Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

HHS Public Access

Author manuscript

Nat Immunol. Author manuscript; available in PMC 2019 April 30.

Published in final edited form as:

Nat Immunol. 2018 December ; 19(12): 1427–1440. doi:10.1038/s41590-018-0238-4.

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

Hiroyuki Hosokawa[*,1], **Maile Romero-Wolf**[*,1], **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. Rothenberg**[†,1]

[1]Division of Biology & Biological Engineering, California Institute of Technology, Pasadena, CA, USA;

[2]Division of Molecular Hematology, Lund University; Sweden

[3]Department of Molecular and Cellular Biology, Medical Institute of Bioregulation, Kyushu University, 3-1-1 Maidashi, Higashi-ku, Fukuoka, 812-8582, Japan

[4]Department of Molecular Diagnosis, Chiba University, 1-8-1 Inohana, Chuo-ku, Chiba 260-8670, Japan

[5]AMED-CREST, Graduate School of Medicine, Chiba University, 1-8-1 Inohana, Chuo-ku, Chiba 260-8670, Japan

SUMMARY

Multipotent progenitors confirm their T cell-lineage identity in the DN2 pro-T cell 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. Proteomic analysis revealed that Bcl11b associates with multiple cofactors, and that its direct action was needed to recruit these cofactors to selective target sites. These sites of Bcl11b-dependent cofactor recruitment were enriched near functionally regulated target genes, and deletion of individual cofactors relieved repression of many Bcl11b-repressed genes. Runx1 collaborated with Bcl11b most frequently for both activation and repression. In parallel, Bcl11b indirectly regulated a subset

Users may view, print, copy, and download text and data-mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use:http://www.nature.com/authors/editorial_policies/license.html#terms

[*]Co-first authors, H.H. and M.R.-W. [†] Corresponding author and lead contact, evroth@its.caltech.edu.

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 helpful discussions; E.V.R. designed and supervised the study, analyzed data and wrote the manuscript.

COMPETING FINANCIAL INTEREST STATEMENT: The authors declare no conflict of interest or competing financial interest affecting this work.

DATA AVAILABILITY

All sequencing data are deposited in GEO: GSE110305, GSE110882, and GSE115744. The additional data that support the findings of this study are available from the corresponding author upon request.

DATA AND SOFTWARE AVAILABILITY

The accession numbers for all the deep-sequencing data reported in this paper are GEO: GSE110305, GSE110882, and GSE115744. The additional data that support the findings of this study are available from the corresponding author upon request.

STATEMENT OF ETHICAL COMPLIANCE

All work in this paper has been performed in compliance with all relevant ethical regulations.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

of target genes by a gene network circuit via *Id2* and *Zbtb16* (encoding PLZF), which were directly repressed by Bcl11b and controlled distinct alternative programs. Thus, this study defines the molecular basis of direct and indirect Bcl11b actions that promote T cell identity and block alternative potentials.

Keywords

Bcl11b; repressor complex; Id2; PLZF; cell fate decision; Runx1

INTRODUCTION

The zinc finger transcription factor Bcl11b is required for development of αβ T cells and most γδ T cells[1–3]. Its expression initiates precisely during T cell-lineage commitment[4], i.e. before T cell antigen receptor expression and between the DN2a ($CD25^+CD44^+$c-Kit^{hi+}) and DN2b ($CD25^+CD44^+$c-Kit^+) pro-T cell stages, and progression through T cell-lineage commitment is blocked or highly abnormal in cells lacking Bcl11b[5–7]. Bcl11b-deficient pro-T cells are less sensitive to Notch signaling and more prone to differentiate into natural killer (NK) cells than wild-type counterparts[6,8]. They also fail to go through β-selection, due to defects in T cell gene expression as well as abnormal persistence of immature features including c-Kit expression[5–9]. Deletion of the *Bcl11b* gene after β-selection causes abnormal activation of effector genes[10,11] and multiple functional defects in later thymocytes and mature T cells[12–14]. While the importance of Bcl11b for T cell development is clear, its exact mechanism of action is not. Bcl11b can bind to GC-rich sequences in DNA[15] and recruit chromatin-modifying NuRD and SIRT1 complexes[16,17], but in pro-T cells it primarily binds Ets and Runx motif-enriched sites in open chromatin[7,18]. Previous work has implicated Bcl11b in both activation and repression[5,6,8,10,12,19,20], with the most consistent effects across development on a core of genes that apparently require repression by Bcl11b in T cells[7,11]. Finally, Bcl11b effects have a striking overlap with effects of the basic helix-loop-helix protein E2A in early T cells[7], yet the basis for this convergence is not known.

This report addresses three questions about Bcl11b roles in establishing T cell commitment. First, what are the directly regulated target genes of 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 based on 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 blocks expression of E-protein antagonist 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 cells[21–26], while PLZF is crucial for innate-type T cells and for non-cytolytic ILCs[24,27–31], and both govern subsets of myeloid and dendritic cells[32–34]. We show that a gene network relating *Id2* and *Zbtb16* to Bcl11b function sheds light on the split between the T and innate immune cell families of developmental programs.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Results

Bcl11b impacts on gene expression in DN2/3 stage thymocytes

We previously showed that Bcl11b regulates a distinctive set of genes during initial T cell-lineage commitment of fetal-liver-derived precursors differentiating *in vitro*, including many targets that appear to be unique to this developmental period[7]. To examine its commitment role *in vivo*, we compared DN pro-T cells in mice where *Bcl11b*$^{fl/fl}$ was conditionally deleted with *Vav1*-iCre[35], expressed in all hematopoietic cells, against those in mice where *Bcl11b* was deleted with *Lck*-Cre (*Lck* proximal promoter), from an early-expressed transgene[36] first activated in DN2 pro-T cells (Fig. 1a, Supplementary Fig. 1a). The mice also contained a Cre-dependent ROSA26R-YFP reporter, which distinguished cells with deleted alleles from normal DN2a cells. In animals with wild-type (WT) *Bcl11b*, *Vav1*-iCre caused all DN thymocytes to express YFP (Fig. 1a, top), whereas *Lck*-Cre activated YFP only in DN2b and later cells (Fig. 1a, below; Supplementary Fig. 1a). Thus, *Lck*-Cre only deleted genes after *Bcl11b* would normally be turned on[4]. Homozygous *Bcl11b*$^{fl/fl}$ mice bred with either of these Cre transgenes showed similar-appearing arrests of T-cell precursors with a c-Kit^{hi+}CD25^{+} phenotype resembling normal DN2a cells (Fig. 1a). In the *Lck*-Cre^{+} *Bcl11b*$^{fl/fl}$ mice, however, the c-Kit^{hi+} DN2a-like cells comprised two populations, a YFP-negative, CD44^{+} one enriched for true DN2a cells, and a much larger YFP^{+}CD44^{lo} one generated only upon *Bcl11b* deletion (Supplementary Fig. 1a,b). Thus, *Bcl11b* excision could generate the YFP^{+} c-Kit^{hi+}CD25^{+} phenotype by retrograde-like differentiation from cells that had previously reached DN2b stage after activating *Bcl11b* initially.

