-selection thymocytes (defined as CD24<sup>hi</sup>TCRβ<sup>lo</sup>) from control cells. In sharp contrast, nearly all pre-selection thymocytes derived from the transferred *Bcl11b*<sup>m/m</sup> progenitors

expressed *Thpok*-GFP. While the two major T cell subsets, CD4<sup>+</sup> helper and CD8<sup>+</sup> cytotoxic, were differentiated from control *Bcl11b*<sup>+/+</sup> progenitors, *Bcl11b*<sup>m/m</sup> progenitors gave rise to predominantly CD4<sup>+</sup> T cells with a helper-related signature (e.g., CD40L induction) in mature thymocyte and peripheral T cell populations (Fig. 2c, d). To address whether a re-direction of MHC-I selected cells to the CD4<sup>+</sup> lineage contributed to this CD4-skewing, we generated chimeras using irradiated MHC-II deficient hosts expressing the Ly9.1 marker, in which only differentiation of MHC-I selected cells was supported. Strikingly, we continued to observe skewing of CD4 differentiation in Ly9.1<sup>-</sup> donor-derived populations from *Bcl11b*<sup>m/m</sup>, but not *Bcl11b*<sup>+/+</sup>, progenitors (Fig. 2e). The shift to CD4 dominant differentiation was not observed when *Bcl11b*<sup>m/m</sup>:*Thpok*<sup>gfp/gfp</sup> progenitors lacking ThPOK expression were transferred (Fig. 2c). Along with de-repression of *Thpok*-GFP in remaining CD8<sup>+</sup> T cells (Fig. 2c), we conclude that aberrant *Thpok* expression redirects MHC class I-selected *Bcl11b*<sup>m/m</sup> thymocytes to the CD4<sup>+</sup> lineage. After these initial characterizations of the hypomorphic *Bcl11b*<sup>m</sup> allele, we generated a true conditional mutant (*Bcl11b*<sup>fl</sup>) allele and a mutant allele but lacks the loxP sequence in its 3'UTR, which we referred to as the *Bcl11b*<sup>HM</sup> allele (Supplementary Fig. 1a).

**Silencer-independent *Thpok* repression by Bcl11b.** To further investigate how *Thpok* expression was de-regulated in *Bcl11b*<sup>m/m</sup> cells, we examined promoter usages. Prior studies have shown that *Thpok* is transcribed from distal P1- and proximal P2-promoters<sup>20</sup>, however, P2-derived transcripts were detected specifically in T-lineage cells and require *PE* activity (Fig. 3a, b). Given the loss of *Thpok*-GFP expression in peripheral T cells upon removal of both *TE* and *PE* (Supplementary Fig. 2a), *TE* is likely to drive P1-promoter activity in T cells. In contrast, *Thpok* transcript in *Bcl11b*<sup>m/m</sup> pre-selection thymocytes was transcribed only from the P1-promoter (Fig. 3c). To understand the underlying mechanisms for unusual P1-promoter activation without P2-promoter activation in *Bcl11b*<sup>m/m</sup> thymocytes, we examined *Thpok*-GFP expression from the *Thpok*<sup>gfp</sup>: $\Delta$ TESPE reporter allele, in which all three T cell-specific regulatory elements, *Sth*, *TE*, and *PE*, were removed from the *Thpok*<sup>gfp</sup> allele by sequential gene targeting (Fig. 3d). In Bcl11b wild-type cells, *Thpok* expression from the enhancer-deficient *Thpok*<sup>gfp</sup>: $\Delta$ TESPE reporter allele was derived solely from the P1-promoter (Supplementary Fig. 2b), and was detected only in cells after the post-selection stages (Fig. 3d). In contrast, despite deletion of the *Sth* silencer, *Thpok*-GFP expression from the *Thpok*<sup>gfp</sup>: $\Delta$ TESPE reporter allele was initiated in *Bcl11b*<sup>m/m</sup> thymocytes at DN3 stage and its expression level was increased during transition to DP stage (Fig. 3d). These data revealed the presence of undefined *Sth*-independent mechanism by which Bcl11b represses *Thpok* expression in DN3 as well as in DP thymocytes.

**Fig. 3** *Sth*-independent *Thpok* repression by Bcl11b in pre-selection DP thymocytes. **a** Quantitative RT-PCR showing amount of promoter-specific *Thpok* transcripts in splenic CD4<sup>+</sup> T cells and splenic B220<sup>+</sup> B lymphocytes. Summary of three measurements. **b** Quantitative RT-PCR for *Thpok*-gfp expression in CD4<sup>+</sup> T cells from mutant reporter alleles, *Thpok*<sup>gfp</sup>: $\Delta$ TE and *Thpok*<sup>gfp</sup>: $\Delta$ PE, lacking the thymic enhancer (*TE*) or proximal enhancer (*PE*), respectively. Levels of P1- and P2 promoter-specific transcripts relative to those from the control *Thpok*<sup>gfp</sup> allele harboring intact enhancers is shown. One representative of two independent experiments. **c** Promoter-specific *Thpok* transcripts in *Bcl11b*<sup>m/m</sup> pre-selection DP thymocytes and CD4<sup>+</sup> T cells relative to those in control CD4<sup>+</sup> T cells. Summary of three measurements. **d** Histograms showing *Thpok*-GFP expression from the *Thpok*<sup>gfp</sup>: $\Delta$ TESPE allele in the indicated cell subsets differentiated in *Rag1*-deficient recipients from control (dotted line) and *Bcl11b*<sup>m/m</sup> (solid line) fetal liver cells. Expression pattern of *Thpok*<sup>gfp</sup> is shown as reference (shaded). Schematic structure of the *Thpok*<sup>gfp</sup>: $\Delta$ TESPE allele is included (upper panel). One representative of two independent experiments. **e** ATAC-seq tracks at the *Thpok* locus in CD24<sup>hi</sup>TCRβ<sup>lo</sup>CD4<sup>+</sup>CD8<sup>+</sup> pre-selection thymocytes of 6-week-old *Bcl11b*<sup>+/+</sup> mice, newborn *Bcl11b*<sup>+/fl</sup>:*Lck-cre*, newborn *Bcl11b*<sup>HM/fl</sup>:*Lck-cre* and newborn *Bcl11b*<sup>HM/HM</sup>:*Lck-cre* mice. ATAC-seq results in immature B cells (GSM2123187) is shown as reference. DP thymocyte- and B lymphocyte-specific peaks are marked with blue and red arrowheads, respectively. Position of the P1 promoter is indicated with black arrowhead

