Bcl11b sets pro-T cell fate by site-specific cofactor recruitment and by repressing Id2 and Zbtb16

Hiroyuki Hosokawa,1,6 Maile Romero-Wolf,1,6 Mary A. Yui,1* Jonas Ungerbäck,1,2 Maria L. G. Quiloan,1 Masaki Matsumoto,3 Keiichi I. Nakayama,3 Tomoaki Tanaka,4,5 and Ellen V. Rothenberg1,*

Multipotent progenitor cells confirm their T cell–lineage identity in the CD4–CD8– double-negative (DN) pro-T cell DN2 stages, when expression of the essential transcription factor Bcl11b begins. In vivo and in vitro stage-specific deletions globally identified Bcl11b-controlled target genes in pro-T cells. Proteomics analysis revealed that Bcl11b associated with multiple cofactors and that its direct action was needed to recruit those cofactors to selective target sites. Regions near functionally regulated target genes showed enrichment for those sites of Bcl11b-dependent recruitment of cofactors, and deletion of individual cofactors relieved the repression of many genes normally repressed by Bcl11b. Runx1 collaborated with Bcl11b most frequently for both activation and repression. In parallel, Bcl11b indirectly regulated a subset of target genes by a gene network circuit via the transcription inhibitor Id2 (encoded by Id2) and transcription factor PLZF (encoded by Zbtb16); Id2 and Zbtb16 were directly repressed by Bcl11b, and Id2 and PLZF controlled distinct alternative programs. Thus, our study defines the molecular basis of direct and indirect Bcl11b actions that promote T cell identity and block alternative potentials.

The zinc-finger transcription factor Bcl11b is required for the development of αβ T cells and most γδ T cells1–3. Its expression initiates precisely during commitment to the T cell lineage4; that is, before expression of the T cell antigen receptor and between the CD4–CD8– double-negative (DN) pro-T cell stages DN2a (CD25+/CD44+c-Ki67+) and DN2b (CD25−/CD44+c-Ki67−). Progression through commitment to the T cell lineage is blocked or highly abnormal in cells that lack Bcl11b5–7. Bcl11b-deficient pro-T cells are less sensitive to Notch signaling and more prone to differentiate into natural killer (NK) cells than are their wild-type counterparts8–10. They also fail to go through β-selection, due to defects in T cell gene expression as well as to abnormal persistence of immature features, including expression of the surface marker c-Ki678. Deletion of Bcl11b after β-selection causes abnormal activation of genes encoding effector molecules8,9,11 and multiple functional defects in later thymocytes and mature T cells12–14. While the importance of Bcl11b in T cell development is clear, its exact mechanism of action is not. Bcl11b can bind to GC-rich sequences in DNA15 and can recruit chromatin-modifying NuRD and SIRT1 complexes16,17, but in pro-T cells it binds mainly to sites in open chromatin showing enrichment for motifs for Ets and Runx transcription factors18–20. Published work has linked Bcl11b to both activation and repression21,22,24,25, with its most consistent effects across development on a core of genes that apparently require repression by Bcl11b in T cells21. Finally, Bcl11b’s effects have striking overlap with the effects of the basic helix-loop-helix (bHLH) transcription factor E2A in early T cells2, yet the basis for this convergence is not known.

This report addresses three questions about Bcl11b’s roles in establishing T cell commitment. First, what are the target genes directly regulated by Bcl11b during T cell commitment? Second, what are the mechanisms that Bcl11b deploys to work as an activator or a repressor at its target sites? We identify direct target loci on the basis of a new criterion for functional sites of Bcl11b action, through its role in recruiting specific cofactors. Finally, how many of the effects of Bcl11b are indirect, and how are they mediated? We show that Bcl11b in pro-T cells blocked expression of the E-protein antagonist Id2 (encoded by Id2) and the innate-response regulator PLZF (encoded by Zbtb16). Id2-mediated suppression of E-protein activity is important for all innate lymphoid cells, including NK cells21–26, while PLZF is crucial for innate-type T cells and for non-cytolytic innate lymphoid cells27–30, and both govern subsets of myeloid and dendritic cells24–26. We find that a gene network that relates Id2 and Zbtb16 to Bcl11b’s function sheds light on the split between the T cell and innate immune cell families of developmental programs.

Results

Impacts of Bcl11b on gene expression in thymocytes at DN2–DN3 stages. Bcl11b regulates a distinct set of target genes during the initial T cell–lineage commitment of fetal-liver-derived precursor cells differentiating in vitro, including many targets that seem to be unique to this developmental period1. To investigate its role during commitment in vivo, we compared DN pro-T cells in mice in which loxp-flanked Bcl11b alleles (Bcl11bfl/fl) were conditionally deleted by Cre recombinase, before or after commitment. For the first we used Cre recombinase expressed in all hematopoietic cells via the hematopoietic compartment–specific gene Vav1 (Vav1-iCre)35; for the second we used Cre expressed via the proximal promoter of the gene encoding the tyrosine kinase Lck (Lck-Cre), from a transgene36 that is first activated in DN2 pro-T cells (Fig. 1a and Supplementary Fig. 1a). The mice also had a Cre-dependent reporter for the expression of yellow fluorescent protein (YFP) from a ubiquitously expressed locus (ROSA26R-YFP), which distinguished cells in which the alleles were deleted from those in which they were not (normal DN2a cells). In mice with wild-type Bcl11b, Vav1-iCre caused all DN thymocytes...
Fig. 1 | Cellular and molecular phenotypes of in vivo deletion of Bcl11b by Vav1-iCre or Lck-Cre. a. Flow cytometry of DN thymocytes from Bcl11b<sup>+/+</sup>, Bcl11b<sup>+/-</sup> and Bcl11b<sup>-/-</sup> (Bcl11b-KO) mice (above plots; representative of n ≥ 5 per genotype) expressing Vav1-iCre (left group) or Lck-Cre (right group) for deletion of Bcl11b and YFP to distinguish cells with (YFP<sup>+</sup>) or without (YFP<sup>-</sup>) deletion. Figure shows gates used for defining DN subpopulations (top row diagrams) and sorting strategy used for purifying YFP CD25<sup>+</sup> and YFP<sup>-</sup>CD25<sup>-</sup> thymocytes used in the RNA expression analysis (bottom row diagrams). Numbers in quadrants (top row) or adjacent to outlined areas (bottom row) indicate percent cells in each. b, c. RNA-seq analysis of genes significantly differentially expressed after deletion of Bcl11b, presented in hierarchically clustered heat maps, showing genes repressed by Bcl11b<sup>+</sup> (b) and Bcl11b-dependent genes (c) (one gene per row; gene symbols, Supplementary Table 1). Samples were from DN2 and DN3 thymocytes sorted from Bcl11b<sup>+/+</sup>Lck-Cre mice (Lck WT) or Bcl11b<sup>+/-</sup>Lck-Cre mice (Lck HE) and YFP<sup>+</sup>CD25<sup>+</sup> cells from Bcl11b<sup>-/-</sup>Vav1-iCre mice (Vav WT), Bcl11b<sup>+/+</sup>Vav1-iCre mice (Vav HE), Bcl11b<sup>-/-</sup> Lck-Cre mice (Lck KO) or Bcl11b<sup>+/+</sup>Vav1-iCre mice (Vav KO) (genotype labels below plot; one column per mouse). Color scales show the differences in expression of each gene in the indicated sample relative to its expression in the Bcl11b<sup>+/+</sup> DN2 sample (log<sub>2</sub> (sample level / control DN2 level) = log<sub>2</sub> FC). In b, c. DEGs are identified in Vav1-iCre deletion samples by comparison of the average of Bcl11b<sup>+/+</sup> Vav1-iCre and Bcl11b<sup>+/+</sup>Vav1-iCre versus Bcl11b<sup>+/+</sup>Vav1-iCre; or, for deletion via Lck-Cre, by the average of Bcl11b<sup>+/+</sup>Lck-Cre DN3 cells versus Bcl11b<sup>+/+</sup>Lck-Cre cells. d, e. Identification of subsets of DEGs (right margin) with lower expression (d) or higher expression (e) in CD25<sup>+</sup> thymocytes after deletion of Bcl11b via Vav1-iCre (Vav KO) than after deletion via Lck-Cre (Lck KO), relative to the average for Bcl11b<sup>+/+</sup>Lck-Cre (Lck KO) samples. Hierarchical clustering analysis showing expression differences of significant DEGs ([log<sub>2</sub> FC] > 1 and FDR < 0.05).
| Genes upregulated in Bcl11b-KO cells at stages DN2 and DN3 (targets repressed by Bcl11b) |
|---------------------------------------------------------------|
| 1700024P16Rik         | 1700112E06Rik         | 2900026A02Rik         | 41340 | 5730559C18Rik         | 9030670O3Rik |
| A4                 | Akap6                | Aoh                 | Apobr | App2                 | Arap3 |
| Arhgap20            | Artf10               | Arl6                | Arsi   | Art2a-ps             | Art2b |
| Arvcf               | Asph                 | Bace1               | Batf3   | Bcar3               | Bcas1os2 |
| Celm1               | Camk2b               | Camkmt              | Card11   | Cass4               | Ccdc157 |
| Ccr2                | Cd160                | Cd163f              | Cd244   | Cd63                | Cd7 |
| Cdt2                | Cd9                  | Ceacam1             | Cers4   | Chad                | Chn2 |
| Chfp                | Cisd3                | Cited4              | Clnd2   | Clnk                | Clvs1 |
| Cnbd2               | Cnks3                | Cnnm2               | Cnr2    | Col9a3              | Colq |
| Coprs               | Cpd                  | Cpeb2               | Cpn7    | Cpt1a               | Ctb2 |
| Cx3cr1              | Cxcr5                | Dab2ip              | Dapi1   | Denn3               | Disc1 |
| Dlx1                | Dlx1as               | Dok3                | Dr7     | Dscam               | Dsp |
| Dyrk3               | Ect1                 | Eea1                | Egln3   | Fam129a             | Fam151a |
| Fam79a              | Fam46a               | Fam71b              | Fbxw8   | Fcer1g              | Fdx1 |
| Fes                 | Fg3f                 | Fina                | Flt3    | Frmd4b              | Gas7 |
| Gimap4              | Gimap6               | Gimap7              | Gimap8  | Glis2               | Goml1 |
| Gpr141              | Gpr183               | Gstm1               | Gstm3   | Gzmc                | Hey2 |
| Hey1                | Hip1r                | Ica1l               | Id2     | Il2rb               | Igap2 |
| Ira3                | Itgb6                | Itgb7               | Ihv1bp  | Ihidc2              | Iklr1 |
| Klf13a              | Klf5a                | Klf3                | c-Kit   | Khdc2               | Khr1 |
| Lag3                | Layn                 | Ltk                 | Lyn     | Man1c1              | Mico1a |
| Myo1e               | Myo1f                | Myo7a               | Nat8l   | Nav2                | Nectin1 |
| Neur13              | Nfl3                 | Npfr1               | Nrf26   | Nrgn                | NHSe |
| Osbp15              | Osbp1                | Paps2               | Pcox1l  | Pde2a               | Pde4a |
| Pear1               | Pik3r2               | Pikg                | Plekhs5 | Polm                | Pou2af1 |
| Ptg                 | Ptprn14              | Ptprn21             | Rab19   | Rapgep2             | Rassf4 |
| Rgs3                | Rhobb1b              | Rnh1                | Rora    | S100a10             | Sprr3 |
| Scoppd              | Scn5a                | Sema3c              | Sema4c  | Serpinb9            | Sh2d1b |
| Shbgrl2             | Siae                 | Slc22a15            | S1c22a3  | S1c206              | S1c35f |
| S1c45a3             | Spa17                | Spaac9              | St6galnac6 | Sult5a1             | Tass1 |
| Tgf3                | Tiam2                | Timp2               | Tiap    | Tj2                 | Tmem126a |
| Tmem198             | Tmem231              | Tmem67              | Tmfrs25  | Tmnl1               | Tfr |
| Trm2                | Trmp1                | Trpm1               | Tspan32  | Tyrobp              | Ubc1p |
| Vsig2               | Wipi1                | Wnt5b               | Xclf1   | Zbb1b6              | Zbb1b7 |
| Zcchc18             | Zfp105               | Zfp296              | Zfp316   | Zfp518b             | Zfp568 |
| Zfp768              |                      |                     |         |                     |       |

| Genes downregulated in Bcl11b-KO cells at stages DN2 and DN3 (Bcl11b-dependent genes) |
|-------------------------------------------------------------------------------------------------|
| 1300002E11Rik         | Abbott2              | Actn1               | Acy3     | Bbof1               | BC025920 |
| Bst1                 | Camkv                | Ccdc153             | Cdc18    | Cd3d                | D3g |
| Cd5                  | Cd6                  | Cldn4               | Comp     | Dap1k               | Dcp1b |
| Def8                 | Dgke                 | Dgkeos              | Dgkg     | Ehd3                | Em1p |
| Evpl                 | F730043M19Rik        | Fam117a             | Fer15    | Frmd4a              | Frmd6 |
| Gbp4                 | Gbp8                 | Glyctk              | Gm15708  | Gm2839              | Gmpr |
| Grasp                | Hist1                | Hmgs2               | Id3      | Ifngr2              | I10a |
| Il21r                | If7i                 | Lgfl1               | Map4k2   | Matk                | Mpi2 |
| Mzb1                 | Mfatt2               | Mjkb2               | Pisd-ps1 | Pisd-ps2            | Pisd-ps3 |
| Pclg1                | Plecka7              | Pllx1d              | Pqlc3    | Rab11fip3           | Rass1 |
| Rgs10                | Rhbdf2               | Sestd1              | Sirt5    | Slc37a1             | Slc5a9 |
| Smim5                | Smox                 | Smyd2               | Sox13    | Spib                | Steap3 |
| Synj2                | Tbx32r               | Tmem131             | Tmem221  | Tmpss4              | Tmfrs1a |
| Tc38                 | Zeb2os               |                     |         |                     |       |