DN thymocyte gene expression patterns in mice with either *Vav1*-iCre or *Lck*-Cre showed that about 300 genes were reproducibly upregulated in YFP^{+} homozygous *Bcl11b* knockout DN2-like thymocytes as compared to YFP^{+} control WT or heterozygous *Bcl11b* DN2 and DN3 thymocytes (Fig. 1b), defining Bcl11b-repressed genes. About 220 genes were significantly downregulated in these *Bcl11b* knockouts (Fig. 1c), defining Bcl11b-dependent genes. These criteria [false discovery rate (FDR) <0.05, |$\log_2$Fold Change (FC)|>1, average reads per kilobase million (RPKM)>1; Supplementary Tables 1,2] defined Bcl11b-regulated "differentially expressed genes" (DEGs) in young adult thymocytes. Although the *Bcl11b* knockout cells resembled normal DN2a thymocytes, their gene expression patterns sharply distinguished the mutant cells from any normal *Bcl11b*$^{+/+}$ subsets. Highly robust effects were seen in multiple samples, e.g. for *Bcl11b* itself, *Zbtb16* (encoding PLZF) and *Id2* (both Bcl11b-repressed), and *Cd6* (Bcl11b-dependent) genes (Supplementary Fig. 2). Interestingly, certain differentially expressed genes also showed partial de-repression in YFP^{+} *Bcl11b* heterozygous cells (Fig. 1b; Supplementary Table 1).

At a subset of these loci, the effects of *Bcl11b* deletion prior to commitment (*Vav1*-iCre) were more severe than the effects of deletion after commitment (*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 *Vav1*-iCre;*Bcl11b* deleted cells, but were expressed somewhat in *Lck*-Cre;*Bcl11b* deleted cells (Fig. 1d), while ~85 Bcl11b repression targets were more overexpressed in the *Vav1*-iCre knockout cells than in the *Lck*-Cre knockout cells (Fig. 1e). This difference suggests that even transient Bcl11b

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

expression in the *Lck*-Cre deletion model provides some needed function for T-cell development. The most timing-sensitive genes included *Cd3g, Cd3d, Cd3e*, and *Dntt* among Bcl11b-dependent genes, and progenitor-associated, γδ-associated, and alternative-lineage associated genes *Pou2af1, Tyrobp, Cd7, Itgae, Itgb7, Klf2, Trpm1, Cd163l1 (Scart1), Cited4*, and *Tnni1* among Bcl11b-repressed ones. However, most DEGs required Bcl11b both during and continuously after commitment, for activation or especially for repression in pro-T cells.

For mechanistic experiments on Bcl11b function, we also defined the genes affected by acute Cas9-mediated disruption of *Bcl11b* in pro-T cells differentiating *in vitro,* as shown in Supplementary Fig. 3a (Supplementary Table 2). Bone marrow precursors from B6.*ROSA26-Cas9; Bcl2-tg* transgenic mice (Cas9-Bcl2 cells), were cultured with OP9-DL1 stroma for 7 days, then transduced with retroviral vectors encoding sgRNA, and 7 d later analyzed and harvested for RNA-seq analysis. Most of these genes also overlapped with the significant DEGs from both *Lck*-Cre and *Vav1*-iCre deletion *in vivo* (Supplementary Fig. 3b; Supplementary Table 3). DEGs significant in all three are listed in Table 1, and include most of the targets previously reported in fetal liver-derived pro-T cells[7] (Supplementary Fig. 3c; see Methods). Importantly, genes encoding transcription factors Id2 and PLZF (*Zbtb16*) were highly significant Bcl11b repression targets in DN2/DN3 cells in every case. Thus, during pro-T cell commitment, Bcl11b activates and represses several hundred genes important for T cell identity.

A challenge of functional specificity

To determine which genes were directly regulated by Bcl11b, we performed Bcl11b ChIP-seq. However, despite the prevalence of Bcl11b repressive effects, binding of Bcl11b across the genome appeared preponderantly at open chromatin sites with active marks[7,18], globally associated with active genes, and showed no distinction between non-regulated loci and those with any response to *Bcl11b* deletion (Supplementary Fig. 4a; Supplementary Table 2). We tested whether distinct motifs were bound at functional sites. Most sites bound by Bcl11b in DN3 pro-T cells are enriched for Ets and Runx motifs[7,18]; we compared these with GC-rich sequences originally reported as a cognate site[15] and with three newly reported motifs for Bcl11b binding to protein-binding microarrays[37]. While matches to the newly defined sites (log odds ≥5) were found at 10–18% of Bcl11b occupancy sites, they were not enriched at DEGs (Supplementary Table 4). Bcl11b binding sites in CpG islands were actually strongly depleted near DEGs (Supplementary Fig. 4b). Bcl11b occupancy and target site motifs were thus unable to distinguish sites of Bcl11b function.

Bcl11b interacts with several “repressor” complexes

We reasoned that the sites where 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, cells of a DN3-like cell line[38,39], Scid.adh.2c2 (Fig. 2a), were transduced with Myc-Flag-tagged Bcl11b, and the Bcl11b-containing protein complexes were subjected into two-step affinity purification followed by SDS/PAGE and silver staining (Fig. 2b). Liquid chromatography-tandem mass spectrometry (LC-MS/MS) analysis identified more than 300 molecules with

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

supra-threshold enrichment (Supplementary Table 5). The bifunctional transcription factor Runx1 and proteins annotated as involved in “negative regulation of gene expression”, “transcriptional regulation”, and “chromatin remodeling” were highly enriched (Fig. 2c). The most enriched Bcl11b interaction partners included multiple members of the Nucleosome Remodeling Deacetylase complex (NuRD), RE-1-silencing transcription factor (Rest) complex (NRSF), and Lysine-specific demethylase 1A (Kdm1a or LSD1) complexes, with lower but still substantial scores for Polycomb repressor complex 1 (PRC1), confirming earlier evidence for Bcl11b-NuRD association[16] (Fig. 2d). Association of major components of these complexes, Chd4, Mta2, Rnf2 (Ring1b), Rest, LSD1 and Hdac2, with Bcl11b was validated by immune co-precipitation (Fig. 2e). Bcl11b has also been found in SWI/SNF complexes[40], and the Smarca4 (Brg1) component was also specifically enriched (Supplementary Table 5), but not most SWI/SNF components. Note that although LSD1, Chd4, Mta2 and Ring1b act as components of repressor complexes, in specific contexts they too may contribute to activation of genes[41–45]. Thus, both “repressor complexes” and Runx1 might play roles in Bcl11b-mediated positive or negative gene regulation.