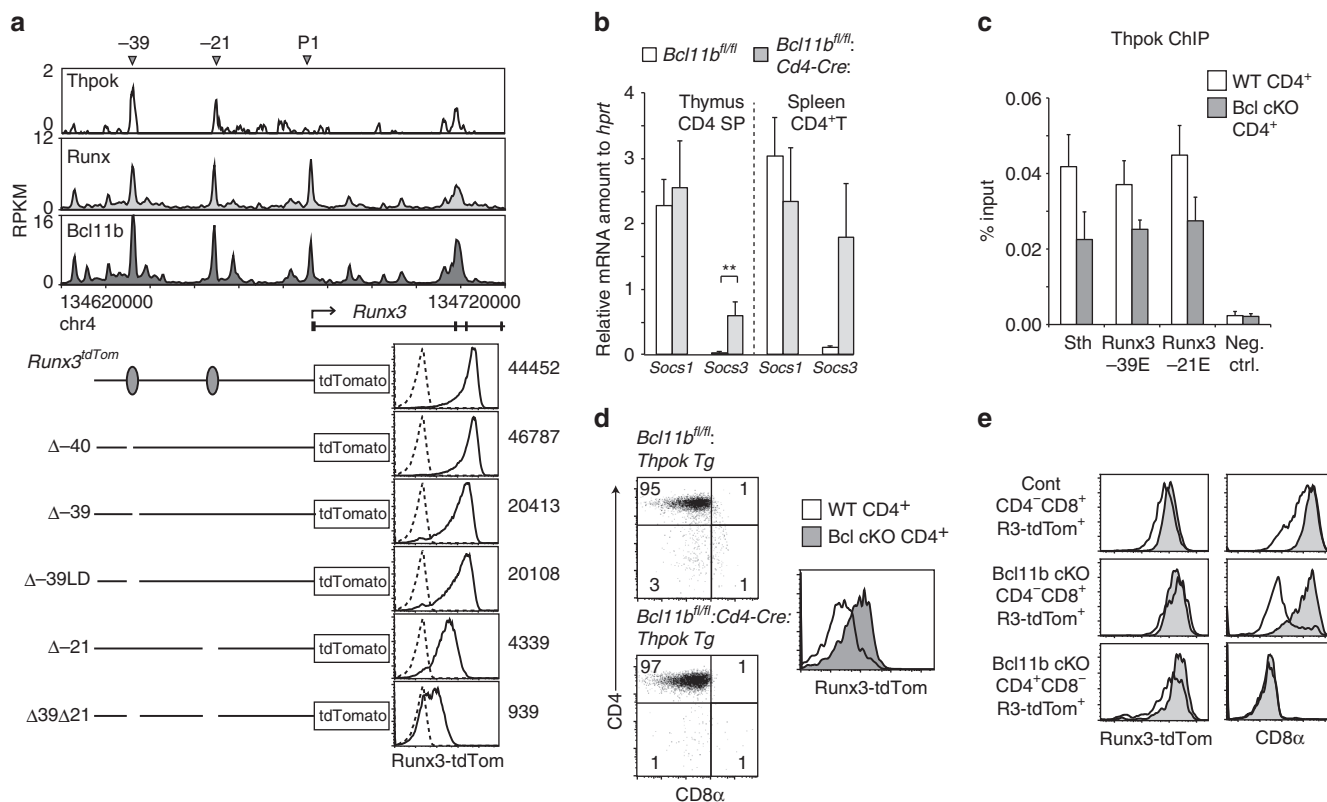




**Fig. 4** Lineage scrambling resulting from inactivation of *Bcl11b* at the DP stage due to sporadic *Thpok* and *Runx3* expression. **a** Flow cytometry analyzing the expression of CD4, CD8, *Thpok*-GFP, and *Runx3*-tdTomato by various thymocyte subsets in mice with the indicated genotypes. The histogram in the middle showing *Thpok*-GFP expression in CD4<sup>+</sup>CD8<sup>-</sup> mature thymocytes from control (open) and *Bcl11b*-deficient (shaded) cells. **b** Graph showing summary of percentage of each subset in mature (CD24<sup>-</sup>TCRβ<sup>hi</sup>) thymocytes population of mice with indicated genotype. **c** Flow cytometry analyzing the expression of CD4, CD8, *Thpok*-GFP and *Runx3*-tdTomato in lymph node T cells of mice with the indicated genotypes. One representative of at least independent mice. **d** Flow cytometry analyzing the expression of CD4, CD8, *Thpok*-GFP and *Runx3*-tdTomato in pre-selection DP thymocytes and lymph node T cells differentiated in Rag1-KO hosts from *Bcl11b*<sup>+/+</sup> and *Bcl11b*<sup>m/m</sup> fetal livers harboring *Thpok*<sup>gfp</sup> and *Runx3*<sup>tdTomato</sup> reporter alleles. One representative of two independent experiments. **e, f** Representative flow cytometry **e** analyzing differentiation of MHC-I and MHC-II selected cells in MHC-II and MHC-I deficient (MHC-II<sup>o</sup> and MHC-I<sup>o</sup>) backgrounds, respectively, along with a graph **f** showing a statistical summary of the percentages and absolute numbers in each cell subset in CD24<sup>-</sup>TCRβ<sup>hi</sup> mature thymocytes population of control (white circles) and *Bcl11b*<sup>fl/fl</sup>.*Cd4-Cre* (filled circles) mice. Partial re-directed differentiation occurred in both MHC-I and -II selected cells. \**P* < 0.05, \*\*\**P* < 0.001 (unpaired *t*-test)

To further address activation status of *Thpok* locus in *Bcl11b* mutant cells, we examined chromatin accessibility in DP thymocytes prepared from a *Bcl11b*<sup>HM/fl</sup>.*Lck-Cre* and *Bcl11b*<sup>HM/HM</sup>.*Lck-Cre* mice by ATAC-seq<sup>37</sup>. Interestingly, the *Thpok* locus in *Bcl11b*<sup>HM/HM</sup> pre-selection thymocytes retained open chromatin signatures around the P1-promoter, resembling those seen in B cells expressing P1-derived transcripts (Fig. 3e). ChIP-seq analyses for *Bcl11b* binding in thymocytes of *Thpok*<sup>gfp</sup>:*ΔTESPE*/*gfp*:*ΔTESPE* mice detected association of *Bcl11b* at intronic regions nearby the P1-promoter (Fig. 1d), which retained accessibility in the *Bcl11b*<sup>HM/HM</sup> cells and whose sequences are evolutionally

conserved (Fig. 1d). Similar open structure was observed in DP cells of *Bcl11b*<sup>HM/fl</sup>.*Lck-Cre* mice. These observations indicate that truncation of *Bcl11b* C-terminal sequences interferes, most specifically, with normal patterns of *Thpok* expression at early thymocyte differentiation, reconfiguring its regulome to mimic expression patterns found in B cells. However, RNA-seq did not detect significant elevation of B cell signature genes in *Bcl11b*<sup>m/m</sup> pre-selection thymocytes (Supplementary Fig. 3a, b), suggesting that *Bcl11b* might be involved in modulating chromatin accessibility at the restricted loci such as *Thpok*.



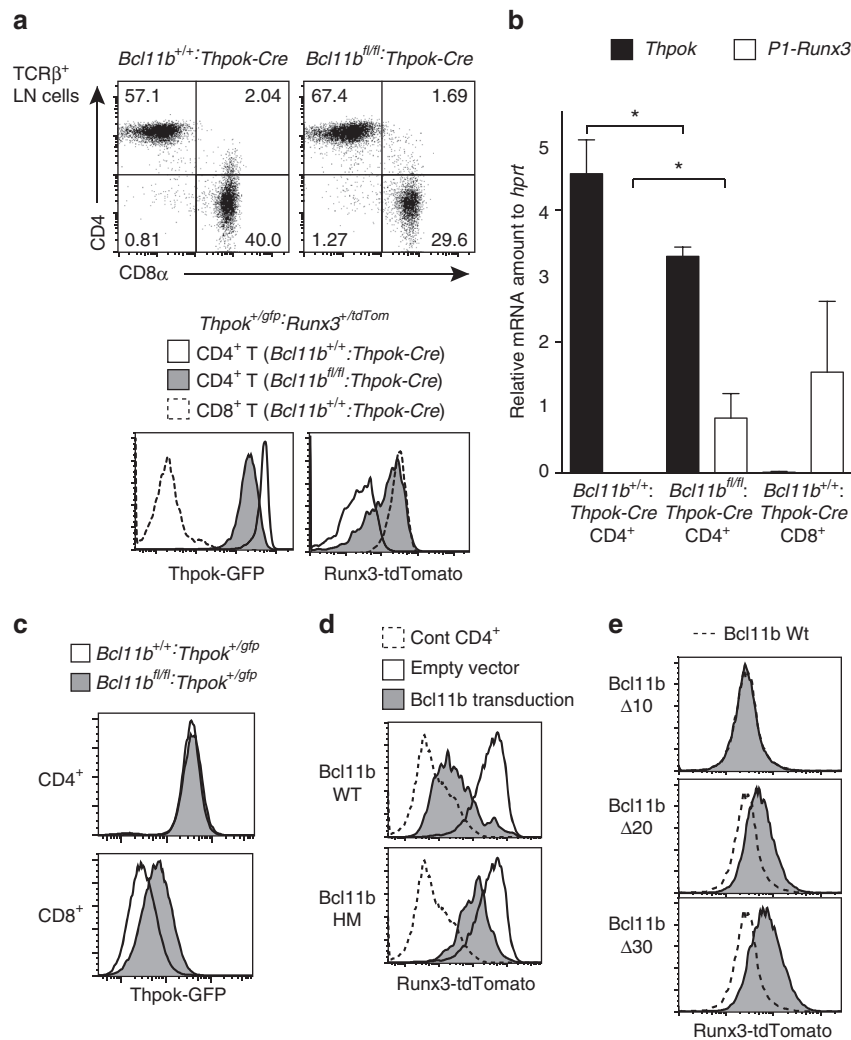
**Fig. 5** Roles of Bcl11b in ThPOK-mediated repression of Runx3. **a** Runx and Bcl11b ChIP-seq tracks at the *Runx3* locus using total thymocytes, along with a ThPOK ChIP-chip track using CD4 SP thymocytes. Schematic structure of mutant *Runx3<sup>tdTomato</sup>* reporter alleles harboring several types of deletion (bottom panel, left). Histograms (bottom panel, right) showing Runx3-tdTomato expression in CD4<sup>+</sup>CD8<sup>+</sup> thymocytes from indicated mutant alleles (solid line) and non-transgenic allele (dotted line). Data shown are representative of at least three independent mice. **b** Amount of *Socs1* and *Socs3* mRNA in CD4 SP thymocytes and CD4<sup>+</sup> T cells from mice of indicated genotypes. Summary of three independent experiments. Mean  $\pm$  SD \*\**P* < 0.01 (unpaired *t*-test). **c** ChIP-qPCR analyzing Thpok binding to the *Sth* in *Thpok* and the -39E and -21E regions in the *Runx3* using chromatin from CD4<sup>+</sup> T cells of Wt (white bars) and *Bcl11b<sup>fl/fl</sup>;Cd4-Cre* (gray bars) mice. Combined data of three independent ChIP experiments. Means  $\pm$  SD. **d** Flow cytometry analyzing the effect of a *Thpok* transgene on Cd8 and Runx3 expression. Dot plots showing CD4/CD8 expression in splenic T cells. Histograms at the right show *Runx3<sup>tdTomato</sup>* reporter expression in Wt (white) and *Bcl11b*-deficient (gray) CD4<sup>+</sup>CD8<sup>-</sup> subset. One representative of three individual mice. **e** Histograms showing CD8 and Runx3-tdTomato expression in control CD4<sup>+</sup>CD8<sup>+</sup>, *Bcl11b*-deficient CD4<sup>+</sup>CD8<sup>+</sup> and CD4<sup>+</sup>CD8<sup>-</sup> T cells after retroviral transduction of Wt (open) and the non-functional HD mutant ThPOK (shaded). One representative of two independent experiments

### Bcl11b is needed for lineage fidelity by MHC-restriction.

Contrary to CD4-skewing from *Bcl11b<sup>m/m</sup>* progenitor, a previous study reported emergence of CD8<sup>+</sup> T cells following *Bcl11b* inactivation at the DP stage using a *Cd4-Cre* driver. We reasoned that the apparent discrepancy could result from a stage-specific requirement for Bcl11b in regulating *Thpok* gene. We therefore next examined effect of *Bcl11b* inactivation by *Cd4-Cre* driver on *Thpok* expression. Although Thpok-GFP was uniformly de-repressed in all *Bcl11b<sup>m/m</sup>* pre-selection thymocytes (Fig. 2c), expression of this reporter was variegated and detected in only half of the pre-selection thymocytes from *Bcl11b<sup>fl/fl</sup>;Cd4-Cre* mice (Fig. 4a). We also observed that CD4<sup>+</sup>CD8<sup>+</sup> and CD4<sup>+</sup>CD8<sup>-</sup> subsets were present in mature thymocytes from *Bcl11b<sup>fl/fl</sup>;Cd4-Cre* mice (Fig. 4a, b). It is noteworthy that, after positive selection, levels of Thpok-GFP in CD4<sup>+</sup>CD8<sup>-</sup> cells differentiated from *Bcl11b<sup>m/m</sup>* progenitors or in *Bcl11b<sup>fl/fl</sup>;Cd4-Cre* mice were lower than those in control counterparts (Fig. 2c and Fig. 4a middle), a shift that was also reflected in attenuated *P2-Thpok* mRNA (Fig. 3c). Thus, loss of Bcl11b function disrupts not only *Thpok* repression, but also distinct mechanisms for its activation. These results suggests that, in a portion of pre-selection thymocytes in the *Bcl11b<sup>fl/fl</sup>;Cd4-Cre* mice, the kinetics of *Thpok* de-repression are delayed compared to *Bcl11b<sup>m/m</sup>* progenitors, which is combined later with low *Thpok* expression levels, allowing the cells to differentiate into CD4<sup>+</sup>CD8<sup>+</sup> and CD4<sup>+</sup>CD8<sup>-</sup> subsets.