Genes with significantly differential expression, relative to that in wild-type cells, in all three of the following conditions: Lck-Cre–mediated deletion of Bcl11b in vivo, Vav1-iCre–mediated deletion of Bcl11b in vivo, and Cas9-mediated deletion of Bcl11b in DN2b–DN3 cells differentiated in vitro from BM-derived precursor cells (expression values, Supplementary Tables 1 and 2); genes listed in alphabetical order here (complete list of genes in each comparison, Supplementary Table 3).
to express YFP (Fig. 1a, left), whereas Lck-Cre activated YFP only in cells at the DN2b stage and later (Fig. 1a and Supplementary Fig. 1a). Thus, Lck-Cre deleted genes only after Bcl11b would normally be turned on. Homozygous Bcl11b−/− mice bred to express either of these Cre-encoding transgenes showed arrest of the T cell lineage; Bcl11b-deficient cells retain a DN2a-like phenotype when Bcl11b is deleted either before commitment or after commitment. Scid.adh.2c2 (bottom right), DN3-like cell line. b, SDS-PAGE and silver staining of total extracts of Scid.adh.2c2 cells mock transduced (left lane) or transduced to express Myc- and Flag-tagged Bcl11b (right lane), followed by two-step affinity purification; middle lane left unloaded; right margin, identification of Bcl11b in complexes; left margin, molecular size, in kilodaltons (kDa). All visible bands were subjected to liquid chromatography–tandem mass spectrometry analysis (Supplementary Table 5). c, Gene-ontology analysis of molecules that interact with Bcl11b, showing the top ten gene-ontology (GO) terms (left) and the enrichment for those terms in those complexes relative to their abundance in the mouse genome (middle left column). d, Bcl11b-repressor complexes detected by liquid chromatography and tandem mass spectrometry; magenta indicates cofactors discussed in detail in Results. e, Immunoblot analysis (IB) of total extracts of Scid.adh.2c2 cells as in b (above lanes) after immunoprecipitation with monoclonal antibody to Flag (IP: anti-Flag) left or of nuclear lysates of such cells without immunoprecipitation (right), probed with antibodies to (anti-) components of the complexes in d (indicated along right margin); gels cropped to focus on protein species migrating near the mobilities of the size markers (left margin). Data are representative of two independent experiments (b,e).

The gene-expression patterns in DN thymocytes from mice with Bcl11b deleted by either Vav1-iCre or Lck-Cre showed that about 300 genes were reproducibly upregulated in YFP+ DN2-like thymocytes homoyzous for knockout of Bcl11b as compared to their expression in YFP+ DN2 and DN3 thymocytes with wild-type Bcl11b or heterozygous knockout of Bcl11b (control cells) (Fig. 1c), which defined the genes normally repressed by Bcl11b. About 220 genes were significantly downregulated in these Bcl11b-knockout cells (Fig. 1c), which defined genes dependent on Bcl11b. These criteria (a false-discovery rate (FDR) of <0.05, an increase or decrease in expression of >2-fold (referred to as |log2 FC| >1), and average reads per kilobase million (RPKM) of >1). Supplementary Tables 1 and 2) defined ‘differentially expressed genes’ (DEGs) regulated by Bcl11b in young adult thymocytes. Although the Bcl11b-knockout

Fig. 2 | Identification of Bcl11b interacting molecules in early T cells. a, Phenotype of Bcl11b-deficient pro-T cells: Bcl11b is turned on during commitment to the T cell lineage; Bcl11b-deficient cells retain a DN2a-like phenotype when Bcl11b is deleted either before commitment or after commitment. Scid.adh.2c2 (Supplementary Fig. 1a,b). Thus, excision of Bcl11b was able to generate the YFP-c-KithiCD25+ phenotype through retrograde-like differentiation from cells that had previously reached DN2b stage after activating Bcl11b initially.

The gene-expression patterns in DN thymocytes from mice with Bcl11b deleted by either Vav1-iCre or Lck-Cre showed that about 300 genes were reproducibly upregulated in YFP+ DN2-like thymocytes homoyzous for knockout of Bcl11b as compared to their expression in YFP+ DN2 and DN3 thymocytes with wild-type Bcl11b or heterozygous knockout of Bcl11b (control cells) (Fig. 1c), which defined the genes normally repressed by Bcl11b. About 220 genes were significantly downregulated in these Bcl11b-knockout cells (Fig. 1c), which defined genes dependent on Bcl11b. These criteria (a false-discovery rate (FDR) of <0.05, an increase or decrease in expression of >2-fold (referred to as |log2 FC| >1), and average reads per kilobase million (RPKM) of >1). Supplementary Tables 1 and 2) defined ‘differentially expressed genes’ (DEGs) regulated by Bcl11b in young adult thymocytes. Although the Bcl11b-knockout
cells resembled normal DN2a thymocytes, their gene-expression patterns distinguished the mutant cells from any normal Bcl11b+/+ cell subsets. Highly robust effects were seen in multiple samples; for example, for Bcl11b itself, for Zbtb16 and Id2 (both repressed by Bcl11b) and for Cd6 (a Bcl11b-dependent gene encoding the surface receptor CD6) (Supplementary Fig. 2). Interestingly, certain DEGs also showed partial de-repression in YFP+ cells heterozygous for knockout of Bcl11b (Fig. 1b and Supplementary Table 1).

At a subset of the loci identified above, the effects of deletion of Bcl11b before commitment (via Vav1-iCre) were more severe than the effects of deletion after commitment (via Lck-Cre), and Vav1-iCre-mediated deletion had more severe effects on total thymus cellularity (Supplementary Fig. 1c). About 40 Bcl11b-dependent genes failed to be turned on in cells in which Bcl11b was deleted via Vav1-iCre (‘Vav1-iCre-knockout cells’) but were expressed somewhat in cells in which Bcl11b was deleted via Lck-Cre (‘Lck-Cre-knockout cells’) (Fig. 1d), while ~85 targets repressed by Bcl11b were overexpressed more in the Vav1-iCre-knockout cells than in the Lck-Cre-knockout cells (Fig. 1e). This difference suggested that even transient Bcl11b expression in the Lck-Cre deletion model provided some function needed for T-cell development. The most timing-sensitive genes included Cd3g, Cd3d, Cd5e and Dnmt3, among the Bcl11b-dependent genes, and Pou2af1, Tyrobp, Cd7, Iggae, Iggb7, Klf2, Trpm1, Cd163l1, Cric1, Cric4 and Tmim1 (genes associated with progenitors, γδ T cells and alternative lineages), among genes repressed by Bcl11b. However, most DEGs required Bcl11b both during commitment and continuously after commitment, for activation or especially for repression in pro-T cells.

For mechanistic experiments assessing Bcl11b function, we also acutely disrupted Bcl11b via the endonuclease Cas9 in pro-T cells differentiating in vitro and defined the genes affected by this (Supplementary Fig. 3a and Supplementary Table 2). Bone marrow (BM) precursor cells from C57BL/6 (B6) mice with expression of Cas9 via ROSA26 and transgenic expression of the anti-apoptotic protein Bcl2 (B6.Rosa26-Cas9;Bcl2-Ig called ‘Cas9-Bcl2 cells’ here) were cultured for 7 d on OP9-DL1 stroma (OP9 mouse BM stromal cells expressing the DL1 ligand for Notch signaling receptors), then were transduced with retroviral vector encoding small guide RNA (sgRNA) targeting Bcl11b and, 7 d later, were harvested for analysis by RNA-based next-generation sequencing (RNA-seq). Most of these genes also overlapped the significant DEGs both in cells with deletion via Lck-Cre and those with deletion via Vav1-iCre in vivo (Supplementary Fig. 3b and Supplementary Table 3). The significant DEGs in all three (Table 1) included most of the targets previously reported in fetal liver–derived pro-T cells (Supplementary Fig. 3c). Notably, the transcription factor–encoding genes Id2 and Zbtb16 were highly significant targets of repression by Bcl11b in DN2–DN3 cells in every case. Thus, during pro-T cell commitment, Bcl11b activated and repressed several hundred genes encoding molecules important for T cell identity.

A challenge of functional specificity. To determine which genes were directly regulated by Bcl11b, we analyzed Bcl11b by chromatin immunoprecipitation followed by deep sequencing (ChIP-seq). However, despite the prevalence of repressive effects of Bcl11b, binding of Bcl11b across the genome seemed to occur mainly at open chromatin sites with active marks18,28 globally associated with active genes and showed no distinction between unregulated loci and those with any response to deletion of Bcl11b (Supplementary Fig. 4a and Supplementary Table 2). We investigated whether distinct motifs were bound at functional sites. Most sites bound by Bcl11b in DN3 pro-T cells show enrichment for Ets and Runx motifs15; we compared those with GC-rich sequences originally reported as a cognate site15 and with three newly reported motifs for the binding of Bcl11b to protein-binding microarrays15. While matches to the newly defined sites (log odds ≥ 5) were found at 10–18% of sites occupied by Bcl11b, DEGs did not show enrichment for those matches (Supplementary Table 4). Genomic regions near DEGs actually showed substantial depletion of Bcl11b-binding sites in CpG islands (Supplementary Fig. 4b). Bcl11b occupancy and target-site motifs were thus insufficient to distinguish sites of Bcl11b function.