Identification of Bcl11b-dependent cofactor binding sites

To test whether differential gene expression in primary pro-T cells was linked to genomic regions where specific cofactor assemblies might be nucleated by Bcl11b, we performed ChIP-seq analysis of wildtype and *Bcl11b*-deleted pro-T cells. To obtain the large cell numbers needed, we used *in vitro* differentiation cultures with Cre-ERT2 activation by 4-OH tamoxifen (4-OHT) to delete $Bcl11b^{fl/fl}$ (Supplementary Fig. 4c). On day 7, before 4-OHT treatment, most cells showed a DN2a/b phenotype ($Lin^{-}CD45^{+}c\text{-}Kit^{+}CD25^{+}$, Supplementary Fig. 4d). Five days after 4-OHT removal, control cells had efficiently progressed into a $c\text{-}Kit^{lo}$ DN3 stage, while *Bcl11b*-deficient cells were still characteristically $c\text{-}Kit^{hi+}$ (Fig. 3a, 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[7], Supplementary Fig. 4f) and these peaks almost completely disappeared in *Bcl11b*-deleted cells (Fig. 3b). Of the factors tested, only Brg1 could not be mapped using commercially available reagents (data not shown). In the control cells, Chd4, Mta2, Rest, Ring1b, LSD1 and Runx1 bound at ~6500 (Rest)–~33,000 (Runx1) sites each. These overlapped with 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 Bcl11b for their recruitment. *Bcl11b* deletion caused large fractions of some cofactor peaks to disappear (Mta2, Rest) or relocate (Chd4, Ring1b, LSD1, Runx1)(Fig. 3c). In each case, most of the Bcl11b-dependent peaks coincided with sites bound by Bcl11b in wildtype cells (Fig. 3c, green), implying that Bcl11b binding itself was needed to recruit that cofactor to such sites. The number of Mta2 peaks was especially sharply reduced by deletion of *Bcl11b,* although Mta2 protein abundance in *Bcl11b*-deficient cells was comparable to that in wildtype DN3 cells (Supplementary Fig. 4g), suggesting that most of its genome-wide associations in these cells depend on Bcl11b.

Only a small subset of Bcl11b sites showed Bcl11b-dependent cofactor binding, among the tens of thousands of Bcl11b peaks across the genome. Some of the genes repressed by Bcl11b not only had Bcl11b-dependent cofactor peaks (Fig. 4a–d, magenta rectangles), but

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

also showed newly generated cofactor peaks occupying distinct sites if *Bcl11b* was deleted (Fig. 4a,b, green rectangles). Motif preference for cofactor binding changed when Bcl11b was absent. Genome-wide, not only ETS and Runx family but also bHLH motifs were highly enriched among Bcl11b-dependent cofactor peaks (Supplementary Fig. 5a,b), and a minority of these sites included the motifs defined recently using 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") much less common and bZIP, HMG family, and other motifs more commonly enriched (Supplementary Fig. 5c). Bcl11b at CpG islands overlapped mostly with sites where Chd4, Ring1b, LSD1 or Runx1 were engaged whether *Bcl11b* was deleted or not (Supplementary Fig. 5d). Thus, Bcl11b both facilitated cofactor binding to a subset of Bcl11b occupancy sites and antagonized cofactor binding to other sites.

Bcl11b alters cofactor binding at functional target loci

Bcl11b-dependent cofactor peaks were found around genes that Bcl11b repressed (Fig. 4a–d, magenta boxes)[7], both in primary DN3 cells and in Scid.adh.2c2 cells (Supplementary Fig. 6a-d). At *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[22]. Similarly, at *Zbtb16* and *Tnni1* only one of several sites appeared to be Bcl11b-dependent, and at *Cd163l1*, different cofactors varied in Bcl11b-dependence at different sites (Fig. 4b–d, Supplementary Fig. 6b–d), suggesting that Bcl11b may interact separately with distinct complex subcomponents. Finally, selective Bcl11b-dependent cofactor recruitment was also seen at genes that were positively regulated by Bcl11b in pro-T cells, such as the *Cd3gde* cluster and *Cd6* (Supplementary Fig. 6e,f). The bifunctional transcription factor Runx1 was frequently recruited to Bcl11b sites around all classes of targets. Thus, Bcl11b binding recruited different cofactors to specific subsets of its genomic sites including both positively and negatively regulated loci.

Statistical evidence implied that this corecruitment was functionally relevant. Unlike Bcl11b binding itself, Bcl11b-dependent cofactor recruitment sites were enriched at Bcl11b-regulated loci (defined in Supplementary Fig. 4a, Supplementary Table 2). Among the DEGs linked to direct Bcl11b binding, Bcl11b-dependent cofactor peaks and "new" cofactor peaks that appeared only when *Bcl11b* was deleted were highly overrepresented (Fig. 4e,f), as compared to genes that also had Bcl11b binding but did not change expression when *Bcl11b* was deleted (Fig. 4e,f, right; non-DEG). Bcl11b-dependent Runx1 recruitment was particularly enriched at Bcl11b-repressed target genes, (Fig. 4e, left), while Bcl11b-dependent Rest was particularly depleted. Mta2 accompanied Bcl11b binding at DEGs and non-DEGs alike, but Bcl11b-dependent Chd4, Ring1b, and LSD1 recruitments were also significantly enriched at Bcl11b repression targets (Fig. 4e, left). As expected for repression mediators, these sites had minimal association with the active histone mark H3K27Ac. At sites linked to Bcl11b-dependent DEGs (Fig. 4e, middle), there was also highly significant enrichment for Bcl11b-dependent Runx1, Chd4, LSD1 and Ring1b recruitment, but here with highly enriched H3K27Ac as well. Interestingly, Bcl11b-repressed DEGs were most