Since previous studies showed release of *Runx3* repression in CD4<sup>+</sup> T cells that differentiated under low levels of ThPOK<sup>27, 38</sup>, we next examined *Runx3* expression using a *Runx3-tdTomato* reporter allele that monitors distal P1-promoter activity, which is CD8-lineage specific and serves as a major source for Runx3 protein expression in T cells<sup>27</sup>. Importantly, we found that most mature thymocytes from *Bcl11b<sup>fl/fl</sup>;Cd4-Cre* mice co-expressed Runx3-tdTomato and Thpok-GFP, regardless of their CD4/CD8 expression profiles (Fig. 4a), whereas expression of these genes was mutually exclusive between CD4<sup>+</sup>CD8<sup>-</sup> and CD4<sup>+</sup>CD8<sup>+</sup> cells from control mice. Contrary to a previous report showing *Runx3* de-repression in CD4<sup>+</sup>CD8<sup>+</sup> DP thymocytes of *Bcl11b<sup>fl/fl</sup>;Cd4-Cre* mice<sup>31</sup>, Runx3-tdTomato was not detectably expressed in CD24<sup>hi</sup>TCR $\beta$ <sup>lo</sup> pre-selection thymocytes (Fig. 4a). Co-expression of Thpok-GFP and Runx3-tdTomato was observed also in CD4<sup>+</sup>CD8<sup>-</sup> and CD4<sup>+</sup>CD8<sup>+</sup> cells in the peripheral T cell pool of *Bcl11b<sup>fl/fl</sup>;Cd4-Cre* mice (Fig. 4c), as well as in CD4<sup>+</sup> T cells differentiated from *Bcl11b<sup>m/m</sup>* progenitors (Fig. 4d). Thus, lineage-specific expression of two major factors, ThPOK and Runx3, which drive the CD4/CD8 lineage dichotomy, requires Bcl11b, or more specifically, Bcl11b function mediated by its C-terminal sequences.

We next examined the MHC specificity of T cells differentiated in *Bcl11b<sup>fl/fl</sup>;Cd4-Cre* mice by crossing with MHC-I or MHC-II deficient mice. In mature thymocytes and peripheral T cell



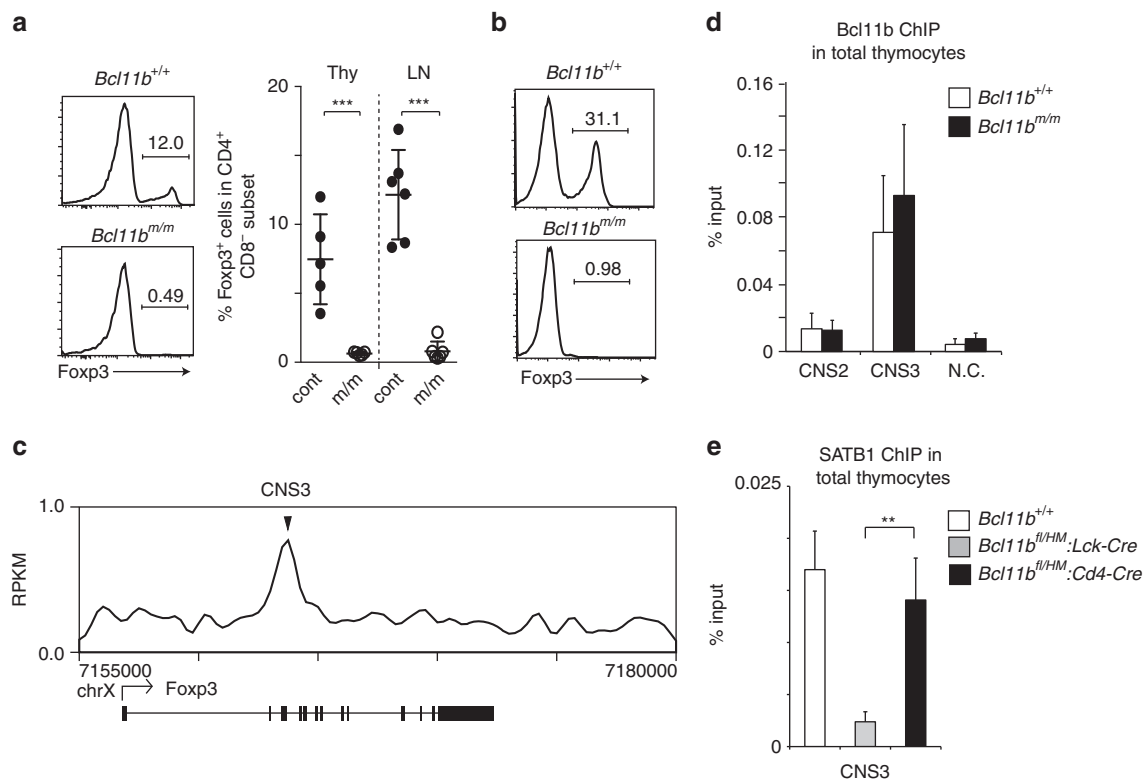
**Fig. 6** Bcl11b function requires its C-terminal zinc-finger motif. **a** Flow cytometry analyzing the expression of CD4, CD8, Thpok-GFP, and Runx3-tdTomato in peripheral T cells from *Bcl11b*<sup>+/+</sup>:*Thpok-Cre* and *Bcl11b*<sup>fl/fl</sup>:*Thpok-Cre* mice harboring *Thpok*<sup>gfp</sup> and *Runx3*<sup>tdTomato</sup> reporter genes. Inactivation of *Bcl11b* during the maturation of CD4-lineage cells by a *Thpok-Cre* transgene resulted in impaired Thpok-GFP expression, as well as Runx3-tdTomato de-repression. **b** RT-qPCR analyses showing relative amounts of *Thpok* (black bars) and *P1-Runx3* (white bars) mRNA in splenic CD4<sup>+</sup>CD8<sup>-</sup> T cells of *Bcl11b*<sup>+/+</sup>:*Thpok-Cre* and *Bcl11b*<sup>fl/fl</sup>:*Thpok-Cre* mice. Results in splenic CD4<sup>+</sup>CD8<sup>+</sup> T cells of *Bcl11b*<sup>+/+</sup>:*Thpok-Cre* are shown as references. Data were summary of three experiments. Means  $\pm$  SD \**P* < 0.05 (unpaired *t*-test). **c** Histograms showing Thpok-GFP expression from CD4<sup>+</sup> and CD8<sup>+</sup> T cells of *Bcl11b*<sup>+/+</sup>:*Thpok*<sup>+/gfp</sup> (white bars) and *Bcl11b*<sup>fl/fl</sup>:*Thpok*<sup>+/gfp</sup> (gray bars) mice 4 days after retroviral Cre transduction. One representative of three experiments. **d** Histograms showing Runx3-tdTomato expression 3 days after retroviral transduction of Wt and hypomorphic (HM) Bcl11b into peripheral CD4<sup>+</sup>CD8<sup>-</sup>Runx3-tdTomato<sup>+</sup> cells from *Bcl11b*<sup>fl/fl</sup>:*Runx3*<sup>+/tdTomato</sup>:*Thpok-Cre* mice. Expression of Runx3-tdTomato in activated control CD4<sup>+</sup> T cells (dotted line) after transduction with an empty vector (open) are shown as controls. **e** Histograms showing Runx3-tdTomato expression after retroviral transduction of mutant Bcl11b proteins lacking the C-terminal zinc-finger motif are shown as in **d**. The dotted line represents Runx3-tdTomato expression after transduction of Wt Bcl11b. One representative of at least two experiments **d**, **e**

populations, a significant proportion of both MHC-I- and MHC-II-specific cells in *Bcl11b*<sup>fl/fl</sup>:*Cd4-Cre* mice expressed surface markers for the alternative lineage, CD8 and CD4, respectively (Fig. 4e, f). We therefore conclude that a substantial number of both MHC-I and MHC-II selected thymocytes failed in their specification for the appropriate developmental pathway, presumably due to spurious expression of ThPOK and Runx3 following positive selection. Ultimately, the chaotic expression of these factors resulted in “lineage scrambling”, unlinking DP precursor thymocyte commitment from MHC specificity.