Bcl11b interacts with several ‘repressor’ complexes. We reasoned that the sites at which Bcl11b exerts regulatory functions could be identified through its local interactions with specific protein factors. To identify the functional components of Bcl11b complexes that control gene expression in pro-T cells, we transduced Scid.adh.2c2 cells, which are a DN3-like cell line29,39 (Fig. 2a), with Myc- and Flag-tagged Bcl11b, then subjected the Bcl11b-containing protein complexes to two-step affinity purification followed by SDS-PAGE and silver staining (Fig. 2b). Analysis by liquid chromatography and tandem mass spectrometry identified more than 300 molecules for which the complexes showed supra-threshold enrichment, based on mass spectrometry (MASCOT) enrichment scores (Supplementary Table 5). The complexes showed considerable enrichment for the bifunctional transcription factor Runx1 and proteins annotated as being involved in ‘negative regulation of gene expression,’ ‘transcriptional regulation’ and ‘chromatin remodelling’ (Fig. 2c). Bcl11b's
**Fig. 4 | Bcl11b-dependent cofactor peaks around the major Bcl11b target genes.** a–d, ChIP-seq analysis of the binding of Bcl11b, Chd4, Hdac2, Mta2, Rest, Ring1b, LSD1, or Runx1 (far left margin) to the loci of Idd2 (a), Zbbt16 (b), Tnni1 (c) and Cdf163I1 (d) (above plots) in wild-type (WT) DN3 cells and Bcl11b-deficient (Bcl11b-KO) CD25+ cells (left margin). Also shown: ChIP-seq of H3K27Ac in such cells (second group from bottom), and RNA-seq analysis of 100kb locus (Supplementary Table 2). Non-DEGs were defined in these samples and samples in which Bcl11b was deleted. [log2 FC] > 1, FDR < 0.05 and RPKM > 1 (Supplementary Table 2). Non-DEGs were defined in these comparisons as [log2 FC] < 0.05 and RPKM > 3. Cofactor and H3K27Ac binding was scored from peaks scored as ‘reproducible’ in two independent experiments. P values, versus non-DEG (two-sided Fisher’s exact test).
Fig. 5 | Effect of cofactor deletion on the expression of Bcl11b target genes. a. Flow cytometry (as in Fig. 3a) of primary DN cells derived from BM precursor cells obtained from Cas9 mice and transduced with control sgRNA (sgControl) or sgRNA targeting Bcl11b (sgBcl11b), Chd4 (sgChd4), Mta1 and Mta2 (sgMta1_2), Rest (sgRest), Ring1a and Ring1b (sgRing1a_b), LSD1 (sgLSD1) or Runx1 (sgRunx1) (above plots) for CRISPR-Cas9-mediated deletion in primary DN cells. Cells were taken after culture for 14 d on OP9-DL1 stroma (Supplementary Fig. 3a). Numbers in quadrants indicate percent cells in each, among Lin–CD45 (sgRest), Ring1a (sgRing1a_b), LSD1 (sgLSD1) or Runx1 (sgRunx1) (above plots) for CRISPR-Cas9–mediated deletion in primary DN cells. Cells were obtained from Cas9 mice and transduced with control sgRNA (sgControl) or sgRNA targeting Bcl11b (sgBcl11b), Mta1 (sgMta1_2), Chd4 (sgChd4) or Bcl11b (sgBcl11b). Percentages of Lin–CD45– cells were obtained as in Fig. 3a, right. b. Proportion of genes that are repressed by cofactors (increased in expression by at least twofold after deletion of cofactors indicated along horizontal axis) among Bcl11b-dependent genes or control genes (key). c, Proportion of genes that are repressed by cofactors (increased in expression by at least twofold after deletion of cofactors indicated along horizontal axis) among genes normally repressed by Bcl11b or control genes (key). d, Proportion of genes that are cofactor dependent (reduced in expression by at least twofold after deletion of cofactors indicated along horizontal axis) among Bcl11b-dependent genes or control genes (key). (from Supplementary Table 6) *P<0.05, **P<0.001 and ***P<0.000001 (two-sided Fisher’s exact test). Data are representative of three independent experiments (a) or are based on peaks called as reproducible in two replicate samples (ChIP-seq) or on two independent replicates (RNA-seq) (b–d).
interaction partners for which the complexes showed the greatest enrichment included multiple members of the NuRD (‘nucleosome-remodeling deacetylase’) histone-deacetylase complex, Rest (‘RE-1-silencing transcription factor’) transcriptional repressor complex (also known as the NRSF (‘neuron-restrictive silencer factor’) complex), and the Kdm1a histone deacetylase–recruitment complex (also known as the LSD1 (‘lysine-specific demethylase 1’) complex), with lower but still substantial scores for the Polycomb repressor complex PRC1, in confirmation of published evidence of a Bcl11b–NuRD association14 (Fig. 2d). The association of major components of those complexes (Chd4, Mta2, Rnf2 (Ring1b), Rest, LSD1 and Hda2) with Bcl11b was confirmed by co-immunoprecipitation (Fig. 2e). Bcl11b has also been found in SWI/SNF chromatin-remodeling complexes\(^{18}\), and the complexes also showed specific enrichment for the SWI/SNF component Smarca4 (Brg1) (Supplementary Table 5) but not most other SWI/SNF components. We note that although LSD1, Chd4, Mta2 and Ring1b act as components of repressor complexes, in specific contexts they too are reported to contribute to gene activation\(^{41–45}\). Thus, both ‘repressor complexes’ and Runx1 might serve roles in Bcl11b-mediated positive or negative regulation of genes.

**Identification of Bcl11b-dependent cofactor-binding sites.** To determine whether the differential gene expression in primary pro-T cells was linked to genomic regions at which specific cofactor assemblies might be nucleated by Bcl11b, we performed ChIP-seq analysis of wild-type pro-T cells and those in which Bcl11b was deleted. To obtain the large number of cells needed, we used in vitro differentiation cultures with deletion of Bcl11b\(^{38}\) effected by activation of the tamoxifen-inducible Cre-ERT2 construct by 4-OH tamoxifen (4-OHT) (Supplementary Fig. 4c). On day 7, before treatment with 4-OHT, most cells showed a DN2a–DN2b phenotype (Lin−CD45+ c-Kit−CD25+; Supplementary Fig. 4d). At 5 d after removal of 4-OHT, Cre-ERT2-expressing control cells with wild-type Bcl11b had efficiently progressed to a c-Kit− DN3 stage, while Bcl11b−/− deficient cells were still characteristically c-Kit+ (Fig. 3a and Supplementary Fig. 4e). ChIP-seq analysis identified more than 25,000 reproducible Bcl11b peaks in Bcl11b+/+Cre-ERT2 (control) DN3 cells (including 82% of peaks previously reported in B6 DN3 cells without 4-OHT treatment\(^{7}\); Supplementary Fig. 5f), and these peaks almost completely disappeared from cells in which Bcl11b was deleted (Fig. 3b). Of the factors assessed, only Brg1 could not be mapped through the use of commercially available reagents (data not shown). In the control cells, Chd4, Mta2, Rest, Ring1b, LSD1 and Runx1 each bound at from ~6,500 sites (Rest) to ~33,000 sites (Runx1). Those peaks overlapped Bcl11b peaks to different extents (Fig. 3c). However, the results clearly showed that distinct subsets of peaks for each cofactor depended on the presence of Bcl1b for their recruitment. Deletion of Bcl1b caused large fractions of some cofactor peaks to disappear (Mta2 and Rest) or relocate (Chd4, Ring1b, LSD1 and Runx1) (Fig. 3c). In each case, most of the Bcl11b-dependent peaks coincided with sites bound by Bcl11b in wild-type cells (Fig. 3c), which indicated that binding of Bcl11b itself was needed to recruit that cofactor to such sites. The number of Mta2 peaks was especially sharply reduced by deletion of Bcl11b, although the abundance of Mta2 protein in Bcl11b-deficient cells was comparable to that in wild-type DN3 cells (Supplementary Fig. 4g); this suggested that most of its genome-wide associations in these cells depended on Bcl11b.

Only a small subset of Bcl11b sites showed Bcl11b-dependent binding of cofactors, among the tens of thousands of Bcl1 peaks across the genome. Some of the genes repressed by Bcl11b not only had Bcl11b-dependent cofactor peaks (Fig. 4a–d) but also showed newly generated cofactor peaks that occupied distinct sites when Bcl11b was deleted (Fig. 4a,b). Motif ‘preference’ for cofactor binding changed when Bcl11b was absent. Bcl11b-dependent cofactor peaks genome-wide showed considerable enrichment for motifs not only for the Ets and Runx families but also for the bHLH family of transcription factors (Supplementary Fig. 5a,b). A minority of these sites included the motifs defined by protein-binding microarrays\(^{37}\), similar to Bcl11b sites overall (Supplementary Table 4). However, the ‘new’ cofactor peaks that appeared specifically when Bcl11b was deleted had sharply altered motif distributions, with bHLH motifs (‘E2A’ or ‘Ptf1a’) being much less common and motifs for the bZIP and HMG families of transcription factors and other motifs more commonly showing enrichment (Supplementary Fig. 5c). Bcl11b at CpG islands overlapped mainly sites at which Chd4, Ring1b, LSD1 or Runx1 was engaged whether Bcl11b was deleted or not (Supplementary Fig. 5d). Thus, Bcl11b both facilitated the binding of cofactors to a subset of sites occupied by Bcl11b and antagonized the binding of cofactors to other sites.

**Bcl11b alters cofactor binding at functional target loci.** Bcl11b-dependent cofactor peaks were found around genes that Bcl11b repressed\(^{3}\) (Fig. 4a–d), both in primary DN3 cells and in Scid.adh.2c2 cells (Supplementary Fig. 6a–d). For Id2, Bcl11b-dependent cofactor peaks were seen not on the gene body itself but substantially upstream and downstream of the Id2 locus (Fig. 4a, Supplementary Fig. 6a), consistent with the extended regulatory system for this gene\(^{21}\). Similarly, for Zbtb16 and Tnni1, only one of several sites seemed to be Bcl11b dependent, and for Cd163l1, the various cofactors varied in Bcl11b dependence at different sites (Fig. 4b–d and Supplementary Fig. 6b–d), which suggested that Bcl11b might interact separately with distinct complex subcomponents. Finally, selective Bcl11b-dependent recruitment of cofactors was also seen at genes that were positively regulated by Bcl11b in pro-T cells, such as the Cd3gde cluster and Cad6 (Supplementary Fig. 6e,f). The bifunctional transcription factor Runx1 was frequently recruited to Bcl11b sites around all classes of targets. Thus, the binding of Bcl11b recruited different cofactors to specific subsets of its genomic sites, including both positively regulated loci and negatively regulated loci.

Statistical evidence indicated that the co-recruitment identified above was functionally relevant. In contrast to results obtained for the simple binding of Bcl11b itself, sites of Bcl11b-dependent recruitment of cofactors were found to be enriched at Bcl11b-regulated loci (defined in Supplementary Fig. 4a and Supplementary Table 2). Among the DEGs linked to direct binding of Bcl11b, Bcl11b-dependent cofactor peaks and ‘new’ cofactor peaks that appeared only when Bcl11b was deleted were considerably over-represented, relative to their representation among genes that also had Bcl11b binding but did not change expression when Bcl11b was deleted (Fig. 4e,f). Target genes normally repressed by Bcl11b showed particular enrichment for Bcl11b-dependent recruitment of Runx1 (Fig. 4e), while they showed depletion for Bcl11b-dependent recruitment of Rest. Mta2 accompanied the binding of Bcl11b at DEGs and non-DEGs alike, but targets repressed by Bcl11b showed significant enrichment for Bcl11b-dependent recruitment of Cd4, Ring1b and LSD1 (Fig. 4e). As expected for mediators of repression, these sites had minimal association with the active histone mark H3K27Ac. At sites linked to Bcl11b-dependent DEGs (Fig. 4e), there was also highly significant enrichment for Bcl11b-dependent recruitment of Runx1, Chd4, LSD1 and Ring1b, but with substantial enrichment for H3K27Ac as well. Interestingly, DEGs normally repressed by Bcl11b showed the most specific enrichment for ‘new’ cofactor peaks (Chd4, Ring1b, LSD1, Runx1 and Ring1b) that appeared only when Bcl11b was absent, as H3K27Ac marking of these genes also increased in cells in which Bcl11b was deleted (Fig. 4f). Although ‘new’ sites for Runx1 and LSD1 themselves rarely overlapped sites with substantial occupancy by Bcl11b in wild-type cells (Figs. 3c and 4a,b), it seemed that Bcl11b’s presence normally inhibited their occupancy. Thus, Bcl11b-dependent localization of cofactors, in the
Fig. 6  |  Id2 is involved in establishment of phenotypes of Bcl11b-deficient cells. a, Flow cytometry (left; as in Fig. 3a) of BM-derived precursor cells transduced with various combinations of sgRNA (above plots) and cultured on OP9-DL1 stroma (Supplementary Fig. 3a). Right, histogram summary of c-Kit expression by those cells (top right) and mean fluorescent intensity (MFI) of c-Kit expression by cells transduced with sgRNA or combinations of sgRNA indicated in key (bottom right). Numbers in quadrants (left) indicate percent cells in each, among Lin–CD45 by those cells (top right) and mean fluorescent intensity (MFI) of c-Kit expression by cells transduced with sgRNA or combinations of sgRNA indicated in key (above plots) and cultured on OP9-DL1 stroma (Supplementary Fig. 3a). Right, histogram summary of c-Kit expression with antibody to PLZF (Anti-PLZF) or isotype-matched control antibody (Isotype) (key), and mean fluorescent intensity of PLZF expression by those cells (right). Each symbol (bottom right) represents an individual replicate. *P = 0.043 (two-sided Student’s t-test). b, Intracellular staining (left) of PLZF in cells transduced with various combinations of sgRNA (key), assessed with antibody to PLZF (Anti-PLZF) or isotype-matched control antibody (Isotype) (key), and mean fluorescent intensity of PLZF expression by those cells (right). Each symbol (bottom right) represents an individual replicate. *P = 0.011 and **P = 0.005 (two-sided Student’s t-test). c, RNA-seq analysis of Bcl11b-regulated genes, showing effects of mutation of Id2 and Zbtb16. Heat maps show a subset of genes repressed by Bcl11b (top) and Bcl11b-dependent genes (bottom) in CD25+ purified DN cells derived from BM precursor cells transduced by sgRNA as in a (above plots) and purified (DN cells), showing hierarchical clustering of expression ratios relative to controls. d, Effects of double knockout of Bcl11b together with Id2 or Zbtb16 compared with knockout of Bcl11b alone. Effects on genes normally repressed by Bcl11b (n = 410), presented as the change in mRNA expression (log2) in cells transduced with sgControl versus (vs) those transduced with sgBcl11b (vertical axis) plotted against that in cells transduced with sgControl versus those transduced with sgBcl11b plus sgId2 (top row) or sgBcl11b plus sgZbtb16 (bottom row) (horizontal axis). Results are shown for genes with the most highly enriched KEGG pathways: ‘Cytokine–cytokine receptor interaction’ pathway (left; P = 1.35 × 10^-13; n = 18 genes) and ‘Natural killer cell–mediated cytotoxicity’ (right; P = 3.09 × 10^-13; n = 10 genes). e, Effects on Bcl11b-dependent genes (n = 384), presented as in d, showing results for the KEGG pathway ‘T cell receptor signaling’ pathway (P = 5.32 × 10^-12; n = 12 genes). Data are representative of three independent experiments with three biological replicates (a,b; mean ± s.d.) or are pooled from two experiments (c-e).
**Fig. 7 | Id2 and Zbtb16 serve key roles in the Bcl11b-mediated exclusion of alternative fates.**

**a.** Frequency of Lin+, NK1.1+ or CD11c+ cells (horizontal axis) among CFP+ cells transduced with various combinations of sgRNA (key) and cultured on OP9-DL1 stroma (top) or OP9-Mig stroma (bottom) (experimental protocol, Supplementary Fig. 11a; representative flow cytometry data, Supplementary Fig. 11d). *P < 0.05 and **P < 0.01 (two-sided Student\'s t-test).