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

specifically enriched for “new” cofactor peaks that appeared only when Bcl11b was absent (Chd4, Ring1b, LSD1 and Runx1 P<2.2E-16, Ring1b P<2.5E-9), as H3K27Ac marking of these genes also increased (Fig. 4f, left). Although “new” sites for Runx1 and LSD1 themselves rarely overlapped with strong Bcl11b occupancy in wildtype cells (Fig. 3c; Fig. 4a-b, green boxes), it appeared that Bcl11b’s presence inhibited their occupancy normally. Thus, Bcl11b-dependent cofactor localization, both recruitment and inhibition of recruitment, was a much stronger predictor of Bcl11b functionality at genomic sites than 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 tested whether Bcl11b-dependent co-recruitment could identify sites which exert particularly strong effects in repression. We disrupted either of two sites flanking *Id2* (Supplementary Fig. 7a, magenta) or one downstream of *Tnni1* (Fig. 4c) in Scid.adh.2C2 cells and scored them for target gene deregulation. Using 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, right). Loss of the *Tnni1* candidate silencer site (Sil +14k) elevated *Tnni1* expression at least 5-fold (*P*=9.52E-6) (Supplementary Fig. 7c). Deletion of one *Id2* candidate site (Sil +40k), though not another (Sil −600k), consistently raised *Id2* expression in the Scid.adh.2c2 cells above background (*P*=9.49E-4), despite the presence of numerous other Bcl11b sites around the locus (Supplementary Fig. 7a, blue, b). Thus, Bcl11b-dependent cofactor recruitment can help to identify functional repression sites in the genome.

Functional impact of cofactors in Bcl11b activities

Recruitment of the cofactors themselves contributed functionally to Bcl11b’s effects, as shown when we compared effects of acute disruption of *Bcl11b* with those of disrupting the genes encoding the cofactors. We generated sgRNAs against coding regions for Chd4, Mta1_2 (Mta1 and Mta2), Rest, Ring1a_b (Ring1a and Ring1b), LSD1, or Runx1, each with a Cyan Fluorescent Protein (CFP) reporter, and these were confirmed to eliminate target protein expression in Cas9-transduced Scid.adh.2c2 cells (Supplementary Fig. 8a)[46]. We then transduced them into Cas9-Bcl2 transgenic primary cells in parallel with sgRNA against *Bcl11b* alone (Supplementary Fig. 3a), to compare their effects directly during *in vitro* T cell development. At the time of transduction, after 7 d of OP9-DL1 coculture, most of the primary Cas9;Bcl2-transgenic cells were in DN2 stage (c-Kit^{+}CD25^{+}, Supplementary Fig. 8b). After seven more days, transduced (CFP^{+}) control cells (Supplementary Fig. 8c) had progressed into DN3 stage (Fig. 5a), whereas Bcl11b sgRNA-transduced cells showed the typical c-Kit^{hi+} DN2a-like phenotype. The effects on surface phenotype of cofactor deletion were milder than those of deletion of *Bcl11b* (Fig. 5a), although RNA transcript structures confirmed the biallelic deletions at the targeted sites in these loci (Supplementary Fig. 8d; arrowheads).

Cofactor deletion specifically affected RNA expression, more frequently at Bcl11b-regulated genes that were directly bound by Bcl11b (Fig. 5b–d; Supplementary Table 6; $|Log_2FC|>1$) than at background genes which were expressed independently of Bcl11b ($|Log_2FC|<0.5$,

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

RPKM>3 in sgControl) and lacked Bcl11b binding (Fig. 5b, right). Most Bcl11b-repressed genes that were directly bound by Bcl11b were also de-repressed upon deletion of at least one of the cofactors, showing distinct gene-specific patterns of cofactor response (Fig. 5b, left). Among Bcl11b-dependent targets, many were downregulated upon deletion of *Mta1*, *Mta2* or *Runx1* (Fig. 5b, middle). As summarized (Fig. 5c,d), deletion of *Mta1*, *Mta2* or *Runx1* caused the highest numbers of significant gene expression changes (Supplementary Table 6, p.adj<0.1), and were concordant with effects of *Bcl11b* deletion in 80–90% of significant Runx1-regulated genes, ~90% of significant Mta1,2-regulated genes. Thus, cofactor recruitment is functionally significant, but with target gene-specific functional requirements.

Cofactor recruitment by Bcl11b in TCR locus marking

Bcl11b-deficient mice can generate γδ lineage T cells but not αβ lineage T cells, associated with a failure of V_β-DJ_β rearrangement[1,3,6,9,10]. In BM-derived pro-T cells developing *in vitro*, Bcl11b and Runx1 co-occupied multiple sites across both the *Tcrb* complex and the *Tcrg* complex. However, these genomic regions differed markedly in their dependence for Runx1 binding on Bcl11b (Supplementary Fig 9). Whereas Runx1 binding across the V_β coding segments was highly Bcl11b-dependent (Supplementary Fig. 9a), its binding across the whole *Tcrg* complex was largely Bcl11b-independent (Supplementary Fig. 9b). Importantly, the difference between these loci was not revealed by differential RNA expression, for the *in vitro* cultured pro-T cells at this early stage showed minimal V_β transcription with or without Bcl11b (Supplementary Fig. 9a, RNA tracks). Instead, the difference in Bcl11b-dependent cofactor recruitment could indicate locus-specific roles in establishing permissive chromosome structure[18] which could underlie later Bcl11b impacts on DNA rearrangement.

Id2 and *Zbtb16* in the Bcl11b gene regulatory network

The direct effects of Bcl11b binding just defined leave open the possibility that Bcl11b impacts on pro-T cell gene expression could include indirect effects as well. Genes encoding the transcription factors Id2 and PLZF, which promote development of innate lymphoid cells including NK cells[24,28,47,48], were always up-regulated in *Bcl11b*-deficient pro-T cells[6,7] (Table 1; Supplementary Tables 1–3). The E2A antagonist, Id2, was of particular interest. We previously reported that *Bcl11b* knockout effects in fetal liver-derived pro-T cells[7] showed a surprisingly large overlap with *Tcf3* (encoding E2A) knockout effects[49]. Whereas the enrichment of bHLH motifs (Ptf1a or E2A) at Bcl11b sites for cofactor recruitment (Supplementary Fig. 5b) suggested possible complex formation, this was not supported by ChIP data: only 104 of 1,430 E2A published occupancy peaks in DN3 cells[50] overlapped with any of the ~26,000 Bcl11b peaks. This raised the possibility that repression of an intermediate regulator like Id2 could contribute to gene expression effects of Bcl11b.