**Functions of Bcl11b in Runx3 regulation.** To understand how Bcl11b regulates *Runx3* expression, we examined whether genomic regions around the *Runx3* locus were occupied by Bcl11b, Runx,

and ThPOK factors. In addition to the P1-promoter, Bcl11b and Runx bound at two regions (−39 and −21 kb), which were also bound by ThPOK (Fig. 5a) that is essential for *Runx3* repression in CD4<sup>+</sup> T cells<sup>22, 27</sup>. To examine potential of these regions as *cis*-regulatory elements for *Runx3* regulation, we removed the regions separately or in combination from the *Runx3-tdTomato* allele using CRISPR/Cas9 technology<sup>39</sup> (Supplementary Fig. 4). Levels of Runx3-tdTomato expression in CD4<sup>+</sup>CD8<sup>+</sup> SP thymocytes were reduced upon removal of either the −39 or −21 kb region, but not a control region of similar length positioned at −40 kb region (Fig. 5a). Expression of the *Runx3* reporter was further attenuated following combined deletion of the −39 and −21 kb regions. These observations indicated essential regulatory functions for these two regions in driving proper levels of *Runx3* expression in CD8<sup>+</sup> T cells. The residual expression when both regions were





**Fig. 7** Regulation of *Foxp3* expression by Bcl11b. **a** Expression of Foxp3 in CD4<sup>+</sup>CD8<sup>-</sup> mature thymocytes differentiated in Rag1-KO hosts from *Bcl11b*<sup>+/+</sup> and *Bcl11b*<sup>m/m</sup> fetal liver progenitors. A representative histogram (left) along with a summary of five independent hosts (right) are shown. \*\*\**P* < 0.001 (unpaired *t*-test). **b** Histogram showing Foxp3 induction from CD62L<sup>+</sup>CD4<sup>+</sup> T cells after in vitro culture under Treg differentiation condition. One representative of two experiments. **c** Bcl11b ChIP-seq tracks at the *Foxp3* gene in DP thymocytes. Gene structure, transcriptional orientation and position of conserved non-coding sequence (CNS) 3 are indicated. **d** Summary of ChIP-qPCR assay for binding of Wt (white bars) and hypomorphic (black bars) Bcl11b to CNS2 and CNS3 regions in the *Foxp3* locus in neonatal total thymocytes. **e** ChIP-qPCR for binding of SATB1 to the CNS3 enhancer in the *Foxp3* gene. Chromatin was collected from DP thymocyte precursors using mice with the indicated genotypes. Combined data from three independent ChIP experiments are shown. \*\**P* < 0.01 (unpaired *t*-test)

removed also indicated the presence of other enhancer(s) in the *Runx3* locus.

SOCS family proteins were proposed to mediate ThPOK-dependent repression of *Runx3* following thymocyte positive selection<sup>29</sup>. However, expression of *Socs1* and *Socs3* was not reduced in Bcl11b-deficient CD4<sup>+</sup> T cells expressing *Runx3* (Fig. 5b). In Bcl11b-deficient cells, ThPOK bindings to -39 and -21 enhancers were also observed, albeit to a lesser extent, presumably due to lower ThPOK expression levels (Fig. 5c). We then tested whether ectopic ThPOK expression restored *Runx3* repression in CD4<sup>+</sup> T cells. Although transgenic ThPOK prevented the development of CD8<sup>+</sup> T cells in *Bcl11b*<sup>fl/fl</sup>:*Cd4-Cre* mice, *Runx3*-tdTomato was not repressed as much in CD4<sup>+</sup>CD8<sup>-</sup> splenic T cells (Fig. 5d). Consistent with this observation, retroviral ThPOK transduction into peripheral CD8<sup>+</sup> T cells reduced *Runx3*-tdTomato expression in control cells, but not in Bcl11b-deficient cells, while CD8 expression was reduced in both cells, as reported previously<sup>40</sup> (Fig. 5e). These results indicated that full *Runx3* repression by ThPOK required Bcl11b after ThPOK bound to *Runx3* enhancers. In contrast, ThPOK could repress *Cd8* expression in the absence of Bcl11b.

**Bcl11b regulates *Thpok* and *Runx3* expression.** To address at which developmental stages Bcl11b is necessary for appropriate *Thpok* and *Runx3* expression, we inactivated the *Bcl11b* gene during maturation of MHC-II selected cells using a *Thpok-Cre*

driver<sup>41</sup>. While CD4 and CD8 expression on peripheral T cells remained unchanged in *Bcl11b*<sup>fl/fl</sup>:*Thpok-Cre* mice, the expression of *Thpok*-GFP and *Runx3*-tdTomato in CD4<sup>+</sup> T cells was decreased or increased, respectively (Fig. 6a, b). When Bcl11b was inactivated in differentiated CD8<sup>+</sup> T cells by retroviral Cre transduction, de-repression of *Thpok*-GFP was also observed (Fig. 6c). These findings indicated that Bcl11b was necessary to maintain lineage-specific expression of both *Thpok* and *Runx3* after primary commitment of post-selection thymocytes to CD4/CD8 lineages.

We also tested the converse scenario, namely, whether retroviral Bcl11b transduction restored *Runx3*-tdTomato repression in differentiated CD4<sup>+</sup> cells of *Bcl11b*<sup>fl/fl</sup>:*Thpok-Cre* mice. Levels of *Runx3*-tdTomato was reduced upon Bcl11b transduction compared with empty vector controls (Fig. 6d); however, transduction with a vector encoding the Bcl11b<sup>HM</sup> protein repressed *Runx3*-tdTomato to lesser extent than wild-type Bcl11b (Fig. 6d). To gain more insights into the domain structure of Bcl11b, we tested the activities of other mutants generated by sequential deletion on its C-terminus (Supplementary Fig. 5). While a Bcl11b mutant retaining the final zinc-finger domain (Bcl11b-Δ10) and wild-type Bcl11b equally repressed *Runx3*-tdTomato expression, other mutants lacking the zinc-finger structure showed weaker activity (Fig. 6e). These results not only confirmed that loss of the last zinc-finger motif, rather than acquisition of aberrant sequences, attenuated Bcl11b function in the Bcl11b<sup>HM</sup> protein, but the new data also indicated that the

last zinc-finger motif is essential for Bcl11b function in preserving lineage-specific *Runx3* expression.

In this regard, we compared structures of Bcl11 family proteins in other species. There are two zinc-finger clusters in Bcl11b, double zinc-fingers in its central region and triple zinc-fingers at its C-terminus. Comparative genomics analyses revealed that some species, such as the Ascidians (*C. intestinalis*) and Worm (*C. elegans*), lacked the middle zinc-finger motif in Bcl11 orthologues, whereas the triplet zinc-finger structure at the C-terminal was highly conserved during evolution (Supplementary Fig. 6), suggesting that C-terminal triplet zinc-fingers could be involved in ancestral and conserved function of Bcl11 family proteins.

### Stage-specific requirement for Bcl11b in *Foxp3* transcription.

Additional analyses of T cell development from *Bcl11b<sup>m/m</sup>* progenitors detected a striking defect in the development of *Foxp3<sup>+</sup>* regulatory T cells (Treg) (Fig. 7a). In line with this finding, we were unable to induce *Foxp3* expression in the mutant CD4<sup>+</sup> T cells even under potent Treg differentiation conditions in vitro (Fig. 7b). Contrary to this finding, a previous study reported the emergence of *Foxp3<sup>+</sup>* Treg cells following *Bcl11b* inactivation at the DP stage using a *Cd4-Cre* driver<sup>8</sup>. This apparent discrepancy could result from a direct and stage-specific requirement for Bcl11b in activating *Foxp3*, as we observed in *Thpok* regulation by Bcl11b. Recent work reported that the CNS3 region in the *Foxp3* gene is essential to initiate *Foxp3* activation by poising the *Foxp3* promoter in an active state prior to the DP stage<sup>42, 43</sup>. Our ChIP-seq detected Bcl11b binding to CNS3 in total thymocytes (Fig. 7c) and our ChIP-qPCR detected similar binding of both Bcl11b and Bcl11b<sup>HM</sup> proteins to CNS3 (Fig. 7d). Since we recently observed that binding of the genome organizer, SATB1, to CNS3 is essential for *Foxp3<sup>+</sup>* Treg development in the thymus<sup>41, 44</sup>, we examined SATB1 association with CNS3 in *Bcl11b* mutant thymocytes. Consistent with development of *Foxp3<sup>+</sup>* cells, SATB1 binding to CNS3 was unaffected in total thymocytes from *Bcl11b<sup>fl/HM</sup>·Cd4-Cre* mice. On the contrary, SATB1 recruitment was significantly decreased in those cells from *Bcl11b<sup>fl/HM</sup>·Lck-Cre* mice (Fig. 7e). These results suggested that priming of the *Foxp3* locus by Bcl11b prior to the DP stage is essential for activation of *Foxp3* gene via promoting SATB1 recruitment to the CNS3 enhancer.

### Discussion

In this study, we demonstrate that a hypomorphic Bcl11b protein (Bcl11b<sup>HM</sup>), lacking only its final zinc-finger domain, supports T-lineage commitment and ILC2 development. In contrast, the hypomorphic *Bcl11b* mutation fails to control the expression of transcription factors, including ThPOK, Runx3, and *Foxp3*, which are important for specification of mature T lineages (CD4-helper, CD8-cytotoxic, and Treg, respectively). Importantly, our findings dissect the functional architecture of Bcl11b, revealing independent regions that are responsible for early T-lineage commitment versus transcriptional regulation of lineage-specifying genes during positive selection.