**b.** Experimental test for cell autonomy of altered differentiation by cells in which Bcl11b was disrupted with or without disruption of Id2 or Zbtb16 (experimental plan, Supplementary Fig. 12a). Flow cytometry of primary DN cells derived from BM precursor cells after culture for 7 d on OP9-DL1 stroma, then split into three aliquots separately transduced with sgRNA targeting the various gene combinations to be tested (left margin) (all detected with a CFP reporter), or with sgRNA targeting Bcl11b alone, with a human nerve growth factor receptor (hNGFR) marker (sgBcl11b/hNGFR), as a reference standard, or with empty vector plus green fluorescent protein (Mock/GFP), as a control. For each experimental comparison, the three populations were then pooled (left margin) 3 d after infection, transferred to OP9-Mig stroma and cultured for 4 d. Results show forward scatter (FSC) versus the expression of Lin, NK1.1 or CD11c in CFP+ cells (above plots). Numbers below outlined areas indicate percent cells in each.

**c, d.** Important roles for Id2 and PLZF in establishing the phenotypes of Bcl11b-deficient pro-T cells: effects on alternative-lineage markers (c) and on specific genes (d). Bcl11b directly represses the expression of Id2, Zbtb16 and Kit, and Id2 is partially involved in Bcl11b-mediated downregulation of the expression of Zbtb16 and Kit. The upregulation of Zbtb16 (encoding PLZF) in Bcl11b-deficient cells supports the generation of NK1.1+ cells and represses CD11c+ cells. On the other hand, Id2 positively contributes to the generation of NK1.1+ and CD11c+ cells. The activation of genes encoding molecules in the ‘cytokine–cytokine receptor interaction’ pathway and antagonism of the E protein–mediated activation of T cell genes (d). The rare repression of an E protein–dependent gene (Bcl11a) by Bcl11b (d) shows the gene specificity of its effects. Data are individual values of three biological replicates (a; mean ± s.d.) or are representative of three independent experiments (b).
context of both recruitment and inhibition of recruitment, was a much stronger predictor of Bcl11b functionality at genomic sites than was Bcl11b binding alone.

Gene repression is complex, and it is not known how many binding sites are usually needed for effective repression of a target by Bcl11b. However, we investigated whether Bcl11b-dependent co-recruitment could identify sites that exerted particularly strong effects in repression. We designed sgRNAs to disrupt either of two sites flanking Id2 (Supplementary Fig. 7a) or one site downstream of Tnni1 (Fig. 4c) in Scid.adh.2c2 cells and tested the cells in which these sites had been disrupted for deregulation of the target gene. Through the use of transduction of Cas9 plus sgRNA followed by cloning, we identified cells in which the target sites were completely disrupted, as shown by genomic qPCR (Supplementary Fig. 7b,c).

Loss of the Tnni1 candidate silencer site 'Sil +14K' elevated Tnni1 expression at least fivefold ($P = 9.52 \times 10^{-4}$) (two-sided Student's t-test); Supplementary Fig. 7c). Deletion of one Id2 candidate site ('Sil +40K') consistently raised Id2 expression in the Scid.adh.2c2 cells above background, despite the presence of numerous other Bcl11b sites around the locus ($P = 9.49 \times 10^{-4}$) (two-sided Student's t-test); however, deletion of another Id2 candidate site ('Sil –600K') did not (Supplementary Fig. 7a). Thus, Bcl11b-dependent recruitment of cofactors helped in identifying repression functions per site in the genome.

**Functional effect of cofactors on the activities of Bcl11b.** Recruitment of the cofactors themselves contributed functionally to Bcl11b's effects, as shown when we compared the effects of acute disruption of Bcl11b with those of the disruption of the genes encoding the cofactors. We generated sgRNAs directed against coding regions for Chd4, Mta1–Mta2, Rest, Ring1a–Ring1b, LSD1 or Runx1, each with a cyan fluorescent protein (CFP) reporter, and confirmed that these eliminated expression of the target protein in Cas9-transduced Scid.adh.2c2 cells (Supplementary Fig. 8a). We then transduced those into Cas9-Bcl2 primary cells, in parallel with sgRNA directed against Bcl11b alone or control sgRNA (Supplementary Fig. 3a), to compare their effects directly during in vitro T cell development. At the time of transcription, after 7 d of culture on OP9-DL1 stromal cells, most of the Cas9-Bcl2 primary cells were at a DN2 stage (c-Kit+CD25+); Supplementary Fig. 8b). After 7 more days, transduced (CFP+) control cells (Supplementary Fig. 8c) had progressed into the DN3 stage (Fig. 5a), whereas cells transduced with Bcl11b-specific sgRNA showed the typical c-Kit++ DN2-like phenotype. The effects of cofactor deletion on surface phenotype were milder than those of Bcl11b deletion (Fig. 5a), although RNA transcript structures confirmed the biallelic deletions at the targeted sites in the loci encoding these cofactors (Supplementary Fig. 8d). Cofactor deletion specifically affected RNA expression, more frequently for genes regulated by Bcl11b that were directly bound by Bcl11b ($\log FC > 1$; Fig. 5b–d and Supplementary Table 6) than for 'background' genes expressed independently of Bcl11b ($\log FC < 0.5$), at substantial levels in cells transduced with control sgRNA (RPKM $> 3$), and without scorable binding of Bcl11b (Fig. 5b, right). Most genes repressed by Bcl11b that were directly bound by Bcl11b were also de-repressed after deletion of at least one of the cofactors (Fig. 5b, left), indicative of distinct gene-specific patterns of cofactor response. Among Bcl11b-dependent targets, many were downregulated after deletion of Mta1, Mta2 or Runx1 (Fig. 5b, middle). As summarized (Fig. 5c,d), deletion of Mta1, Mta2 or Runx1 caused the largest number of significant changes in gene expression (adjusted $P$ value, $< 0.1$ (EdgeR); Supplementary Table 6) among both Bcl11b-repressed genes and Bcl11b-dependent genes, and these effects were concordant with the effects of the deletion of Bcl11b for $80–90\%$ of genes significantly regulated by Runx1 and $\sim 90\%$ of genes significantly regulated by Mta1 and Mta2. Thus, cofactor recruitment was functionally significant but with target-gene-specific functional requirements.

**Cofactor recruitment by Bcl11b in marking of the loci encoding TCRβ and TCRγ.** Bcl11b-deficient mice can generate T cells of the $\gamma$δ lineage but not those of the $\beta$ chain, associated with a failure of rearrangement of the $\beta$-chain variable region and diversity–joining regions ($\gamma_v$-$\delta_J$) of the T cell antigen receptor (TCR) ($\beta$-$\gamma$). In BM-derived pro-T cells developing in vitro, Bcl11b and Runx1 co-occupied multiple sites across both the complex encoding the TCR $\beta$-chain (Tcrb) and the complex encoding the TCR $\gamma$-chain (Tcrg). However, these genomic regions differed considerably in the extent to which their binding of Runx1 depended on Bcl11b (Supplementary Fig. 9). Whereas the binding of Runx1 across the $\gamma_v$-encoding segments was highly dependent on Bcl11b (Supplementary Fig. 9a), its binding across the whole Tcrg complex was largely independent of Bcl11b (Supplementary Fig. 9b). Notably, the difference between these loci was not revealed by differential RNA expression, as the pro-T cells cultured in vitro to this early DN3-like stage still showed minimal transcription of genes encoding $\gamma_v$ regions, with or without Bcl11b (Supplementary Fig. 9a). Instead, the difference in the Bcl11b-dependent recruitment of cofactors might indicate locus-specific roles in establishing permissive chromosome structure that could underlie later effects of Bcl11b on DNA rearrangement.
Table 7), which indicated that most of these 410 genes targeted for repression by Bcl11b did not depend on Id2 or PLZF for their expression (DEGs with FDR < 0.05, versus control (410 for Bcl11b-KO and 299 for either DKO); Fig. 6c and Supplementary Table 7). In contrast, among the 349 Bcl11b-dependent genes, the downregulation of many genes encoding molecules involved in the T cell program was ameliorated when Bcl11b was deleted together with Zbtb16 or especially with Id2 (Fig. 6c). Double deletion of Id2 with Bcl11b provided protection to well over half of the Bcl11b-dependent genes (349 DEGs for Bcl11b-KO with FDR < 0.05 versus control, but only 72 for Bcl11b-Id2 DKO; Supplementary Table 7), which suggested that many of these genes were dependent on E proteins.

KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway analysis showed that deletion of Bcl11b preferentially upregulated genes encoding molecules in the ‘cytokine–cytokine receptor interaction’ and ‘NK cell–mediated cytotoxicity’ pathways, while it reduced expression of genes encoding molecules in the ‘T cell receptor signaling’ pathway. Although genes encoding molecules in the ‘NK cell–mediated cytotoxicity’ pathway were upregulated in Bcl11b-KO cells with or without Id2 or Zbtb16, there was much less upregulation of genes encoding molecules in the ‘cytokine–cytokine receptor interaction’ pathway in DKO samples, especially in Bcl11b-Id2-DKO samples, than in Bcl11b-KO samples. Double deletion also resulted in less inhibition of genes encoding molecules in the ‘T cell receptor signaling’ pathway (Fig. 6d,e). Thus, despite the weak overall effects of Id2 and PLZF on the gene expression profiles of Bcl11b-deficient cells, they had effects specifically on the regulation of selective sets of genes encoding molecules related to the cytokine–cytokine receptor and TCR signaling pathways.

Double deletion of Id2 and Bcl11b did not always attenuate the effects of the loss of Bcl11b; for some targets repressed by Bcl11b, it amplified the effects of the disruption of Bcl11b (Supplementary Table 7). Certain genes that are preferentially expressed in TCRγδ+ intraepithelial lymphocytes, including Heyl and Cited4, as well as Bcl11a (which encodes a B cell– and progenitor-cell–associated regulatory molecule), were upregulated much more in Bcl11b-Id2-DKO samples than in Bcl11b-KO samples. These results indicated that Bcl11a, Heyl and Cited4 depended on positive regulation by E proteins even though they were also targets of repression by Bcl11b. Thus, the upregulation of Id2 could sometimes mediate but in other cases mask the full spectrum of the regulatory effects of Bcl11b.