To examine roles of Id2 and PLZF in *Bcl11b*-deficient cells, we carried out single and double deletion experiments. Cas9-Bcl2 transgenic BM precursors were co-transduced with sgRNAs targeting *Bcl11b* and/or *Id2* or *Zbtb16*, singly or in combination (as in Supplementary Fig. 3a). *Id2* and *Zbtb16* showed increased expression in *Bcl11b*-deficient cells but not in *Bcl11b-Id2* or *Bcl11b-Zbtb16* double KO (DKO) cells, respectively. While

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

the characteristic upregulation of c-Kit was comparable in *Bcl11b* KO cells and *Bcl11b-Zbtb16* DKO cells, *Bcl11b-Id2* DKO cells expressed slightly lower c-Kit (Fig. 6a), implying that Id2 may be involved in the abnormal elevation of c-Kit in *Bcl11b*-deficient cells. Interestingly, up-regulation of *Zbtb16* in *Bcl11b* KO cells was also weakened in *Bcl11b-Id2* DKO cells (Supplementary Fig. 10), and this was confirmed at the PLZF protein level (Fig. 6b). Thus, Id2 is involved in up-regulation of *Zbtb16* expression in *Bcl11b*-deficient cells.

Most genes up-regulated by *Bcl11b*-deletion (Bcl11b-repressed genes) showed overall similar responses in *Bcl11b* KO, *Bcl11b-Id2* DKO or *Bcl11b-Zbtb16* DKO samples (Fig. 6c, Supplementary Table 7), indicating that most of these 410 Bcl11b repression target genes do not depend on Id2 or PLZF for their expression (Supplementary Table 7; DEGs with FDR<0.05 vs. control: 410 for *Bcl11b* KO; 299 for DKO)(Fig. 6c, top). In contrast, among 349 Bcl11b-dependent genes, the down-regulation of many T cell program genes was ameliorated if *Bcl11b* was deleted together with *Zbtb16* or especially *Id2* (Fig. 6c, bottom). Double deletion of *Id2* with *Bcl11b* gave protection to well over half of the Bcl11b-dependent genes (Supplementary Table 7; DEGs with FDR<0.05 vs. control: 349 for *Bcl11b* KO; only 72 for DKO), suggesting that many of these genes are dependent on E proteins.

Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis showed that *Bcl11b* deletion up-regulated genes in ‘Cytokine-cytokine receptor interaction’ and ‘Natural killer cell mediated cytotoxicity’ pathways preferentially, while reducing expression of ‘T cell receptor signaling’ pathways. Although the ‘Natural killer cell mediated cytotoxicity’ genes were upregulated in *Bcl11b* KO cells with or without *Id2* or *Zbtb16*, there was markedly less upregulation of the ‘Cytokine-cytokine receptor interaction’ genes in DKO samples, especially the *Bcl11b-Id2* DKO, than in single *Bcl11b* KO samples, and less inhibition of ‘T cell receptor signaling’ genes (Fig. 6d,e). Thus, despite the weak overall effects of *Id2* and *Zbtb16* on gene expression profiles of *Bcl11b*-deficient cells, they had impacts on regulation of selective sets of genes, related to cytokine-cytokine receptor and TCR signaling pathways specifically.

Id2-Bcl11b double deletion did not always attenuate the effects of *Bcl11b* loss; for some Bcl11b repression targets, it amplified the effects of *Bcl11b* disruption (Supplementary Table 7). Certain genes highly enriched in TCRγδ IEL, including *Heyl and Cited4,* and the B cell and progenitor-cell-associated regulatory gene *Bcl11a*, were upregulated substantially more in *Bcl11b-Id2* DKO than in *Bcl11b* single KO samples. These results imply that *Bcl11a, Heyl* and *Cited4* depend on positive regulation by E proteins even though they are also a target of repression by Bcl11b. Thus, the upregulation of Id2 could sometimes mediate but in other cases mask the full spectrum of regulatory effects of Bcl11b.

Id2 and Zbtb16 promote different alternative programs

Bcl11b deletion upregulates genes associated with various alternative fates, and it has not been clear whether this response represents one coordinated program or several. Confirming previous reports[5,6,8], we repeatedly detected more Lin^+ cells in *Bcl11b*-deficient cells than control cells, especially cells expressing NK1.1 and CD11c, even in T cell developmental conditions (Supplementary Fig. 11a,b, OP9-DL1)(cf. ref[6]). When *Bcl11b*-deficient cells were removed from Notch signals on day 10, to allow non-T cell fates to be expressed, the

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

percentage of Lin^+ cells rose over the next 4d with NK1.1 and CD11c expression most prominent (Supplementary Fig. 11a,b, OP9-Mig). *Bcl11b*-deficient pro-T cells express these markers less strongly than mature NK or dendritic cells, respectively (Supplementary Fig. 11c), but they indicate distinct altered regulatory states, and double deletion experiments showed that Id2 and PLZF play distinct roles in these states. Both the percentage and the cell number of Lin^+ cells were significantly decreased in *Bcl11b-Id2* DKO cells, indicating that they depend on this E protein antagonist. However, Lin^+ cells were unexpectedly increased in *Bcl11b-Zbtb16* DKO cells, both on OP9-DL1 and on OP9-Mig (note scale change: Fig. 7a; Supplementary Fig. 11d,e). Generation of $NK1.1^+$ cells was attenuated in *Bcl11b-Id2* DKO cells and reduced further in *Bcl11b-Zbtb16* DKO cells. In contrast, while generation of $CD11c^+$ cells was also reduced in *Bcl11b-Id2* DKO, it was sharply enhanced in *Bcl11b-Zbtb16* DKO cells (Fig. 7a, Supplementary Fig. 11d,e).

The effects of these single and double deletions were cell-intrinsic (Fig. 7b and Supplementary Fig. 12a,b). When we cocultured *Bcl11b* KO and *Bcl11b-Id2* or *Bcl11b-Zbtb16* dKO cells together with distinctly marked reference standard *Bcl11b* single KO cells and GFP^+ empty-vector control cells in the same well (Supplementary Fig. 12a), the patterns of altered developmental phenotypes were determined only by the genetic perturbation history of the cells themselves (Fig. 7b, quantitated in Supplementary Fig. 12b).