The distinct effects on expression of *Thpok* and *Foxp3* genes of the germline hypomorphic mutation when compared with a DP stage-specific inactivation of its gene indicate that Bcl11b acts on these two genes at an early developmental stage, presumably at the DN to DP transition, to control appropriate expression of these lineage-specifying transcription factors at later stage. For *Thpok*, while a half of precursors still repressed *Thpok* gene after conditional loss of Bcl11b expression, hypomorphic Bcl11b mutation caused *Thpok* expression in all precursors that retains non-T cells like open chromatin structure around the

P1-promoter. This observation indicates that early *Thpok* repression at the DN to DP transition by Bcl11b involves chromatin closing. Our results using the *Thpok<sup>gfp:ΔTESPE</sup>* mutant reporter allele lacking *Sth* silencer revealed that Bcl11b regulates such chromatin closing independently of the *Sth*, which is essential to silence *Thpok* expression at later stage specifically in cytotoxic T cells<sup>35</sup>. Given that the *Sth* is sufficient to repress reporter transgenes driven by heterologous enhancer/promoter in DP thymocytes<sup>20</sup> and is essential to prevent expression of the endogenous gene in Bcl11b-sufficient DP cells<sup>19</sup>, both *Sth*-independent and *Sth*-dependent pathways operate to repress *Thpok* in precursor DP thymocytes. Considering that ectopic ThPOK expression in DP precursor thymocytes not only disturbs helper/cytotoxic lineage choice<sup>16</sup> but also increases a risk for lymphomagenesis<sup>45</sup>, it might be beneficial to have two independent mechanisms to secure no leaky ThPOK expression in DP precursors. It is noteworthy that *Sth*-independent *Thpok* repression was reversible, since *Thpok*-GFP expression from the *Thpok<sup>gfp:ΔTESPE</sup>* allele was induced in mature T cells, suggesting that irreversible epigenetic changes such as DNA methylation are unlikely to be involved in early *Sth*-independent *Thpok* repression by Bcl11b. Given that *Sth* could stably silence *Thpok* transcript, particularly from the P2-promoter, in CD8-lineage cells<sup>35</sup>, mechanisms of *Sth*-dependent and *Sth*-independent *Thpok* repression could be different.

The precise involvement of Bcl11b in *Sth*-independent *Thpok* repression remains elusive. As detected by ChIP-seq, Bcl11b covers regions around the P1-promoter in an *Sth*-independent manner. Along with evolutionary conservation in sequences of intronic region just downstream of the P1-promoter, there might exist uncharacterized regulatory element(s) that suppresses P1-promoter activity upon expression of Bcl11b. Binding of mutant Bcl11b<sup>HM</sup> protein to such intronic region suggest that the C-terminal zinc-finger motif in Bcl11b is necessary to close P1-promoter region after Bcl11b binding to putative regulatory element(s). Alternatively, Bcl11b may be involved indirectly in the *Sth*-independent *Thpok* repression through an uncharacterized intermediate factor, however, the similar gene expression profiles between wild-type and *Bcl11b<sup>m/m</sup>* precursors disfavors this possibility.

De-repression of *Thpok* in mature CD8<sup>+</sup> populations that emerge from *Bcl11b<sup>fl/fl</sup>·Cd4-Cre* precursors suggests that Bcl11b is required also for *Sth*-mediated *Thpok* repression. Thus, Bcl11b regulates *Thpok* repression by both *Sth*-independent and *Sth*-dependent mechanisms. Based on our data, we presume that Bcl11b acts initially to inhibit P1-promoter activity in an *Sth*-independent manner. Either sequentially or simultaneously, Bcl11b primes three T cell specific regulatory regions, *Sth* and two enhancers, *TE* and *PE*. Upon such priming, *Sth* initiates to suppress *TE* and *PE* activity, thereby removal of *Sth* results in *Thpok* de-repression in Bcl11b-sufficient precursors in which *TE/PE* enhancers also becomes functional after Bcl11b-mediated priming. In this regard, an *Sth*-independent mechanism alone is likely to be insufficient for extinguishing *TE/PE* enhancer activities. Thus, we assume that pre-conditioning of T cell specific regulatory regions by Bcl11b is an essential process for dissecting TCR signals that arise from engagement by distinct MHCs, and for converting them into establishment of “on” and “off” *Thpok* expression status for inducing helper- or cytotoxic-fate via inactivation or maintaining *Sth* activity, respectively.

Bcl11b similarly primes the *Foxp3* locus for later activation upon exposure to agonistic selection signals<sup>46</sup>. A recent study showed that the pioneering enhancer, CNS3<sup>43</sup>, confers a poised state to the *Foxp3* promoter in precursor cells in response to TCR stimuli<sup>42</sup>. In the present study, we show that binding of SATB1 to CNS3 in precursors, an essential process for activation of *Foxp3*

gene in the thymus<sup>41, 44</sup>, requires Bcl11b-mediated priming before transition to the DP stage. Whether common Bcl11b-mediated mechanisms operate to prime the *Thpok* and *Foxp3* loci remains unclear; however, we establish that the C-terminal zinc-finger motif of Bcl11b is involved in both processes. Collectively, our results reveal a set of novel functions for the T-lineage commitment factor Bcl11b, which “pre-conditions” precursors for integration of environmental cues, specifically TCR signals, into a developmental program that dissects effector lineages and shapes the primary T cell pool.

Another striking finding from our studies is that the lack of Bcl11b inhibits both the positive and negative regulatory regions near *Thpok*. One possible model to explain how Bcl11b controls multiple regulatory regions would be its function as a scaffold to facilitate communication between the distal elements and the *Thpok* promoters. Using insertion ChIP technology<sup>47</sup>, we recently observed that the *Sth* and *PE* regions were positioned in close proximity during differentiation into CD8-lineage cells (I.T. manuscript in preparation). Thus, it is possible that Bcl11b is involved in shaping promoter-regulatory elements communications on the *Thpok* locus.

Contrary to *Thpok* regulation, Bcl11b is merely involved in repression of *Runx3* in CD4<sup>+</sup> T cells<sup>22, 27</sup>. Prior studies invoked an indirect mechanism for *Runx3* repression by ThPOK via the induction of SOCS family proteins<sup>29</sup>. However, our finding that ThPOK associates with functional *Runx3* enhancers raises the possibility that ThPOK directly antagonizes enhancer function. Indeed, ThPOK bound to the enhancers even in Bcl11b-deficient cells that were de-repressing *Runx3* albeit normal level of *Socs* family gene expression. These data suggest that the antagonistic action of ThPOK against *Runx3* expression requires Bcl11b after its binding to the *Runx3* enhancers. It is conceivable that Bcl11b assists ThPOK in preventing the formation of looping between the enhancers and the P1-*Runx3* promoter. Nevertheless, it will be important in future studies to examine whether lineage-specific chromatin structures are formed in *Thpok* and *Runx3* loci and whether Bcl11b regulates these topologies.

The C<sub>2</sub>H<sub>2</sub> zinc-finger motif not only serves as a DNA binding domain that recognizes specific sequences, but also as a docking module for RNA and proteins<sup>48</sup>. We found that the Bcl11b<sup>HM</sup> protein retains its ability to associate with regulatory regions, indicating that its final zinc finger is dispensable for DNA binding. However, it remains possible that the other four zinc-fingers domains have compensatory functions in this regard. There are two zinc-fingers clusters in mammalian Bcl11 family proteins, double zinc-fingers in its central portion and a trio of zinc-fingers at its C-terminus. Although there are Bcl11-related proteins in lower species that lacks the central dual zinc-finger domain, a trio of zinc-fingers at the C-terminus is evolutionally well conserved. Our results indicate that this domain is essential for gene regulation, presumably through modulating topological structures. Bcl11a has been shown to play a key role in the switch from fetal  $\gamma$ - to adult  $\beta$ -globin in human<sup>49</sup>. It is conceivable that common Bcl11 family-dependent mechanisms play a key role in developmentally programmed switching of regulatory regions by modulating topological structures through the C-terminal triplet zinc-finger stretch. Isolation of interacting molecules with it will further elucidate regulatory mechanism by Bcl11 family.

## Methods

**Mice.** *Runx1*<sup>Δ446</sup> mice<sup>50</sup>, *Runx3*<sup>fl</sup> mice<sup>51</sup>, *Bcl11b*<sup>+/-</sup> mice<sup>36</sup>, *Thpok*<sup>gfp</sup> mice<sup>22</sup>, *Thpok*<sup>gfp:241-401RM</sup> mice<sup>35</sup> and *Thpok* transgenic mice<sup>19</sup> have been described. *Lck-Cre* mice, and *Cd4-Cre* mice were from Dr J. Takeda, and Dr C. Wilson, respectively.  $\beta$ 2m-deficient mice (stock No:002070), Rag1-deficient mice (Stock