### Discussion

The importance of Bcl11b for commitment to the T cell lineage has raised the question of how this factor works to promote and consolidate T cell identity. Unlike the effects of the transcription factors Pax5 and EBF1 on the B cell lineage, the effects of Bcl11b on gene expression in T cells seem to be more limited and more related to immune system–activation thresholds (‘adaptive’ versus ‘innate-like’) than to T cell identity itself. A published study has shown that Bcl11b binds to numerous sites throughout the active topological domains of the genome in early T cells, suggestive of a possible global organizational role for Bcl11b, but this leaves open its gene-specific regulatory role. Here we have used proteomics and genome-wide transcriptome and factor-binding analysis to demonstrate molecular mechanisms through which the binding of Bcl11b to specific genomic sites controlled target-gene expression during the commitment of T cells. We found that Bcl11b was able to repress target genes directly by nucleating complexes of co-repressors on the DNA at specific sites where they would not otherwise assemble. Such sites of Bcl11b-dependent recruitment or redirection of cofactors turned out to be a far better statistical discriminator of genes that Bcl11b actually controlled than was the binding of Bcl11b alone. At the same time, a substantial minority of the effects of Bcl11b were apparently indirect and were mediated through its repression of the regulatory factor–encoding genes Id2 and Zbtb16. Acute double deletion through the Cas9 system revealed that many effects of the deletion of Bcl11b were responses to the resultant increases in Id2 and/or PLZF. In particular, many genes expressed specifically by the T cell lineage, which apparently depend on Bcl11b for activation during commitment, in fact required Bcl11b largely for the suppression of Id2. Thus, in addition to its action on direct genomic targets, Bcl11b is an indirectly acting but critical member of the E protein–Id gene regulatory network in pro-T cells during commitment.

The best statistical enrichment criteria available for the identification of loci that Bcl11b regulates, positively or negatively, turned
Kueh, H. Y. et al. Asynchronous combinatorial action of four regulatory factors activates Bcl11b for T cell commitment. Nat. Immunol. 17, 956–965 (2016).

Ikawa, T. et al. An essential developmental checkpoint for production of the T cell lineage. Science 329, 93–96 (2010).

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).

Longabaugh, W. J. R. et al. Bcl11b and combinatorial resolution of cell fate in the T-cell gene regulatory network. Proc. Natl Acad. Sci. USA 114, 5800–5807 (2017).

Li, P. et al. Reprogramming of T cells to natural killer-like cells upon Bcl11b deletion. Science 329, 85–89 (2010).

Inoue, J. et al. Expression of TCRβ partly rescues developmental arrest and apoptosis of αβ T cells in Bcl11b-/- mice. J. Immunol. 176, 5871–5879 (2006).

Kastner, P. et al. Bcl11b represses a mature T-cell gene expression program in immature CD4+CD8+ thymocytes. Eur. J. Immunol. 40, 2143–2154 (2010).

Hirose, S. et al. Bcl11b prevents the intrathymic development of innate CD8+ T cells in a cell intrinsic manner. Int. Immunol. 27, 205–215 (2015).

Avram, D. & Califano, D. The multifaceted roles of Bcl11b in thymic and peripheral T cells: impact on immune diseases. J. Immunol. 193, 2059–2065 (2014).

Koopmann et al. Priming of lineage-specifying genes by Bcl11b is required for lineage choice in post-selection thymocytes. Nat. Commun. 8, 702 (2017).

Uddin, M. N. et al. Transcription factor Bcl11b sustains INK1T and INK2T cell programs, restricts INK1T cell development, and governs INK7 cell survival. Proc. Natl Acad. Sci. USA 113, 7608–7613 (2016).

Svárová, D., Fields, A., Senawong, T., Topark-Ngarm, A. & Leid, M. COUP-TF (chicken ovalbumin upstream promoter transcription factor)-interacting protein 1 (CTIP1) is a sequence-specific DNA binding protein. Biochem. J. 368, 555–563 (2002).

Gismasa, V. B. et al. Bcl11b functionally associates with the NuRD complex in T lymphocytes to repress targeted transcript. Oncogene 24, 6753–6764 (2005).

Senawong, T. et al. Involvement of the histone deacetylase SIRT1 in chicken ovalbumin upstream promoter transcription factor (COUP-TF)-interacting protein 2-mediated transcriptional repression. J. Biol. Chem. 278, 43041–43050 (2003).

Hu, G. et al. Transformation of accessible chromatin and 3D nucleosome underlies lineage commitment of early T cells. Immunity 48, 227–242 (2018).

Califano, D. et al. Transcription factor Bcl11b controls identity and function of mature Type 2 innate lymphoid cells. Immunity 43, 354–368 (2015).

Yu, Y. et al. Single-cell RNA-seq identifies a PD-1+ ILC progenitor and defines its development pathway. Nature 539, 102–106 (2016).

Miyazaki, M. et al. The E-Id protein axis specifies adaptive lymphoid cell identity and suppresses thymic innate lymphoid cell development. Immunity 46, 818–834 (2017).

Mowl, W. K. et al. Group 1 innate lymphoid cell lineage identity is determined by a cis-regulatory element marked by a long non-coding RNA. Immunity 47, 435–447 (2017).

Delconte, R. B. et al. The helix-loop-helix protein ID2 governs NK cell fate by tuning their sensitivity to Interleukin-15. Immunity 44, 103–115 (2016).

Sellet, C. et al. Deciphering the innate lymphoid cell transcriptional program. Cell Rep. 17, 436–447 (2016).

Rankin, L. & Belz, G. T. Diverse roles of inhibitor of differentiation 2 in adaptive immunity. Clin. Dev. Immunol. 2011, 281569 (2011).

Boos, M. D., Yokota, Y., Eberl, G. & Kee, B. L. Mature natural killer cell and lymphoid tissue-inducing cell development requires Id2-mediated suppression of E protein activity. J. Exp. Med. 204, 1119–1130 (2007).

Constantinides, M. G. et al. PLZF expression maps the early stages of ILC development. Proc. Natl Acad. Sci. USA 112, 5123–5128 (2015).

Constantinides, M. G., McDonald, B. D., Verhoef, P. A. & Bendelac, A. A committed precursor to innate lymphoid cells. Nature 508, 397–401 (2014).

Raberger, J. et al. The transcriptional regulator PLZF induces the development of CD44+ memory phenotype Tcells. Proc. Natl Acad. Sci. USA 105, 17919–17924 (2008).

Savage, A. K. et al. The transcription factor PLZF directs the effector program of the NKT cell lineage. Immunity 29, 391–403 (2008).

Kovalovsky, D. et al. The BTB-zinc finger transcriptional regulator PLZF controls the development of invariant natural killer T cell effector functions. Nat. Immunol. 9, 1055–1064 (2008).

Miller, J. C. et al. Deciphering the transcriptional network of the dendritic cell lineage. Nat. Immunol. 13, 888–899 (2012).

Satpathy, A. T., Wu, X., Albring, J. C. & Murphy, K. M. Re(de)fining the dendritic cell lineage. Nat. Immunol. 13, 1145–1154 (2012).

Doutlov, S. et al. PLZF is a regulator of homeostatic and cytokine-induced myeloid development. Genes Dev. 23, 2076–2087 (2009).

de Boer, J. et al. Transgenic mice with hematopoietic and lymphoid specific expression of Cre. Eur. J. Immunol. 33, 314–325 (2003).

Lee, P. F. et al. A critical role for Dnmt1 and DNA methylation in T cell development, function, and survival. Immunity 15, 763–774 (2001).
37. Liu, N. et al. Direct promoter repression by BCL11A controls the fetal to adult hemoglobin switch. *Cell* 173, 430–442.e17 (2018).
38. Dionne, C. J. et al. Subversion of T lineage commitment by PU.1 in a clonal cell line system. *Dev. Biol.* 280, 448–466 (2005).
39. Del Real, M. M. & Rothenberg, E. V. Architecture of the lymphoid developmental switch controlled by PU.1, Notch and Gata3. *Development* 140, 1207–1219 (2013).
40. Kadoch, C. & Crabtree, G. R. Mammalian SWI/SNF chromatin remodeling complexes and cancer: Mechanistic insights gained from human genomics. *Sci. Adv.* 1, e1500447 (2015).
41. Kang, H. et al. Bivalent complexes of PRC1 with orthologs of BRD4 and MOZ/MORF target developmental genes in *Drosophila*. *Genes Dev.* 31, 1988–2002 (2017).
42. Gao, Z. et al. An AUTS2-Polycomb complex activates gene expression in the CNS. *Nature* 516, 349–354 (2014).
43. Hosokawa, H. et al. Functionally distinct Gata3/Chd4 complexes coordinately establish T helper 2 (Th2) cell identity. *Proc. Natl Acad. Sci. USA* 110, 4691–4696 (2013).
44. Miccio, A. et al. NuRD mediates activating and repressive functions of GATA-1 and POG-1 during blood development. *EMBO J.* 29, 442–456 (2010).
45. Williams, C. J. et al. The chromatin remodeler Mi-2β is required for CD4 expression and T cell development. *Immunity* 20, 719–733 (2004).
46. Hosokawa, H. et al. Transcription factor PU.1 represses and activates gene expression in early T cells by redirecting partner transcription factor binding. *Immunity* 48, 1119–1134.e7 (2018).
47. Ishizuka, I. E. et al. Single-cell analysis defines the divergence between the innate lymphoid cell lineage and lymphoid tissue-inducer cell lineage. *Nat. Immunol.* 17, 269–276 (2016).
48. Di Santo, J. P. Staying innate: transcription factor maintenance of innate lymphoid cell identity. *Immunol. Rev.* 261, 169–176 (2014).
49. Xu, W. et al. E2A transcription factors limit expression of Gata3 to facilitate T lymphocyte lineage commitment. *Blood* 121, 1534–1542 (2013).
50. Miyazaki, M. et al. The opposing roles of the transcription factor E2A and its antagonist Id3 that orchestrate and enforce the naive fate of T cells. *Nat. Immunol.* 12, 992–1001 (2011).
51. Revilla-I-Domingo, R. et al. The B-cell identity factor Pax5 regulates distinct transcriptional programmes in early and late B lymphopoiesis. *EMBO J.* 31, 3130–3146 (2012).
52. McMamus, S. et al. The transcription factor Pax5 regulates its target genes by recruiting chromatin-modifying proteins in committed B cells. *EMBO J.* 30, 2388–2404 (2011).
53. Boller, S. & Grosschedl, R. The regulatory network of B-cell differentiation: a focused view of early B-cell factor 1 function. *Immunol. Rev.* 261, 102–115 (2014).
54. Treiber, T. et al. Early B cell factor 1 regulates B cell gene networks by activation, repression, and transcription– independent poising of chromatin. *Immunity* 32, 714–725 (2010).
55. Albu, D. I. et al. BCL11B is required for positive selection and survival of double-positive thymocytes. *J. Exp. Med.* 204, 3003–3015 (2007).
56. Mingueuena, M. et al. The transcriptional landscape of αβ T cell differentiation. *Nat. Immunol.* 14, 619–632 (2013).
57. Heng, T. S. P. & Painter, M. W., The Immunological Genome Project Consortium. The Immunological Genome Project: networks of gene expression in immune cells. *Nat. Immunol.* 9, 1091–1094 (2008).

Acknowledgements

We thank D. Perez, T. Tijerina, and R. Diamond for cell sorting and advice; I. Soto for mouse colony care; V. Kumar for library preparation and sequencing; H. Amrhein and D. Trout for computational assistance; I. Antoshechkin for sequencing management; X. Wang for related exploratory experiments; and members of the Rothenberg group for discussions and reagents. Supported by the Manpei Suzuki Diabetes Foundation (H.H.), the US Public Health Service (R01AI083514 and R01HD076915 to E.V.R.), Grants-in-Aid for Advanced Research and Development Programs for Medical Innovation (T.T.), the Takeda Science Foundation (T.T.), the SENSINE Medical Research Foundation (T.T.), the Swedish Research Council (J.U.), the California Institute of Regenerative Medicine Bridges to Stem Cell Research Program (Pasadena City College and Caltech; M.R.-W.), the L. A. Garfinkle Memorial Laboratory Fund and the AI Sherman Foundation, special project funds from the Provost and Division of Biology & Biological Engineering of Caltech, and the Albert Billings Ruddock Professorship (E.V.R.). This work was performed in part in the Collaborative Research Project Program of the Medical Institute of Bioregulation, Kyushu University.

Author contributions

H.H. designed the study, performed experiments, analyzed data and wrote the manuscript; M.R.-W. performed experiments, analyzed data and wrote the manuscript; M.A.Y., J.U., M.L.G.Q., M.M., K.I.N., T.T. performed experiments, analyzed data and provided discussions; and E.V.R. designed and supervised the study, analyzed data and wrote the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41590-018-0238-4.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to E.V.R.

© The Author(s), under exclusive licence to Springer Nature America, Inc. 2018
Methods

Mice. C57BL/6 (B6) mice, B6.Cg-Tg(BCL2)25Wehi/J (Bcl2-tg) mice and B6.Gt(Rosa26)J(Cg-cre-iCre)1Cwi N9 mice were purchased from the Jackson Laboratory. Vav1-iCre mice (B6.Cg-Comm10(CB.1)-lck(Lck)-proximal promoter-iCre mice developed by C. Wilson's group (B6.Cg-Tg(Lck-Cre)) were purchased from Taconic Laboratories. The Lck-Cre (B6.Cg-Tg(Lck-Cre)) mouse was maintained in our laboratory.