Thus, PLZF not only supported generation of $NK1.1^+$ cells but also restrained expression of CD11c in *Bcl11b* KO cells (Fig. 7a,c), forcing the *Bcl11b* KO phenotype normally closer to an NK-like profile. In turn, Id2 had stronger roles in other features of *Bcl11b*-deficient cells including DN2a-like high c-Kit expression, up-regulated *Zbtb16* expression, increased expression of cytokine and chemokine receptors, decreased expression of T cell program genes and opening the way for generation of $NK1.1^+$ and $CD11c^+$ cells (Fig. 7c,d). The ability of Bcl11b to repress both *Id2* and *Zbtb16* is thus crucial for completing conventional T cell commitment.

DISCUSSION

The importance of Bcl11b for T cell lineage commitment has raised the question of how this factor works to promote and consolidate T-cell identity[5,6,8]. Unlike Pax5 and EBF1 in the B cell lineage[51–54], Bcl11b effects on T-cell gene expression appear more limited and more related to immune activation thresholds – “adaptive” vs. “innate-like” – than to T-cell identity *per se*[6–8,10,12,55]. A recent study[18] showed Bcl11b to bind to numerous sites throughout the active topological domains of the genome in early T cells, suggesting a possible global organizational role, but leaving 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 Bcl11b binding to specific genomic sites controls target gene expression during T cell commitment. We show that Bcl11b can repress targets directly by nucleating complexes of corepressors on the DNA at specific sites where they do not otherwise assemble. Such sites of Bcl11b-dependent cofactor recruitment and redirection turn out to be a far better statistical discriminator of genes that Bcl11b actually controls than Bcl11b binding alone. At the same time, a substantial minority of Bcl11b effects are apparently indirect, mediated through its repression of the regulatory genes *Id2*

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

and *Zbtb16*. Acute double-deletion analysis using the Cas9 system has revealed that many effects of *Bcl11b* deletion are responses to the resultant increases in Id2 and/or PLZF. In particular, many T lineage-specific genes that apparently depend on Bcl11b for activation during commitment in fact require Bcl11b largely to suppress Id2. Thus, in addition to its 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 available statistical criteria for enriching loci that Bcl11b regulates, positively or negatively, have turned out to be the presence of sites where assembly of chromatin modifying complexes depended on Bcl11b binding. Such complexes often included Chd4, Mta2, Ring1b, LSD1, and Runx1; future work should also test Brg1. Bcl11b-dependent nucleation of cofactor complexes was only seen at a small minority of Bcl11b sites and a minority of the sites for most of these cofactors. Deletion analysis of individual chromatin modification complex components showed that these components, stably expressed throughout commitment[56,57], were often important for Bcl11b effects, especially at loci that Bcl11b represses. Among factors redeployed by Bcl11b, Runx1 was most potent for both activation and repression of target genes. Notably, the components that depended most on Bcl11b for recruitment differed from one genomic site to another. Although much remains to be learned about the mechanistic rules for transcriptional repression, this suggests that Bcl11b can interact with a variety of subunit assemblies individually, not only with pre-formed NuRD or PRC1 complexes as a whole.

New cofactor peaks also appeared in *Bcl11b*-deleted cells at sites never bound by Bcl11b in pro-T cells, with a distinct motif enrichment signature, and these were strongly associated with abnormal activation of Bcl11b-repressed genes. Thus, in addition to recruiting repression complexes, Bcl11b may also repress target genes by preventing cofactors from assembling at other neighboring sites that could otherwise serve different, activating transcription factors. Gene regulation by redirection of limiting pools of cofactors, with or without direct DNA binding, is consistent with recent findings on PU.1-mediated gene regulation in early pro-T cells[46]. Therefore, gene regulation by partner factor redeployment could be a common mechanism for transcription factors that have crucial roles in cell fate decisions.

In summary, we have shown that the advent of Bcl11b expression during commitment re-centers T cell regulatory circuits in multiple ways. Bcl11b physically redirects the binding of Runx1 and chromatin modulating complexes across the genome in site-specific ways that are highly concentrated around Bcl11b-regulated target genes. These biochemical collaborations, especially with Runx1, are likely to be directly functional, especially for repression. At the same time, Bcl11b's repression of PLZF and Id2 blocks implementation of at least two alternative programs, distinctly regulated by these factors. Finally, its control of Id2 gives it a position of indirect power in a pro-T cell network where the main direct effectors are E proteins and their own interaction partners. This network is particularly important for TCR gene expression and the stringent cell biology of β-selection, and may well explain the importance of Bcl11b for these crucial milestones.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

METHODS

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 Life Sciences Reporting Summary that accompanies this paper.

Mice

C57BL/6 (referred to as B6), B6.Cg-Tg(BCL2)25Wehi/J (Bcl2-tg)[58] and B6.Gt(ROSA)26$^{Sortm1.1(CAG-cas9*,-EGFP)Fezh}$/J (Cas9)[59] mice were purchased from the Jackson Laboratory. *Vav1*-iCre mice[35] (B6N.Cg-*Commd10*$^{Tg(Vav1-icre)A2Kio}$/J) were purchased from Jackson Laboratories and *Lck*-Cre (*Lck* proximal promoter) mice developed by C. Wilson's group (B6.Cg-Tg(*Lck*-Cre)1Cwi N9)[36] were purchased from Taconic Laboratories. The Cre activity reporter allele *ROSA26R-eYFP*[60] was also used in our previous studies of *Bcl11b* gene function[6,7,61]. Except for *Vav1*-iCre, which was maintained in heterozygotes, the indicated transgenes were bred to homozygosity alone or in combinations on the B6 background. *Bcl11b*$^{fl/fl}$;*Rosa26-Cre-ERT2* mice[8] were derived from stock originally kindly provided by P. Liu (Wellcome Sanger Institute, Cambridge, UK) as previously described[7], 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.

Cells and cell culture

Thymuses from 4–6wk old *Bcl11b*$^{+/+}$, *Bcl11b*$^{fl/+}$, and *Bcl11b*$^{fl/fl}$ *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 (H57597), 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). Eluted DN cells were stained with CD117 (c-Kit)-PE (2b8), CD25-APCe780 (PC61.5) and CD44-e450, and YFP^{+}CD25^{+} and YFP^{-} CD25^{+} cells were sorted by FACSAria (BD Bioscience). For flow cytometric analysis, eluted DN cells were stained as above for FACS purification and analyzed with FlowJo software (Tree Star).