No:002216) were from Jackson laboratory, and I-A $\beta$  deficient mice and *Il2rg*<sup>-/-</sup>: *Rag1*<sup>-/-</sup> were from Taconic. In order to generate a *Thpok*<sup>gfp:ΔTESPE</sup> allele, *TE/Sth* and *PE* regions were sequentially removed from the *Thpok*<sup>gfp</sup> reporter allele by using homologous recombination in ES cells. In order to construct the target vector for *Bcl11b*<sup>fl</sup> allele, genomic fragment isolated from the phage library (Stratagene) was used as starting material. The neomycin-resistance gene (*neo*<sup>r</sup>) flanked with two loxP sequences was cut out from *pL2-Neo(2)* vector and thymidine kinase (TK) gene was isolated from pNT vector. A 5' short homology region was amplified by PCR and was ligated with the *neo*<sup>r</sup> fragment, followed by sequential ligation to add 3' homology region. The third loxP sequences placed in the 3' untranslated region (UTR) were derived from synthetic oligo-nucleotide. The single nucleotide deletion that causes a frameshift mutation in the *Bcl11b*<sup>fl</sup> allele was accidentally incorporated into the target vector at some point during above sequential ligation steps. To construct a real target vector for *Bcl11b*<sup>fl</sup> mutation, correct exon 4 sequences were replaced in the *Bcl11b*<sup>fl</sup> target vector. A fragment that harbors the identical *Bcl11b*<sup>fl</sup> single nucleotide deletion but lacks the third loxP sequences was used to construct the target vector for *Bcl11b*<sup>HM</sup> allele. In order to generate *Runx3*<sup>tdTomato</sup> reporter allele, we constructed the targeting vector by replacing YFP cDNA fragment with cDNA fragment encoding tdTomato fluorescent protein in pR31-KIN cassette vector<sup>27</sup>, which was used to generate the *Runx3*<sup>YFP</sup> allele<sup>27</sup>. Those targeting vectors were transfected into M1 ES cells as previously described<sup>22</sup>. ES clones underwent homologous recombination were identified by PCR with appropriate primers sets. In order to delete -39 kb and -21 kb genomic regions from the *Runx3*<sup>tdTomato</sup> reporter allele by CRISPR/Cas9 technology<sup>52, 53</sup>, we selected two single guide RNA (sgRNA) target sequences that flanks target genomic regions, which were shown in Supplementary Fig. 3. Custom sgRNA, in which CRISPR RNAs was fused to a normally trans-encoded tracrRNA, were transcribed from a T7 promoter in the *pUC18* vector by in vitro transcription with MEGAscript T7 kit (Life Technologies, AM1354) and both the Cas9 mRNA and the sgRNAs were purified using MEGAclear kit (Life Technologies, AM1908). The sgRNAs and mRNA encoding Cas9 were co-injected into cytoplasm of fertilized eggs that were prepared by in vitro fertilization with *Runx3*<sup>tdTomato/tdTomato</sup> sperm and wild-type oocytes. Cas9-mediated double-stranded DNA breaks resolved by non-homologous end joining (NHEJ) ablated the intervening sequences. Founder offspring that had deletion of the target genome region were crossed to wild-type mice, and F1 founders that harbor both deletion mutation and *Runx3*<sup>tdTomato</sup> allele were selected for establishing mouse line and were analyzed. In order to generate allele harboring double deletions at -39 kb and -21 kb regions, sgRNA pair used for generation of *Runx3*<sup>Δ21E</sup> mutation were injected into eggs obtained by in vitro fertilization between *Runx3*<sup>tdTomato:Δ39LD/+</sup> sperm and wild-type oocytes. RNA injection was performed by the Animal Facility Group at RIKEN, IMS. All mice were maintained in the specific pathogen free animal facility at the RIKEN IMS, and all animal procedures were in accordance with institutional guidelines for animal care and with the protocol (28-017) approved by the safety section in RIKEN Yokohama Campus.

**DNA pull-down assay.** In order to analyze binding of Bcl11b to *Thpok* silencer (*Sth*) core sequences in vitro, nuclear extract was prepared from  $10 \times 10^7$  total thymocytes and were mixed with 10  $\mu$ g of biotin-labeled synthetic oligo-nucleotides probe (Eurofins Genomics) in 400  $\mu$ l of affinity purification (AP) buffer (20 mM HEPES-KOH (pH 7.5), 80 mM KCl, 10% Glycerol, 0.1% Triton X, 0.5 mM PMSF, 1 $\times$  cOmplete protease inhibitor Cocktail tablets (Roche)) on ice for 3 h, followed by mixture with 100  $\mu$ l of Dynabeads M-280 Streptavidin (Thermo Fisher Scientific) for additional 3 h on ice. After one time wash with AP buffer, beads were incubated with 500 pmol of mutant non-labeled oligo-nucleotide for 30 min at 4  $^{\circ}$ C, and were washed with AP buffer three times. Captured protein complexes by beads were then released into 50  $\mu$ l of SDS sample buffer and were applied for western blot analyses. Sequences for oligo-nucleotides probes used were as follows:

Wt:5'-TGGCAGCCACCGCCTCTTCAGGTGGGTGGGGCGGTCTGCGGTAGGGGTTCTGGGGGCGGGAGGGAGGGGGCTGCGGTCTGAG-3'

Mutant:5'-TGGCAGCCACCGCCTCTTCAGCACCCATGGGCGCAGCGGTAGGGGTTCTGGGGGCGGGAGGGAGGGGGCACGCCACTGAG-3'.

**Co-immune precipitation and western blot.** Nuclear extract was prepared from  $10 \times 10^6$  total thymocytes with NE-PER kit (Thermo Fisher Scientific) in the presence of cOmplete protease inhibitor cocktail tablets (Roche), followed by incubation with 2  $\mu$ g of anti-Bcl11b (A300-383A, Bethyl Laboratories) antibody, which was pre-mixed with 50  $\mu$ l of Dynabeads ProteinA, overnight at 4  $^{\circ}$ C with gentle rotation. Antibody bound proteins were released into 20  $\mu$ l of SDS sample buffer and was loaded onto SDS-PAGE gel. In western blot, anti-Runx1 serum<sup>54</sup> and antibodies that reacts with N-terminal sequences (A300-383A, Bethyl Laboratories) and C-terminal sequences (A300-385A, Bethyl Laboratories) of Bcl11b were used to detect protein.

**Chromatin immunoprecipitation.** For ChIP-seq,  $3 \times 10^7$  thymocytes that were freshly prepared from C57BL/6 mice at 4–6 weeks old were washed once with PBS supplemented with 1% FCS, cOmplete protease inhibitor cocktail (Roche) and 1



mM 4-(2-Aminoethyl) benzenesulfonyl fluoride hydrochloride (Sigma-Aldrich), and were cross-linked by incubation in a 1% paraformaldehyde solution for 10 min with gentle rotation at RT. The reaction was stopped by addition of glycine to 0.15 M. Cells were then washed with ice-cold PBS containing 1% FCS for 10 min with gentle rotation at 4 °C, and were lysed in Lysis Buffer 1 (50 mM HEPES pH 7.5, 140 mM NaCl, 1 mM EDTA, 10% Glycerol, 0.5% NP-40, 0.25% Triton X-100) supplemented with cOmplete protease inhibitor cocktail tablets (Roche) for 10 min at 4 °C with gentle rotation. Nuclei were pelleted and were washed by Lysis Buffer 2 (10 mM Tris-HCl pH 8.0, 200 mM NaCl, 1 mM EDTA and 0.5 mM EGTA) supplemented with cOmplete protease inhibitor (Roche). Pelleted chromatin was resuspended in 300 µl of Lysis Buffer 3 (10 mM Tris-HCl pH 8.0, 100 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 0.1% Sodium deoxycholate and 0.5% N-laurylsarcosine sodium salt), and was sonicated using a model XL2000 ultrasonic cell disruptor (MICROSON) at output level 6 for 15 s for 10 times. After removing debris by centrifugation, 30 µl of 10% Triton X-100 (Nacalai tesque) was added to 270 µl of supernatant (final 1%) and sonicated chromatin was incubated overnight at 4 °C with 5 µg of anti-Bcl11b rabbit polyclonal antibody (A300-383A, Bethyl Laboratories) that was pre-conjugated with 50 µl of Dynabeads M-280 Sheep anti-Rabbit IgG (Thermo Fisher Scientific). After washing beads with ChIP-RIPA (50 mM HEPES (pH 7.6), 500 mM LiCl, 1 mM EDTA, 1% NP-40, 0.7% sodium deoxycholate) and TE buffer supplemented with 50 mM NaCl, immunoprecipitates were eluted from beads into 100 µl of elution buffer (50 mM Tris-HCl pH 8.0, 10 mM EDTA, 1% SDS) by incubation for 15 min at 65 °C with vigorous shaking. Eluted immunoprecipitates were then incubated at 65 °C overnight for reverse-crosslinking. Input DNA and ChIP DNA were treated with RNase A (Thermo Fisher Scientific) at 37 °C for 1 h, followed by incubation with Proteinase K (Thermo Fisher Scientific) at 55 °C in the presence of 6 mM CaCl<sub>2</sub> for one hour. DNA was purified by Phenol/Chloroform extraction for ChIP-seq or ChIP DNA Clean and Concentrator kit (ZYMO RESEARCH) for ChIP-qPCR. For ChIP-seq analysis, purified DNA was subjected to re-sonication with a Covaris S220 to produce DNA fragments with an average size of 200 bp, and was used for library construction with NEBNext ChIP-seq Library Prep Master Mix set for Illumina Kit (NEB). Sequencing was performed by the RIKEN IMS sequence facility with Illumina HiSeq 1500 or at Hiroshima University using a GAIIx (Illumina). To detect ThPOK binding regions in the murine *Runx3* locus, we performed ChIP-on-chip experiment. Chromatin DNA isolated from 10 × 10<sup>6</sup> CD4<sup>+</sup> SP thymocytes from FH-ThPOK mice<sup>22</sup> was immunoprecipitated with anti-Flag (M2; Sigma-Aldrich), and were hybridized against custom microarrays generated by Agilent that tiled through the murine *Runx3* locus up to ~565 kb upstream and ~66 kb downstream by means of 60-nucleotide probes. Probe hybridization and scanning of oligo-nucleotide array data were done according to manufacturer's protocol (Agilent). For analytical ChIP, we used 5 µg of rabbit anti-Bcl11b antibody (A300-383A, Bethyl Laboratories), 5 µg of rabbit anti-ThPOK polyclonal antibody<sup>27</sup> and 5 µg of rabbit anti-SATB1 polyclonal antibody (ab70004, Abcam). Quantitative PCR was performed using the StepOnePlus Real-Time PCR system (Applied Biosystems) with SYBR Green detection system. Primers sequences for quantitative PCR are listed in the Supplementary Table 1.