To study Bcl11b murine embryonic stem cells were infected with a retrovirus encoding a Vav1-iCre or Lck-Cre and single-cell suspensions were made. Lineage marker–positive (Lin+) cells were depleted by staining with biotinylated antibodies (all from eBioscience except as otherwise indicated) to CD90 (53-6.7, eBioscience), CD44 (15A11, BioLegend), B220 (RA3-6B2; eBioscience), c-Kit (2B8; eBioscience), CD25 (PC61.5; eBioscience) and a biotin-conjugated lineage cocktail (anti-CD8a, anti-MHC class II, anti-CD4, anti-CD8, anti-TCRβ and anti-TcRγδ) and were subjected to further analysis. For sorting, cells were stained with anti-CD45 and anti-CD25 and a biotin-conjugated lineage cocktail (antibodies to CD8a, CD11b, CD11c, Gr-1, Ter-119, NK1.1, CD19, TCRβ, TCRγδ) (all antibodies identified above), and were sorted for infected CD25+ cells (Lin−CD45−CD25+CFP+).

Two-step affinity purification of Bcl11b complexes from the DN3-like cell line Scid.ad2.c2 cells infected with Myc-Flag-tagged Bcl11b encoding retrovirus were solubilized with a protease inhibitor–containing immunoprecipitation buffer (50 mM Tris-HCl (pH 7.5), 150 mM NaCl, 10% glycerol, 0.1% Tween, 1 mM EDTA, 10 mM NaF, 1 mM DTT and a protease inhibitor cocktail (Roche Applied Science)), then were lysed on ice for 30 min with gentle shaking and sonicated on a Misonix S-4000 sonicator (Qsonica) for three cycles, amplitude 20 for 30 s, followed by 30 s of rest. We did not add benzamidine or ethylidene bromide to exclude DNA- or RNA-mediated interactions because we were interested in the functionally relevant complexes that Bcl11b forms as it is working on the DNA. The insoluble materials were removed by centrifugation, and immunoprecipitation with anti-Flag M2 agarose (Sigma–Aldrich) was performed overnight at 4 °C. Immunocomplexes were eluted from the agarose by 3xFlag peptide (Sigma–Aldrich), and the eluted Bcl11b complexes were subjected to denaturing sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE) and transferred to polyvinylidene fluoride (PVDF) membranes. Western blots were developed with 100 mM bicarbonate in acetonitrile, and the proteins were digested with trypsin. After 0.1% formic acid was added to the supernatant, the peptides were harvested with liquid chromatography–tandem mass spectrometry with an Advance Ultra (Bruker) and an Orbitrap Velos Pro Mass Spectrometer (Thermo Fisher Scientific). The resulting tandem mass spectrometry dataset was analyzed with the Mascot software program (Matrix Science).

Gene Ontology and KEGG pathway analysis. Gene Ontology (GO) and (KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway annotation) pathway annotation was performed using the DAVID analysis tool (https://david.ncifcrf.gov/).

Immunoprecipitation and immunoblot analysis. Protein extracts from Scid.ad2.c2 cells infected with Myc-Flag-tagged Bcl11b were subjected to denaturing sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE) and transferred to polyvinylidene fluoride (PVDF) membranes. Western blots were developed with 100 mM bicarbonate in acetonitrile, and the proteins were digested with trypsin. After 0.1% formic acid was added to the supernatant, the peptides were harvested with liquid chromatography–tandem mass spectrometry with an Advance Ultra (Bruker) and an Orbitrap Velos Pro Mass Spectrometer (Thermo Fisher Scientific). The resulting tandem mass spectrometry dataset was analyzed with the Mascot software program (Matrix Science).

Gene Ontology and KEGG pathway analysis. Gene Ontology (GO) and (KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway annotation) pathway annotation was performed using the DAVID analysis tool (https://david.ncifcrf.gov/).

Immunoprecipitation and immunoblot analysis. Protein extracts from Scid.ad2.c2 cells infected with Myc-Flag-tagged Bcl11b were subjected to denaturing sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE) and transferred to polyvinylidene fluoride (PVDF) membranes. Western blots were developed with 100 mM bicarbonate in acetonitrile, and the proteins were digested with trypsin. After 0.1% formic acid was added to the supernatant, the peptides were harvested with liquid chromatography–tandem mass spectrometry with an Advance Ultra (Bruker) and an Orbitrap Velos Pro Mass Spectrometer (Thermo Fisher Scientific). The resulting tandem mass spectrometry dataset was analyzed with the Mascot software program (Matrix Science).

Gene Ontology and KEGG pathway analysis. Gene Ontology (GO) and (KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway annotation) pathway annotation was performed using the DAVID analysis tool (https://david.ncifcrf.gov/).

Immunoprecipitation and immunoblot analysis. Protein extracts from Scid.ad2.c2 cells infected with Myc-Flag-tagged Bcl11b were subjected to denaturing sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE) and transferred to polyvinylidene fluoride (PVDF) membranes. Western blots were developed with 100 mM bicarbonate in acetonitrile, and the proteins were digested with trypsin. After 0.1% formic acid was added to the supernatant, the peptides were harvested with liquid chromatography–tandem mass spectrometry with an Advance Ultra (Bruker) and an Orbitrap Velos Pro Mass Spectrometer (Thermo Fisher Scientific). The resulting tandem mass spectrometry dataset was analyzed with the Mascot software program (Matrix Science).

Gene Ontology and KEGG pathway analysis. Gene Ontology (GO) and (KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway annotation) pathway annotation was performed using the DAVID analysis tool (https://david.ncifcrf.gov/).
ChIP and ChIP-seq. 1 × 10^7 BM-derived DN3 cells (or CD25− cells for Bcl11b–
knockout cells) were fixed with 1% formaldehyde in 10 mM MEM for 10 min
(H3K27Ac) or with 1 mg/ml DSg (Thermo Scientific) in PBS for 30 min at ~2 °C,
followed by an additional 10 min with addition of formaldehyde up to 1% (Bcl11b–
Chd4, Mta2, HDA2C, Rest, Ring1b, LSD1 and Runx1). The reaction was quenched
by the addition of 1/10 volume of 0.125 M glycine, and the cells were washed
with HBSS (Gibco). Pelleted nuclei were dissolved in lysis buffer (0.5% SDS,
by the addition of 1/10 volume of 0.125 M glycine, and the cells were washed
Chd4, Mta2, HDAC2, Rest, Ring1b, LSD1 and Runx1). The reaction was quenched
followed by an additional 10 min with addition of formaldehyde up to 1% (Bcl11b–
Chd4, Mta2, HDA2C, Rest, Ring1b, LSD1 and Runx1). The samples were
incubated overnight at 4 °C, then were washed and eluted for 6 h at 65 °C in ChIP
elution buffer (20 mM Tris-HCl, pH 7.5, 5 mM EDTA 50 mM NaCl, 1% SDS,
and 50 μg/ml proteinase K). Precipitated chromatin fragments were cleaned with
Zymo ChIP DNA Clean & Concentrator. ChIP-seq libraries were constructed using
NEBNext ChIP-Seq Library Preparation Kit (E6240, NEB) and were sequenced
on Illumina HiSeq2500 in single read mode with a read length of 50 nt. Analysis
pipelines used are described below (ChIP-seq analysis and
RNA-seq analysis).

mRNA-preparation and RNA-seq. Total RNA was isolated from 3 × 10^6 of
cultured cells and 2 × 10^5 to 5 × 10^5 of ex vivo cells using RNeasy Micro Kit
(Qiagen). Libraries were constructed using NEBNext Ultra RNA Library Prep
Kit for Illumina (New England Biolabs) from ~ 1 μg of total RNA following manufacturer's
instructions. Libraries were sequenced on Illumina HiSeq2500 in single read
mode with the read length of 50 nt. Base calls were performed with RTA 1.13.48.0
followed by conversion to FASTQ with bcl2fastq 1.8.4 and produced approximately
30 million reads per sample.

ChIP-seq analysis. Base calls were performed with RTA 1.13.48.0, followed by
conversion to FASTQ with bcl2fastq 1.8.4, and produced approximately
30 million reads per sample. ChIP-seq data were mapped to the mouse genome
build NCBI37/mm9 using Bowtie (v1.1.1; http://bowtie-bio.sourceforge.net/index.
html) with ‘–v 3 –k 11 –m 10 –t best –strata settings, and HOMER tags were created with makeTagDirectory and visualized in the UCSC-genome browser
(http://genome.ucsc.edu)67. The NCBI37/mm9 assembly was chosen for ChIP-seq
sample mapping in this study to ease comparisons with previous data tracks
from our lab and others. ChIP peaks were identified with ‘findPeaks.pl’
against a matched control sample using the settings ‘-P 1 –LFP 1 -poisson -1 -style
factor’. The identified peaks were annotated to genes with the ‘annotatePeaks.pl’
command against the mm9 genomic build in the HOMER package. Peak calls
were always based on data from at least two independent biological replicates.
Peak reproducibility was determined by a HOMER adaptation of the IDR
( Irreproducibility Discovery Rate) package according to ENCODE guidelines
(https://encodeproject.org/). Only reproducible high-
quality peaks with a normalized peak score of ≥15 were considered for further
analysis. motif-enrichment analysis was performed with the ‘findMotifsGenome.pl’
command in the HOMER package using a 200-bp window.

RNA-seq analysis. RNA-seq reads were mapped to the mouse genome build
NCBI37/mm9 with STAR (v2.4.0)70 and were post-processed with RSEM (v1.2.25;
http://deweylab.github.io/RSEM/) according to the settings in the ENCODE long-
RNA-seq-pipeline (https://github.com/ENCODE-DCC/long-rna-seq-pipeline/blob/master/DEG/STAR_RSEM.sh) with the minor modifications that the setting ‘–
output-genome-bam—sampling—for—bam’ was added to rsem—calculate—expression.
STAR and RSEM relevant RNA-seq libraries were created against NCBI37/mm9
with the Ensembl gene model file Mus_musculus.NCBIM37.66. gtf. The resulting bam files were used to create HOMER tag directories
(makeTagDirectory with ‘–keepAll setting). For analysis of statistical significance
among DEGs, the raw gene counts were derived from each tag directory with
‘analyzeRepeats.pl’ with the ‘–noad—condenseGenes’ options, followed by the getDiffExpression.pl command using EdgeR (v3.6.8; http://bioconductor.org/
packages/release/bioc/html/edgeR.html)70. For data visualization, RPMK normalized
reads were derived using the ‘analyzeRepeats.pl’ command with the options ‘–count
exons –condenseGenes –rpkni, followed by log transformation. The normalized
reads were hierarchically clustered with ‘average’ linkage and were visualized in Matlab (clustergram).

UCSC Genome Browser bigwig visualization. BigWigs were generated from
the aligned SAM or BED-file formats using Samtools82, Bedtools60 and the UCSC
genomeCoverageBed and bedGraphPlotBigWig and were normalized to 1 million
reads. For visualization of RNA-seq tracks, bamToBed and genomeCoverageBed
were used with the ‘split’ setting enabled. BigWig files were up-loaded to the
UCSC-genome browser (http://genome.ucsc.edu) for visualization.

Position-weight matrix (PWM) scanning of Bcl11b associated peak lists. New
sequence motifs that are specific Bcl11b-recognition targets, on the basis of in vitro
protein-binding microarray reactivity of purified full-length Bcl11b (‘XL’), the
middle two fingers of Bcl11b (Bcl11b-23) or the carboxy-terminal three zinc
fingers of Bcl11b (Bcl11b-456) have been reported36. To use HOMER to quantify
the representation of these Bcl11b motifs in Bcl11b or Bcl11b cofactor ChIP peaks,
23-residue enriched protein-binding microarray-defined probability matrices
(Bcl11b-23_rep2, Bcl11b-456_rep1, Bcl11b-456_rep2, Bcl11b-456_rep3, Bcl11b-456_rep4,
and Bcl11b-456_rep5) were downloaded from ref.15. From these 23-residue matrices, a HOMER-compatible 12-residue PWM was
thus derived by shifting from the seventh position to the eighteenth position in each of the
23-residue matrices, with the log-odds detection threshold manually set to 5, for relatively
inclusive scoring. To detect presence of these PWMs in the peak tracks, PWM
searches were carried out with ‘annotatePeaks.pl’ with the options ‘–m [motif files’.