Scid.adh.2c2 cells [38] were cultured in RPMI1640 with 10% fetal bovine serum (FBS, Sigma-Aldrich), sodium pyruvate (Gibco), non-essential amino acids (Gibco), Pen-Strep-Glutamine (Gibco) and 50 μM β-mercaptoethanol (Sigma-Aldrich).

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

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

through a magnetic column (Miltenyi Biotec). Then, $Lin^{-}Sca1^{+}c\text{-}Kit^{+}$ (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.

In samples that were tested for developmental lineage alterations, the final passage at day 10 was either onto fresh OP9-DL1 or onto OP9-control monolayers without the Notch ligand DL1 (OP9-Mig, referred to as OP9-control in ref. [39]).

Acute deletion of *Bcl11b* in T-cell development cultures

We used two methods to delete *Bcl11b* acutely *in vitro*, at a known timepoint and developmental stage, for the perturbation experiments in this study.

To generate sufficient numbers of control and Bcl11b-deleted cells for ChIP-seq analysis, we used acute deletion of floxed *Bcl11b* in cells with Cre-ERT2 activated by 4-OH tamoxifen. BM-derived precursor cells from control $Bcl11b^{+/+}$Cre-ERT2 mice and from $Bcl11b^{fl/fl}$Cre-ERT2 mice were cultured in parallel on OP9-DL1 stroma to day 7 as described above. Upon reculture, both controls and experimental samples were supplemented with 120 nM of 4-OH tamoxifen (4-OHT, Sigma-Aldrich). Thus, both control and experimental samples contained the Cre-ERT2 transgenes, were exposed to 4-OHT, and subject to Cre activity. Two days later 4-OHT was removed, and cells were cultured for an additional 5 days (to day 14 overall) on fresh OP9-DL1 with 1 ng/ml of IL-7, 1 ng/ml of Flt3L, as shown in Supplementary Fig. 4c. They were then harvested for further analysis and preparation of DNA for ChIP.

RNA expression responses to real-time perturbation were measured using Cas9 plus guide RNA for acute mutational disruption of *Bcl11b, Runx1, Id2, Zbtb16,* and/or genes encoding cofactors. This method was superior to the Cre-ERT2 method for RNA-seq comparisons, because all perturbations could be carried out on the identical genetic background and the potentially toxic effects of 4-OHT with Cre[62] could be eliminated. To generate input cells, Cas9 mice were first bred to *Bcl2*-tg mice to generate heterozygotes for both transgenes. Bone marrow cells from these *Cas9*;*Bcl2*-tg animals (called Cas9-Bcl2 mice) were then used to seed *in vitro* differentiation cultures as above. At day 7, the cells were transduced with retroviral vectors encoding a fluorescent reporter (CFP or hNGFR) and the indicated guide RNAs (sgRNAs) as detailed below, and then returned to OP9-DL1 culture. Cells were analyzed after another 7 days of culture (to day 14 overall), or as indicated in specific experiments.

Flow cytometry analysis

For staining of sgRNA-introduced BM cells, surface antibodies against CD45, c-Kit, CD25 and a biotin-conjugated lineage cocktail (CD8α, CD11b, CD11c, Gr1, TER-119, NK1.1,

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

CD19, TCRβ, TCRγδ) were used for staining. Prior to cell surface staining cells were treated with 2.4G2 cell supernatant. Where indicated, intracellular staining using the BD cytofix/cytoperm Kit (BD Bioscience) was carried out with anti-PLZF AlexaFluor 647 (BD Bioscience). All of the cells were analyzed using a flow cytometer, MacsQuant 10 (Miltenyi) with FlowJo software (Tree Star).

Cloning

Myc-Flag-tagged cDNA for *Bcl11b* was inserted into a multi-cloning site of the pMxs-IRES-GFP vector. The Cas9-GFP expression vector (pQCXIN-EF1a-mNeonGreen-P2A-Cas9) and sgRNA-CFP expression vector (E42-dTet-CFP, in which mTurquoise is the Cyan Fluorescent Protein reporter) were described previously[46]. 19-mer sgRNAs were designed using the CHOPCHOP web tool (https://chopchop.rc.fas.harvard.edu/) and inserted into the empty sgRNA-expression vector by PCR-based insertion (sgRNA sequences provided in Supplementary Table 8). Three sgRNA-expression vectors were generated for one gene, and pooled retroviral plasmids were used to make retroviral supernatant.

Retroviral infection

The methods used to generate the virus supernatant and for infection were described previously[63]. Cultured Lin^- BM cells for 7 days were disaggregated, filtered through 40-μm nylon mesh, transferred onto RetroNectin-coated virus bound plates, and cultured with OP9 medium supplemented with 5 ng/ml of IL-7, 5 ng/ml of Flt3L and 10 ng/ml of SCF. Infected cells were cultured for an additional 7 days on OP9-DL1 and subjected to further analysis. For sorting, cells were stained with CD45, CD25, and a biotin-conjugated lineage cocktail (CD8α, CD11b, CD11c, Gr1, TER-119, NK1.1, CD19, TCRβ, TCRγδ), and were sorted for $CD25^+$ infected cells ($Lin^-CD45^+CD25^+CFP^+$).

Two-step affinity purification of Bcl11b complexes from the DN3-like cell line Scid.adh.2c2

Scid.adh.2c2 cells[38,39] were infected with either mock control (pMxs-IRES-GFP) or Myc-Flag-Bcl11b-containing retrovirus. Three days after infection, Myc-Flag-tagged Bcl11b-infected Scid.adh.2c2 cells were solubilized with the following protease inhibitor-containing IP 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), and lysed on ice for 30 min with gentle shaking and sonicated on a Misonix S-4000 sonicator (Qsonica) for 3 cycles, amplitude 20 for 30 sec. followed by 30 sec. rest. We did not add benzonase or ethidium 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 ºC. Immune complexes were eluted from the agarose by 3xFlag peptide (Sigma-Aldrich), and the eluted Bcl11b complexes were subjected 2nd immunoprecipitation with anti-Myc gel (MBL). Immune complexes were eluted from the gel with Myc peptide (MBL) and separated by SDS-PAGE. The bands were excised from the gel and subjected to a mass spectrometric analysis to identify corresponding proteins. The gel pieces were washed twice with 100 mM bicarbonate in acetonitrile and the proteins were digested with trypsin. After adding 0.1% formic acid to the supernatant, the peptides were analyzed by liquid

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

chromatography-tandem mass spectrometry (LC-MS/MS) with an Advance UHPLC (Bruker) and an Orbitrap Velos Pro Mass Spectrometer (Thermo Fisher Scientific). The resulting MS/MS data set was analyzed using the Mascot software program (Matrix Science).