**Flow cytometry analyses.** Thymus, spleen, and lymph nodes were removed from mice at 4–8 weeks of age, and were mashed through a 70 µm cell strainer to make single-cell suspensions. Cells were stained with following antibodies purchased from BD-Bioscience: CD4 (RM4-5), CD8 (53-6.7), CD24 (M1/69), CD25 (PC61), CD154 (MR1), IL1RL1 (DJ8), Sca1 (E13-161.7), KLRG1 (2F1), and TCRβ (H57-597). Foxp3 staining buffer Set (00-5523-00) and anti-mouse Foxp3 antibody (FJK-16s) from affymetrix eBioscience was used to stain Foxp3. Antibodies were used at a concentration of 2.5 µg ml<sup>-1</sup>. Multi-color flow cytometry analysis was performed using a FACSCANTO II (BD-Bioscience) and data were analyzed using FlowJo (Tree Star) software. Cell subsets were sorted using a FACSAria II (BD Biosciences) by cell sorting facility at RIKEN IMS.

**In vitro T cell culture.** Purified T cells were cultured in custom ordered Dulbecco's Modified Eagle Medium (D-MEM, KOHJIN BIO) supplemented with 10% heat inactivated FBS (Hyclone). 1.0 × 10<sup>6</sup> cells were stimulated in 24-well plates pre-coated with 2 µg ml<sup>-1</sup> anti-CD3ε antibody (553058, BD Bioscience) with 2 µg ml<sup>-1</sup> soluble anti-CD28 antibody (553295, BD Bioscience) for two days. For induction of Treg differentiation, 10 ng ml<sup>-1</sup> TGFβ1 (7666-MB, R&D systems), 5 µg ml<sup>-1</sup> anti-IL4 (554433, BD Biosciences) and 5 µg ml<sup>-1</sup> anti-IFNγ (554409, BD Biosciences) were added throughout the culture.

**Reconstitution of T cell development in host mice.** Total liver cells with either *Bcl11b*<sup>+/+</sup> or *Bcl11b*<sup>m/m</sup> genotype were prepared from 13.5 to 14.5 dpc embryos, and were suspended in 400 µl of D-MEM supplemented with 10% heat inactivated FBS (Hyclone). 200 µl of these fetal liver cells were intravenously injected into sub-lethally irradiated (6.5 Gy, Gammacell 40 Exactor, MDS Nordion) Rag1-deficient or I- $\beta$  deficient mice that expressing Ly9.1. For ILC2 reconstitution, *Il2rg*<sup>-/-</sup>:*Rag1*<sup>-/-</sup> mice were used as recipients. Host mice were supplied with antibiotics (1 mg ml<sup>-1</sup> neomycin and 100 unit ml<sup>-1</sup> polymyxin B) containing water for first 2 weeks, and were analyzed 8 weeks after the fetal liver cell injection. In I-

$\beta$  deficient host mice, donor derived lymphocytes were identified as Ly9.1-negative cells.

**RNA isolation, RT-qPCR and RNA-seq.** After total cellular RNA extraction by using Trizol reagent (Thermo Fisher Scientific), samples were incubated with RNase-free DNase I (Thermo Fisher Scientific) before cDNA synthesis with the SuperScriptIII First Strand Synthesis System (Invitrogen). Quantitative RT-PCR was performed using the StepOnePlus Real-Time PCR system (Applied Biosystems) with an internal fluorescent TaqMan probe, Universal ProbeLibrary (Roche) or SYBR Green detection system. Primers sequences and probe sequences for quantitative RT-PCR are listed in the Supplementary Table 2. Primer sequence to measure total *Thpok* mRNA was previously described<sup>16</sup>. For RNA-seq analyses, purified 1.0 × 10<sup>6</sup> cells were used to prepare mRNA and 200–300 ng total RNA were used for the library construction with SureSelect Strand Specific RNA-Seq LibraryPreparation kit (Agilent Technologies) according to the manufacturer's protocol. Sequencing was performed by genomics facility at RIKEN IMS with Illumina HiSeq 1500.

**Assay of transposase-accessible chromatin sequencing.** Assay of transposase-accessible chromatin sequencing (ATAC-seq) samples were prepared from 25,000 sorted cells. The transposase reactions were carried as previously described<sup>55</sup> with 12 or 13 total PCR cycles. Amplified DNA fragments were purified with QIAGEN MinElute PCR Purification Kit and size-selected twice with Agencourt AMPure XP. Libraries were quantified with KAPA Library Quantification Kit for Illumina Sequencing Platforms (KAPA Biosystems), and size distribution was checked on a Bioanalyzer (Agilent High Sensitivity DNA chip, Agilent Technologies). Libraries were sequenced on Illumina HiSeq 2500.

**Retroviral transduction into T cells and T cell culture.** Retroviral vectors encoding Cre recombinase or ThPOK in pMSCV-GFP vector was previously described<sup>56</sup>. Retroviral vectors encoding mutant Bcl11b proteins were generated by insertion of corresponding cDNA into pMSCV-GFP vector after conformation of sequences of PCR amplified fragment. Those vectors were transfected into Plat-E packaging line (a gift from Dr T. Kitamura at the University of Tokyo) by FUGENE6 Reagent (Promega, E2691) and supernatant collected two days after transfection were used for transduction into activated T cells by in vitro TCR stimulation with immobilized anti-CD3 and soluble anti-CD28 by spin infection at 2400 r.p.m. for 60 min in the presence of 4 µg of polybrene (Sigma-Aldrich). Transduced T cells were identified according to GFP expression for FACS analyses and RNA extraction.

**Data and statistical analyses.** Sequences retrieved in ChIP-seq experiments and ATAC-seq were aligned on the mouse genome (mm9) using the bowtie2 (<http://bowtie-bio.sourceforge.net/index.shtml>) with default parameters, and accumulated reads were normalized using total mapped reads. RNA-seq reads were mapped using tophat2 (<https://ccb.jhu.edu/software/tophat/index.shtml>) to the mouse genome and gene expression was estimated with fragments per mapped reads per kilobase (FPKM) values calculated using cufflinks (<http://cole-trapnell-lab.github.io/cufflinks/>) and gene annotation provided by iGenome. Primary component analysis (PCA) was carried out using our in-house program with FPKM values of gene sets whose annotation was provided by Gene Ontology database (<http://geneontology.org/>). Statistical analysis was performed by F-test and unpaired *t*-test with or without Welch's correction using GraphPad Prism6 (Graphpad software). Figures display means and SD.

**Comparative genomics analyses.** Homologous sequences of Bcl11 family were searched with human Bcl11b protein sequences as a query against sequences listed at the Ensemble genome browser 85 and Genebank with "TBLASTN" algorithm. As for the Japanese lamprey search, Bcl11 protein sequences for human, mouse, and elephant shark were used as queries to search the Japanese lamprey genome assembly (<http://jlampreygenome.imcb.a-star.edu.sg/>) with "TBLASTN" algorithm. Genomic regions that showed high similarity were extracted and searched against the non-redundant protein database at NCBI to confirm their identity. Alignments of C-terminal sequences of Bcl11b orthologue of each species were performed by a computer program Clustal Omega (<http://www.clustal.org/omega/>) with default options for amino acid sequences using zinc-finger regions obtained proposed by BLASTP algorithm and NCBI conserved domain database (<https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>).

**Data availability.** RNA-seq, ChIP-seq and ATAC-seq data that support the findings of this study have been deposited with accession GSE90134, GSE90949, and GSE90989, respectively. All other relevant data are available from the authors.

Received: 24 January 2017 Accepted: 27 July 2017  
Published online: 26 September 2017