Quantification and statistical analysis. Differentially expressed genes (DEGs)
were defined using EdgeR typically with FDR < 0.05, a change in expression
([log2] of > 1-fold, and RPKM > 1, except where indicated otherwise, on the basis
of measurements from at least two biologically independent replicates for
each sample type. The statistical significance of differences between datasets
was determined by two-sided Student's t-test, Pearson's correlation coefficient or
two-sided Fisher's exact test using Excel or the R package. Statistical details
of experiments can be found in the figure legends. The statistical methods
and methods for ensuring reproducibility are also reported in the Life Sciences
Reporting Summary for this paper.

Data availability

Additional data that support the findings of this study are available from the
corresponding author upon request. In addition to the complete description
and explanation of the methods presented here, reagent lists and some general
methods are also repeated, along with statistical checklists, in the Nature Research
Reporting Summary that accompanies this paper. The GEO accession codes for all
the deep-sequencing data reported in this paper are GSE110305, GSE110882 and
GSE11574.

References

58. Strasser, A., Harris, A. W. & Cory, S. bcl-2 transgene inhibits T cell death and
perturbs thymic self-censorship. Cell 67, 889–899 (1991).
59. Platt, R. J. et al. CRISPR-Cas9 knockin mice for genome editing and cancer
modeling. Cell 159, 440–455 (2014).
60. Srinivas, S. et al. Cre reporter strains produced by targeted insertion of EYFP
and ECFP into the ROSA26 locus. BMC Dev. Biol. 1, 4 (2001).
61. Scripture-Adams, D. D. et al. GATA-3 dose-dependent checkpoints in early
T cell commitment. J. Immunol. 193, 3470–3491 (2014).
62. Higashi, A. Y. et al. Direct hematological toxicity and illegitimate
cromosomal recombination caused by the systemic activation of CreERT2.
J. Immunol. 182, 5633–5640 (2009).
63. Champakar, A. et al. Regulation of early T-lineage gene expression and
developmental progression by the progenitor cell transcription factor PU.1.
Genes Dev. 29, 832–848 (2015).
64. Speir, M. L. et al. The UCSC Genome Browser database: 2016 update. Nucleic
Acids Res. 44, D717–D725 (2016).
65. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics
30, 15–21 (2013).
66. Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA-Seq
data with or without a reference genome. BMC Bioinformatics 12, 323 (2011).
67. Heinz, S. et al. Simple combinations of lineage-determining transcription
factors prime cis-regulatory elements required for macrophage and B cell
identities. Mol. Cell 38, 576–589 (2010).
68. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor
package for differential expression analysis of digital gene expression data.
Bioinformatics 26, 139–140 (2010).
69. Li, H. et al. The Sequence Alignment/Map format and SAM Tools.
Bioinformatics 25, 2078–2079 (2009).
70. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for
comparing genomic features. Bioinformatics 26, 841–842 (2010).
Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

- **n/a** Confirmed
- □ [ ] **The exact sample size** (n) for each experimental group/condition, given as a discrete number and unit of measurement
- □ [ ] An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- □ [ ] The statistical test(s) used AND whether they are one- or two-sided
  - *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
- □ [ ] A description of all covariates tested
- □ [ ] A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- □ [ ] A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- □ [ ] For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted. Give P values as exact values whenever suitable.
- □ [ ] For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- □ [ ] For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- □ [ ] Estimates of effect sizes (e.g. Cohen’s d, Pearson’s r), indicating how they were calculated
- □ [ ] Clearly defined error bars
  - *State explicitly what error bars represent (e.g. SD, SE, CI)*

*Our web collection on statistics for biologists may be useful.*

Software and code

Policy information about availability of computer code

**Data collection**
- Flow cytometry data:
  - Operating system software for Miltenyi Biotech MACSQuant 10 Flow Cytometer, BD FACS Aria II Cell Sorter, iCyt Mission Technology Reflection Cell Sorter.
  - Sequencing acquisition: operating system software for HiSeq2500; RTA 1.13.48.0; bcl2fastq 1.8.4

**Data analysis**
- bedGraphToBigWig
- Bedtools (v.2.17.0)
- Bioconductor (v3.4)
- Bowtie (v1.1.1)
- Cluster3 (v1.52)
- EdgeR (v.3.16.5)
- EaSeq
- FlowJo (v10.0.8)
- HOMER (v4.8)
- HOMER-IDR
- Limma (v.3.30.11)
- MatLab (R2016a)
- R (v3.3.2)
- RSEM (v1.2.25)
Rstudio (v1.0.136)  
Samtools (v0.1.19-96b5f2294a)  
STAR (v2.4.0)  
TreeView (v1.1.6r4)  
Mascot

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:
- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

GEO: GSE110305, GSE110882 and GSE115744.

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Life sciences

Study design

All studies must disclose on these points even when the disclosure is negative.

Sample size  
Sample size was determined empirically, at least two independent experiments were performed.

Data exclusions  
No data were excluded, except for two RNA-seq samples with very small library size and symptoms of amplification artifacts.

Replication  
The experimental findings were reliably reproduced.

Randomization  
No randomization in this study.

Blinding  
No blinding test in this study.

Materials & experimental systems

Policy information about availability of materials

n/a Involved in the study

Antibodies  
Anti-human/mouse CD44 PE, ebioscience, Cat#12-0441-83, Clone IM7, Lot#4312996, 1:300
Anti-mouse CD117 (cKit) APC, ebioscience, Cat#17-1171-82, Clone 2B8, Lot#4299769, 1:100
Anti-mouse CD25 APCe780, ebioscience, Cat#47-0251-82, Clone PC61.5, Lot#1942453, 1:200
Anti-mouse CD45 PECy7, ebioscience, Cat#25-0451-82, Clone PK136, Lot#4329704, 1:600
Anti-mouse Ter119 Biotin, ebioscience, Cat#13-0452-85, Clone RA3-6B2, Lot#4273327, 1:200

Anti-mouse B220 Biotin, Biolegend, Cat#108710, Clone PK136, Lot#B191787, 1:100

Anti-mouse CD19 Biotin, ebioscience, Cat#13-0193-85, Clone eBio1D3, Lot#4300688, 1:300

Antibodies used

Antibodies used  
Anti-human/mouse CD44 PE, ebioscience, Cat#12-0441-83, Clone IM7, Lot#4312996, 1:300
Anti-mouse CD117 (cKit) APC, ebioscience, Cat#17-1171-82, Clone 2B8, Lot#4299769, 1:100
Anti-mouse CD25 APCe780, ebioscience, Cat#47-0251-82, Clone PC61.5, Lot#1942453, 1:200
Anti-mouse CD45 PECy7, ebioscience, Cat#25-0451-82, Clone 3O-F11, Lot#4329704, 1:600
Anti-mouse NK1.1 Biotin, Biolegend, Cat#108704, Clone PK136, Lot#B234365, 1:300
Anti-mouse NK1.1 APC, Biolegend, Cat#108710, Clone PK136, Lot#B191787, 1:100

Anti-mouse B220 Biotin, Biolegend, Cat#108710, Clone 2B8, Lot#4299769, 1:100

Anti-mouse CD19 Biotin, ebioscience, Cat#13-0193-85, Clone eBio1D3, Lot#4300688, 1:300

Antibodies

Antibodies

Antibodies used

Antibodies used  
Anti-human/mouse CD44 PE, ebioscience, Cat#12-0441-83, Clone IM7, Lot#4312996, 1:300
Anti-mouse CD117 (cKit) APC, ebioscience, Cat#17-1171-82, Clone 2B8, Lot#4299769, 1:100
Anti-mouse CD25 APCe780, ebioscience, Cat#47-0251-82, Clone PC61.5, Lot#1942453, 1:200
Anti-mouse CD45 PECy7, ebioscience, Cat#25-0451-82, Clone 3O-F11, Lot#4329704, 1:600
Anti-mouse NK1.1 Biotin, Biolegend, Cat#108704, Clone PK136, Lot#B234365, 1:300
Anti-mouse NK1.1 APC, Biolegend, Cat#108710, Clone PK136, Lot#B191787, 1:100

Anti-mouse B220 Biotin, Biolegend, Cat#108710, Clone 2B8, Lot#4299769, 1:100

Anti-mouse CD19 Biotin, ebioscience, Cat#13-0193-85, Clone eBio1D3, Lot#4300688, 1:300

Anti-mouse Ter119 Biotin, ebioscience, Cat#13-5921-85, Clone Ter-119, Lot#4300555, 1:300
Validation

Antibodies were chosen based on the validation statements for species (mouse) and application (IB, ChIP or FACS) on the manufacturer's website.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s) HEK293T (obtained from ATCC), Scid.adh.2c2 (previously established in our lab, Dionne et al., 2005, Devel Biol), OP9-DL1 & OP9-Mig (created and sent to us by Schmitt & Zuniga-Pflucker, 2002, Immunity). Scid.adh.2c2 have been used by our lab subsequently in Del Real and Rothenberg 2013 Development, Scripture-Adams et al 2014 J Immunol, and Champhekar et al 2015 Genes Dev. Cocultures with OP9-DL1 and OP9-Mig were also used by us in Taghon et al 2005 Genes Dev, Franco et al 2006 PNAS, Taghon et al 2006 Immunity, Taghon et al 2007 Nat Immunol, Li et al Science 2010, Yui et al 2010 J Immunol, Del Real and Rothenberg 2013 Development, Scripture-Adams et al 2014 J Immunol, Champhekar et al 2015 Genes Dev, Kueh et al 2016 Nat Immunol, and Longabaugh et al 2017 PNAS).

Authentication Functionally in repeated tests; by cell surface phenotype; and in cases of Scid.adh.2c2, OP9-DL1, & OP9-Mig, by RNA-seq.

Mycoplasma contamination All cell lines were negative for mycoplasma contamination.

Commonly misidentified lines (See ELAC register) No commonly misidentified cell lines were used.

Research animals

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Animals/animal-derived materials C57BL/6 (referred to as B6), B6.Cg-Tg(BCL2)25Wehi/J (Bcl2-tg) and B6.Gt(ROSA)26Sortm1.1(CAG-cas9*,-EGFP)Fezh/J (Cas9) mice were purchased from the Jackson Laboratory. Vav1-iCre mice (B6N.Cg-Cmmd10Tg(Vav1-icre)A2Kio/J) were purchased from Jackson Laboratories and pLck-Cre mice developed by Christopher Wilson's group (B6.Cg-Tg(Lck-cre)1Cwi) were purchased from Taconic Laboratories. The Cre activity reporter allele ROSA26R-eYFP was also used. Except for Vav1-iCre, which was maintained in heterozygotes, the indicated transgenes were bred to homozygosity alone or in combinations on the B6 background. Bcl11bfl/fl-Rosa26-Cre-ERT2 mice were derived from stock kindly provided by Pentao Liu (Cambridge, UK), and maintained as a separate line. All animals were bred and maintained in the California Institute of Technology Laboratory Animal Facility, under specific pathogen free conditions, and the protocol supporting animal breeding for this work was reviewed and approved by the Institute Animal Care and Use Committee of the California Institute of Technology.

Method-specific reporting

n/a Involved in the study

-  ChIP-seq
-  Flow cytometry
-  Magnetic resonance imaging
Data deposition

- Confirm that both raw and final processed data have been deposited in a public database such as GEO.
- Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links

May remain private before publication.