## References

1. Yui, M. A. & Rothenberg, E. V. Developmental gene networks: a triathlon on the course to T cell identity. *Nat. Rev. Immunol.* **14**, 529–545 (2014).
2. Yang, Q., Jeremiah Bell, J. & Bhandoola, A. T-cell lineage determination. *Immunol. Rev.* **238**, 12–22 (2010).
3. Li, L., Leid, M. & Rothenberg, E. V. An early T cell lineage commitment checkpoint dependent on the transcription factor Bcl11b. *Science* **329**, 89–93 (2010).
4. Ikawa, T. et al. An essential developmental checkpoint for production of the T cell lineage. *Science* **329**, 93–96 (2010).
5. Li, P. et al. Reprogramming of T cells to natural killer-like cells upon Bcl11b deletion. *Science* **329**, 85–89 (2010).
6. Kueh, H. Y. et al. Asynchronous combinatorial action of four regulatory factors activates Bcl11b for T cell commitment. *Nat. Immunol.* **17**, 956–965 (2016).
7. Albu, D. I. et al. Transcription factor Bcl11b controls selection of invariant natural killer T-cells by regulating glycolipid presentation in double-positive thymocytes. *Proc. Natl Acad. Sci. USA* **108**, 6211–6216 (2011).
8. Vanvalkenburgh, J. et al. Critical role of Bcl11b in suppressor function of T regulatory cells and prevention of inflammatory bowel disease. *J. Exp. Med.* **208**, 2069–2081 (2011).
9. Walker, J. A. et al. Bcl11b is essential for group 2 innate lymphoid cell development. *J. Exp. Med.* **212**, 875–882 (2015).
10. Yu, Y. et al. The transcription factor Bcl11b is specifically expressed in group 2 innate lymphoid cells and is essential for their development. *J. Exp. Med.* **212**, 865–874 (2015).
11. Califano, D. et al. Transcription factor Bcl11b controls identity and function of mature type 2 innate lymphoid cells. *Immunity* **43**, 354–368 (2015).
12. Hogquist, K. A. & Jameson, S. C. The self-obsession of T cells: how TCR signaling thresholds affect fate ‘decisions’ and effector function. *Nat. Immunol.* **15**, 815–823 (2014).
13. Singer, A., Adoro, S. & Park, J. H. Lineage fate and intense debate: myths, models and mechanisms of CD4- versus CD8-lineage choice. *Nat. Rev. Immunol.* **8**, 788–801 (2008).
14. Taniuchi, I. & Ellmeier, W. Transcriptional and epigenetic regulation of CD4/CD8 lineage choice. *Adv. Immunol.* **110**, 71–110 (2011).
15. Ellmeier, W. & Taniuchi, I. The role of BTB-zinc finger transcription factors during T cell development and in the regulation of T cell-mediated immunity. *Curr. Top. Microb. Immunol.* **381**, 21–49 (2014).
16. He, X. et al. The zinc finger transcription factor Th-POK regulates CD4 versus CD8 T-cell lineage commitment. *Nature* **433**, 826–833 (2005).
17. Sun, G. et al. The zinc finger protein cKrox directs CD4 lineage differentiation during intrathymic T cell positive selection. *Nat. Immunol.* **6**, 373–381 (2005).
18. He, X., Park, K. & Kappes, D. J. The role of ThPOK in control of CD4/CD8 lineage commitment. *Annu. Rev. Immunol.* **28**, 295–320 (2010).
19. Setoguchi, R. et al. Repression of the transcription factor Th-POK by Runx complexes in cytotoxic T cell development. *Science* **319**, 822–825 (2008).
20. He, X. et al. CD4-CD8 lineage commitment is regulated by a silencer element at the ThPOK transcription-factor locus. *Immunity* **28**, 346–358 (2008).
21. Muroi, S., Tanaka, H., Miyamoto, C. & Taniuchi, I. Cutting edge: fine-tuning of Thpok gene activation by an enhancer in close proximity to its own silencer. *J. Immunol.* **190**, 1397–1401 (2013).
22. Muroi, S. et al. Cascading suppression of transcriptional silencers by ThPOK seals helper T cell fate. *Nat. Immunol.* **9**, 1113–1121 (2008).
23. Sakaguchi, S. et al. The zinc-finger protein MAZR is part of the transcription factor network that controls the CD4 versus CD8 lineage fate of double-positive thymocytes. *Nat. Immunol.* **11**, 442–448 (2010).
24. Sakaguchi, S. et al. MAZR and runx factors synergistically repress ThPOK during CD8<sup>+</sup>T cell lineage development. *J. Immunol.* **195**, 2879–2887 (2015).
25. Wang, L. et al. Distinct functions for the transcription factors GATA-3 and ThPOK during intrathymic differentiation of CD4<sup>+</sup> T cells. *Nat. Immunol.* **9**, 1122–1130 (2008).
26. Steinke, F. C. et al. TCF-1 and LEF-1 act upstream of Th-POK to promote the CD4<sup>+</sup> T cell fate and interact with Runx3 to silence Cd4 in CD8<sup>+</sup> T cells. *Nat. Immunol.* **15**, 646–656 (2014).
27. Egawa, T. & Littman, D. R. ThPOK acts late in specification of the helper T cell lineage and suppresses Runx-mediated commitment to the cytotoxic T cell lineage. *Nat. Immunol.* **9**, 1131–1139 (2008).
28. Park, J. H. et al. Signaling by intrathymic cytokines, not T cell antigen receptors, specifies CD8 lineage choice and promotes the differentiation of cytotoxic-lineage T cells. *Nat. Immunol.* **11**, 257–264 (2010).
29. Luckey, M. A. et al. The transcription factor ThPOK suppresses Runx3 and imposes CD4<sup>+</sup> lineage fate by inducing the SOCS suppressors of cytokine signaling. *Nat. Immunol.* **15**, 638–645 (2014).
30. Taniuchi, I. Views on helper/cytotoxic lineage choice from a bottom-up approach. *Immunol. Rev.* **271**, 98–113 (2016).
31. Kastner, P. et al. Bcl11b represses a mature T-cell gene expression program in immature CD4<sup>+</sup>CD8<sup>+</sup>thymocytes. *Eur. J. Immunol.* **40**, 2143–2154 (2010).
32. Seo, W. et al. Roles of VWRPY motif-mediated gene repression by Runx proteins during T-cell development. *Immunol. Cell Biol.* **90**, 827–830 (2012).
33. Huang, G. et al. PU.1 is a major downstream target of AML1 (RUNX1) in adult mouse hematopoiesis. *Nat. Genet.* **40**, 51–60 (2008).
34. Tang, B. et al. Genome-wide identification of Bcl11b gene targets reveals role in brain-derived neurotrophic factor signaling. *PLoS ONE* **6**, e23691 (2011).
35. Tanaka, H. et al. Epigenetic Thpok silencing limits the time window to choose CD4<sup>+</sup>helper-lineage fate in the thymus. *EMBO J.* **32**, 1183–1194 (2013).
36. Wakabayashi, Y. et al. Bcl11b is required for differentiation and survival of alpha beta T lymphocytes. *Nat. Immunol.* **4**, 533–539 (2003).
37. Buenrostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y. & Greenleaf, W. J. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods* **10**, 1213–1218 (2013).
38. Wang, L. et al. The zinc finger transcription factor Zbtb7b represses CD8-lineage gene expression in peripheral CD4<sup>+</sup> T cells. *Immunity* **29**, 876–887 (2008).
39. Zetsche, B. et al. Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell* **163**, 759–771 (2015).
40. Jenkinson, S. R. et al. Expression of the transcription factor cKrox in peripheral CD8 T cells reveals substantial postthymic plasticity in CD4-CD8 lineage differentiation. *J. Exp. Med.* **204**, 267–272 (2007).
41. Kakugawa, K. et al. Essential roles of SATB1 in specifying T lymphocyte subsets. *Cell Rep.* **19**, 1176–1188 (2017).
42. Feng, Y. et al. A mechanism for expansion of regulatory T-cell repertoire and its role in self-tolerance. *Nature* **528**, 132–136 (2015).
43. Zheng, Y. et al. Role of conserved non-coding DNA elements in the *Foxp3* gene in regulatory T-cell fate. *Nature* **463**, 808–812 (2010).
44. Kitagawa, Y. et al. Guidance of regulatory T cell development by Satb1-dependent super-enhancer establishment. *Nat. Immunol.* **18**, 173–183 (2017).
45. Lee, H. O. et al. Disregulated expression of the transcription factor ThPOK during T-cell development leads to high incidence of T-cell lymphomas. *Proc. Natl Acad. Sci. USA* **112**, 7773–7778 (2015).
46. Kronenberg, M. & Rudensky, A. Regulation of immunity by self-reactive T cells. *Nature* **435**, 598–604 (2005).
47. Fujita, T. & Fujii, H. Direct identification of insulator components by insertional chromatin immunoprecipitation. *PLoS ONE* **6**, e26109 (2011).
48. Iuchi, S. Three classes of C<sub>2</sub>H<sub>2</sub> zinc finger proteins. *Cell Mol. Life Sci.* **58**, 625–635 (2010).
49. Sankaran, V. G. et al. Developmental and species-divergent globin switching are driven by BCL11A. *Nature* **460**, 1093–1097 (2009).
50. Nishimura, M. et al. VWRPY motif-dependent and -independent roles of AML1/Runx1 transcription factor in murine hematopoietic development. *Blood* **103**, 562–570 (2004).
51. Naoe, Y. et al. Repression of interleukin-4 in T helper type 1 cells by Runx/Cbfb binding to the *Il4* silencer. *J. Exp. Med.* **204**, 1749–1755 (2007).
52. Mali, P. et al. RNA-guided human genome engineering via Cas9. *Science* **339**, 823–826 (2013).
53. Wang, H. et al. One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell* **153**, 910–918 (2013).
54. Egawa, T., Tillman, R. E., Naoe, Y., Taniuchi, I. & Littman, D. R. The role of the Runx transcription factors in thymocyte differentiation and in homeostasis of naive T cells. *J. Exp. Med.* **204**, 1945–1957 (2007).
55. Buenrostro, J. D., Wu, B., Chang, H. Y. & Greenleaf, W. J. ATAC-seq: a method for assaying chromatin accessibility genome-wide. *Curr. Protoc. Mol. Biol.* **109**, 21.29.1–21.29.9 (2015).
56. Mucida, D. et al. Transcriptional reprogramming of mature CD4<sup>+</sup> helper T cells generates distinct MHC class II-restricted cytotoxic T lymphocytes. *Nat. Immunol.* **14**, 281–289 (2013).

## Acknowledgements

We thank C. Miyamoto and R. Chihara for mouse genotyping, N. Yoza for cell sorting, T. Ishikura for ES cell aggregation, Y. Iizuka for microinjection of RNA for CRISPR/Cas9 mediated genome editing. We are grateful to Dr Thomas Boehm for helpful suggestion on analyses of lamprey genome and Dr Eugene Oltz for critical reading of the manuscript. This work was supported by JSPS KAKENHI Grant Number JP 21229008 (I.T.), JP 26293109 (I.T.) and JP 22021045 (I.T.), the Uehara Memorial Foundation (I.T.), Takeda Science Foundation (I.T.) the National Institutes of Health grant R01AI097244-01A1 (T.E.).

## Author contributions

S.K., H.T., W.S., K.K., Y.N., and I.T. performed experiments for phenotypic characterization of *Bcl11b* mutant mice. M.T. and K.M. performed reconstitution experiments. S.M. isolated ES cell clones. T.A.E. performed statistical and comparative genomics analyses. K.N. made mutant *Bcl11b* expression vectors. S.K., A.K., and T.I. performed



sequencing of ChIP libraries. T.E. provided anti-ThPOK antibody. Y.L. and A.M. performed ATAC-seq. B.V. analyzed Japanese lamprey genome. Y.K. and R.K. provided *Bcl11b* germline KO mice. I.T. designed experiments, interpreted data and wrote the manuscript.

### Additional information

**Supplementary Information** accompanies this paper at doi:[10.1038/s41467-017-00768-1](https://doi.org/10.1038/s41467-017-00768-1).

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

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

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017