To review GEO accession GSE110305:
Go to https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE110305
Enter token czaonuxpctpmn into the box

To review GEO accession GSE110882:
Go to https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE110882
Enter token gzzazmmocdtsbnwl into the box

To review GEO accession GSE115744:
Go to https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE115744
Enter token mfrhplmwrqtsvf into the box

Files in database submission

ChIP-seq data

WT_DN3_aChd4 HOMER IDR peaks.csv
Bcl11bKO_aChd4 HOMER IDR peaks.csv
WT_DN3_aMta2 HOMER IDR peaks.csv
Bcl11bKO_aMta2 HOMER IDR peaks.csv
WT_DN3_aHdac2 HOMER IDR peaks.csv
Bcl11bKO_aHdac2 HOMER IDR peaks.csv
WT_DN3_aRest HOMER IDR peaks.csv
Bcl11bKO_aRest HOMER IDR peaks.csv
WT_DN3_aRing1b HOMER IDR peaks.csv
Bcl11bKO_aRing1b HOMER IDR peaks.csv
WT_DN3_aLD1s1 HOMER IDR peaks.csv
Bcl11bKO_aLD1s1 HOMER IDR peaks.csv
WT_DN3_aRunx1 HOMER IDR peaks.csv
Bcl11bKO_aRunx1 HOMER IDR peaks.csv
WT_DN3_aBcl11b HOMER IDR peaks.csv
Bcl11bKO_aBcl11b HOMER IDR peaks.csv
Scid.adh.2c2_shControl_aChd4 HOMER peaks.csv
Scid.adh.2c2_shBcl11b_aChd4 HOMER peaks.csv
Scid.adh.2c2_shControl_aHdac2 HOMER peaks.csv
Scid.adh.2c2_shBcl11b_aHdac2 HOMER peaks.csv
Scid.adh.2c2_shControl_aMta2 HOMER peaks.csv
Scid.adh.2c2_shBcl11b_aMta2 HOMER peaks.csv
Scid.adh.2c2_shControl_aRest HOMER peaks.csv
Scid.adh.2c2_shBcl11b_aRest HOMER peaks.csv
Scid.adh.2c2_shControl_aChd4-ChIP-rep1.fastq.gz
Scid.adh.2c2_shBcl11b_aChd4-ChIP-rep1.fastq.gz
WT_DN3_aChd4-ChIP-rep1.fastq.gz
WT_DN3_aChd4-ChIP-rep2.fastq.gz
WT_DN3_aChd4-ChIP-rep3.fastq.gz
Bcl11bKO_aChd4-ChIP-rep1.fastq.gz
Bcl11bKO_aChd4-ChIP-rep2.fastq.gz
Bcl11bKO_aChd4-ChIP-rep3.fastq.gz
WT_DN3_aMta2-ChIP-rep1.fastq.gz
WT_DN3_aMta2-ChIP-rep2.fastq.gz
Bcl11bKO_aMta2-ChIP-rep1.fastq.gz
Bcl11bKO_aMta2-ChIP-rep2.fastq.gz
WT_DN3_aHdac2-ChIP-rep1.fastq.gz
WT_DN3_aHdac2-ChIP-rep2.fastq.gz
WT_DN3_aHdac2-ChIP-rep3.fastq.gz
Bcl11bKO_aHdac2-ChIP-rep1.fastq.gz
Bcl11bKO_aHdac2-ChIP-rep2.fastq.gz
Bcl11bKO_aHdac2-ChIP-rep3.fastq.gz
WT_DN3_aRest-ChIP-rep1.fastq.gz
WT_DN3_aRest-ChIP-rep2.fastq.gz
WT_DN3_aRest-ChIP-rep3.fastq.gz
Bcl11bKO_aRest-ChIP-rep1.fastq.gz
Bcl11bKO_aRest-ChIP-rep2.fastq.gz
Bcl11bKO_aRest-ChIP-rep3.fastq.gz
WT_DN3_aRunx1-ChIP-rep1.fastq.gz
WT_DN3_aRunx1-ChIP-rep2.fastq.gz
Bcl11bKO_aRunx1-ChIP-rep1.fastq.gz
Bcl11bKO_aRunx1-ChIP-rep2.fastq.gz
Bcl11bKO_aRunx1-ChIP-rep3.fastq.gz
WT_DN3_aBcl11b-ChIP-rep1.fastq.gz
WT_DN3_aBcl11b-ChIP-rep2.fastq.gz
Bcl11bKO_aBcl11b-ChIP-rep1.fastq.gz
Bcl11bKO_aBcl11b-ChIP-rep2.fastq.gz
WT_DN3_1%_input-rep1.fastq.gz
WT_DN3_1%_input-rep2.fastq.gz
Bcl11bKO_1%_input-rep1.fastq.gz
Bcl11bKO_1%_input-rep2.fastq.gz
WT_DN3_aH3K27Ac-ChIP-rep1.fastq.gz
WT_DN3_aH3K27Ac-ChIP-rep2.fastq.gz
Bcl11bKO_aH3K27Ac-ChIP-rep1.fastq.gz
Bcl11bKO_aH3K27Ac-ChIP-rep2.fastq.gz
WT_DN3_1%_input-rep1.fastq.gz
WT_DN3_1%_input-rep2.fastq.gz
Bcl11bKO_1%_input-rep1.fastq.gz
Bcl11bKO_1%_input-rep2.fastq.gz
Scid.adh.2c2_shControl_aChd4.fastq.gz
Scid.adh.2c2_shControl_aHdac2.fastq.gz
Scid.adh.2c2_shControl_aMta2.fastq.gz
Scid.adh.2c2_shControl_aRest.fastq.gz
Scid.adh.2c2_shControl_1%_input.fastq.gz
Scid.adh.2c2_shControl_1%_input.fastq.gz
RNA-seq data
Bcl11b_cofactorsKO_RNA-seq_RPKM_table.txt
DN3_sgBcl11b_sgid2_sgZbtb16_RNA-seq_RPKM_table.txt
Lck_Bcl11b_KO_RNA-seq_RPKM_Table
sgControl_RNA-rep1.fastq.gz
sgControl_RNA-rep2.fastq.gz
sgBcl11b_RNA-rep1.fastq.gz
sgBcl11b_RNA-rep2.fastq.gz
sgChd4_RNA-rep1.fastq.gz
sgChd4_RNA-rep2.fastq.gz
sgMta1_2_RNA-rep1.fastq.gz
sgMta1_2_RNA-rep2.fastq.gz
sgRest_RNA-rep1.fastq.gz
sgRest_RNA-rep2.fastq.gz
sgRing1a_b_RNA-rep1.fastq.gz
sgRing1a_b_RNA-rep2.fastq.gz
sgLSD1_RNA-rep1.fastq.gz
sgLSD1_RNA-rep2.fastq.gz
sgRunx1_RNA-rep1.fastq.gz
sgRunx1_RNA-rep2.fastq.gz
sgControl_RNA-rep3.fastq.gz
sgControl_RNA-rep4.fastq.gz
sgBcl11b_RNA-rep3.fastq.gz
sgBcl11b_RNA-rep4.fastq.gz
sgId2_RNA-rep1.fastq.gz
sgId2_RNA-rep2.fastq.gz
sgZbtb16_RNA-rep1.fastq.gz
sgZbtb16_RNA-rep2.fastq.gz
sgBcl11b_Id2_RNA-rep1.fastq.gz
sgBcl11b_Id2_RNA-rep2.fastq.gz
sgBcl11b_Zbtb16_RNA-rep1.fastq.gz
sgBcl11b_Zbtb16_RNA-rep2.fastq.gz
Lck_Bcl11b_WT_DN2_RNA-rep1.fastq.gz
Lck_Bcl11b_WT_DN2_RNA-rep2.fastq.gz
Lck_Bcl11b_WT_DN2_RNA-rep3.fastq.gz
Lck_Bcl11b_WT_DN2_RNA-rep4.fastq.gz
Lck_Bcl11b_WT_DN3_RNA-rep1.fastq.gz
Lck_Bcl11b_WT_DN3_RNA-rep2.fastq.gz
Lck_Bcl11b_WT_DN3_RNA-rep3.fastq.gz
Lck_Bcl11b_WT_DN3_RNA-rep4.fastq.gz
Lck_Bcl11b_WT_DN3_RNA-rep5.fastq.gz
Lck_Bcl11b_WT_DN2_RNA-rep1.fastq.gz
Lck_Bcl11b_WT_DN2_RNA-rep2.fastq.gz
Lck_Bcl11b_WT_DN2_RNA-rep3.fastq.gz
Lck_Bcl11b_WT_DN2_RNA-rep4.fastq.gz
Lck_Bcl11b_WT_DN2_RNA-rep5.fastq.gz
Lck_Bcl11b_HET_DN3_RNA-rep1.fastq.gz
Lck_Bcl11b_HET_DN3_RNA-rep2.fastq.gz
ChIP-seq data were mapped to the mouse genome build NCBI37/mm9 using Bowtie (v1.1.1; http://bowtie-bio.sourceforge.net/index.shtml) with “-v 3 -k 11 -m 10 -t --best –strata” settings and HOMER tagdirectories were created with makeTagDirectory and visualized in the UCSC-genome browser (http://genome.ucsc.edu).

**Methodology**

**Replicates**
Data are based on reproducible ChIP-seq peaks in two replicate samples

**Sequencing depth**
ChIP-seq libraries were sequenced on Illumina HiSeq2500 in single read mode with the read length of 50 nt. Base calls were performed with RTA 1.13.48.0 followed by conversion to FASTQ with bcl2fastq 1.8.4 and produced approximately 30 million reads per sample.

**Antibodies**
- Anti-Chd4 Bethyl Cat#A301-081A
- Anti-Mta2 Santa Cruz Cat#sc-9447
- Anti-Hdac2 Abcam Cat#ab12169
- Anti-Rest Caltech Protein Expression Center Cat#12C11-1B11
- Anti-Ring1b Bethyl Cat#A302-869A
- Anti-LSD1 Abcam Cat#ab17721
- Anti-Runx1 Abcam Cat#ab23980
- Anti-Bcl11b Abcam Cat#ab17721
- Anti-Bcl11b Bethyl Cat#A300-385A
- Anti-Bcl11b CST Cat#121120
- Anti-H3K27Ac Abcam Cat#ab4729

**Peak calling parameters**
ChIP peaks were identified with findPeaks.pl against a matched control sample using the settings “-P .1 -LP .1 -poisson .1 -style factor”. The identified peaks were annotated to genes with the annotatePeaks.pl command against the mm9 genomic build in the HOMER package.

**Data quality**
Peak reproducibility was determined by a HOMER adaptation of the IDR (Irreproducibility Discovery Rate) package according to ENCODE guidelines (https://sites.google.com/site/anshulkundaje/projects/idr). Only reproducible high quality peaks, with a normalized peak score ≥ 15, were considered for further analysis.

**Software**
- bedGraphToBigWig
- Bedtools (v.2.17.0)
- Bowtie (v1.1.1)
- EdgeR (v.3.16.5)
- HOMER (v4.8)
- HOMER-IDR

**Flow Cytometry**

**Plots**
- Confirm that:
  - The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
  - The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a ‘group’ is an analysis of identical markers).
  - All plots are contour plots with outliers or pseudocolor plots.
  - A numerical value for number of cells or percentage (with statistics) is provided.
Methodology

Sample preparation

Thymuses from 4-6wk old Bcl11b+/+, Bcl11bfl/+ and Bcl11bfl/fl ROSA26R-YFP mice with Vav1-iCre or Lck-Cre were removed, and single-cell suspensions were made. Lineage-positive cells were depleted by staining with biotinylated antibodies to CD8α (53-6.7), TCRγδ (GL3), TCRβ (HS7597), Ter119 (Ter119), NK1.1 (PK136), Dx5, and CD11c (N418), CD11b (M1/70), after which the cells were incubated with streptavidin-coated magnetic beads and then passed through an LS magnetic column in accordance with the manufacturer’s instructions (Miltenyi Biotec).

For in vitro differentiation of pro-T cells, bone marrow hematopoietic progenitors were used for input. Bone marrow (BM) was removed from the femurs and tibiae of 2-3 month-old mice. Suspensions of BM cells were prepared and stained for lineage markers using biotin-conjugated lineage antibodies (CD11b, CD11c, Gr1, TER-119, NK1.1, CD19, CD3ε, B220), then incubated with streptavidin-coated magnetic beads (Miltenyi Biotec), and passed through a magnetic column (Miltenyi Biotec). Then, Lin-Sca1+Kit+ (LSK) cells were sorted on a FACSAria (BD Bioscience). LSK cells were cultured on OP9-DL1 monolayers using OP9 medium (α-MEM, 20% FBS, 50 μM β-mercaptoethanol, Pen-Step-Glutamine) supplemented with 10 ng/ml of IL-7 (Pepro Tech Inc) and 10 ng/ml of Flt3L (Pepro Tech Inc). On day 7, cultured cells were disaggregated, filtered through 40-μm nylon mesh, and re-cultured on new OP9-DL1 monolayers with medium containing 5 ng/ml of IL-7 and 5 ng/ml of Flt3L. In cultures that were continued for longer times, cells were passaged onto fresh OP9-DL1 monolayers at day 10 and maintained up to day 14 in 1 ng/ml each of IL-7 and Flt3L.

Instrument

Miltenyi Biotech MACSQuant 10 Flow Cytometer
BD FACS Aria II Cell Sorter
iCyt Mission Technology Reflection Cell Sorter

Software

FlowJo (v10.0.8)

Cell population abundance

The abundance of the post-sort fractions were higher than 98%.

Gating strategy

Doublets were excluded using forward light-scatter gating followed by gating on lymphocytes based on FSC/SSC. Dead cells were excluded by gating on 7AAD negative cells. These cells were further gated as indicated in Supplementary Figures.

☐ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.