Gene Ontology and KEGG pathway analysis

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

Immunoprecipitation and immunoblotting

Protein extracts from Myc-Flag-tagged Bcl11b-infected Scid.adh.2c2 cells were subjected to immunoprecipitation as described previously[43]. Nuclear extracts were prepared using NE-PER Nuclear and Cytoplasmic Extraction Reagents (Thermo Scientific). The antibodies used for the immunoblot analyses were anti-Chd4 (A301–081A, Bethyl), anti-Mta2 (sc-9447, Santa Cruz Biotechnology, Inc.), anti-HDAC2 (ab12169, Abcam), anti-Rest (12C11–1B11, Caltech Protein Expression Center), anti-Ring1b (A302–869A, Bethyl), anti-LSD1 (ab17721, Abcam), anti-Runx1 (ab23980, Abcam), anti-Bcl11b (ab18465, Abcam), anti-Lamin B (sc-6217, Santa Cruz Biotechnology, Inc.), and anti-Myc (My3, MBL).

Chromatin Immunoprecipitation (ChIP) and ChIP-sequencing

1×10^7 BM-derived DN3 cells were fixed with 1% formaldehyde in α-MEM for 10 min (H3K27Ac), or with 1 mg/ml DSG (Thermo Scientific) in PBS for 30 min at ~22°C followed by an additional 10 min with addition of formaldehyde up to 1% (Bcl11b, Chd4, Mta2, HDAC2, Rest, Ring1b, LSD1 and Runx1). The reaction was quenched by 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, 10 mM EDTA, 0.5 mM EGTA, 50 mM Tris-HCl (pH 8) and PIC) and sonicated on a Bioruptor (Diagenode) for 18 cycles of 30sec sonication followed by 30sec rest, with max power. Six μg per 10^7 cells of anti-Bcl11b Abs (a mixture of A300–383A (Bethyl), A300–385A (Bethyl), ab18465 (Abcam) and 12120 (CST)), or anti-Chd4 Ab (A301–081A), anti-Mta2 Ab (sc-9447), anti-HDAC2 Ab (ab12169), anti-Rest Ab (12C11–1B11), anti-Ring1b Ab (A302–869A), anti-LSD1 Ab (ab17721), anti-Runx1 Ab (ab23980), or anti-H3K27Ac Ab (ab4729) were each separately pre-bound to Dynabeads anti-Rabbit, Dynabeads anti-Mouse or Dynabeads Protein A/G (Invitrogen) and then added individually to the diluted chromatin complexes in parallel aliquots. The samples were incubated overnight at 4 °C, then 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 up using Zymo ChIP DNA Clean & Concentrator. ChIP-seq libraries were constructed using NEBNext ChIP-Seq Library Preparation Kit (E6240, NEB) and sequenced on Illumina HiSeq2500 in single read mode with the read length of 50 nt. Analysis pipelines used are described below under ChIP-seq Analysis and RNA-seq Analysis.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

mRNA-preparation and RNA-sequencing

Total RNA was isolated from 3×10^5 of cultured cells and $2–5 \times 10^4$ of ex vivo cells using RNeasy Micro Kit (Qiagen). Libraries were constructed using NEBNext Ultra RNA Library Prep Kit for Illumina (E7530, NEB) 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.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) [64]. The NCBI37/mm9 assembly was chosen for ChIP-seq sample mapping in this study to ease comparisons with numerous 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 -L.P.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://sites.google.com/site/anshulkundaje/projects/idr). Only reproducible high quality peaks, with a normalized peak score ≥ 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-sequenced reads were mapped onto the mouse genome build NCBI37/mm9 with STAR (v2.4.0) [65] and post-processed with RSEM (v1.2.25; http://deweylab.github.io/RSEM/) [66] according to the settings in the ENCODE long-rna-seq-pipeline (https://github.com/ENCODE-DCC/long-rna-seq-pipeline/blob/master/DAC/STAR_RSEM.sh) with the minor modifications that settings *"--output-genome-bam --sampling-for-bam*" was added to *rsem-calculate-expression*. STAR and RSEM reference libraries were created from genome build NCBI37/mm9 together with the Ensembl gene model file Mus_musculus.NCBIM37.66.gtf. The resulting bam-files were used to create HOMER [67] tag directories (*makeTagDirectory* with –keepAll setting). For analysis of statistical significance among differentially expressed genes the raw gene counts were derived from each tag directory with *analyzeRepeats.pl* with the *–noadj -condenseGenes* options followed by the *getDiffExpression.pl* command using EdgeR (v3.6.8; http://bioconductor.org/packages/release/bioc/html/edgeR.html) [68]. For data visualization, rpkm normalized reads were derived using the *analyzeRepeats.pl* command with the options *–count exons –condenseGenes –rpkm* followed by log transformation. The normalized datasets were hierarchically clustered with "average" linkage and visualized in MatLab (clustergram).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

UCSC Genome Browser bigwig visualization

BigWigs were generated from the aligned SAM or BED-file formats using *Samtools* [69], *Bedtools* [70] and the *UCSC genomeCoverageBed* and *bedGraphToBigWig* and 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) [64] for visualization.

Position-weight matrix (PWM) scanning of Bcl11b associated peak lists

Liu and coworkers[37] recently reported new sequence motifs to be specific Bcl11b recognition targets, based on *in vitro* protein binding microarray reactivities of purified Bcl11b full length ("XL"), or the middle two zinc fingers of Bcl11b (Bcl11b_23), or the C-terminal three zinc fingers of Bcl11b (Bcl11b_456). To use HOMER to quantify the representation of these Bcl11b motifs in Bcl11b or Bcl11b co-factor ChIP peaks, 23-mer enriched protein binding microarray-defined probability matrices (Bcl11b_23_rep2, Bcl11b _456_rep1, Bcl11b _XL) were downloaded from ref. [37]. From these 23-mer matrices, a Homer compatible 12-mer PWM was then created from the 7th to 18th positions in each of the 23-mer matrices with the log-odds detection threshold manually set to 5, for relatively inclusive scoring. To detect presence of these PWMs in the peak files, 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, $|\log_2 FC|>1$, and RPKM>1 except where otherwise indicated, based on 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.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

We thank D. Perez, J. 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 valuable discussion and reagents. A fellowship from the Manpei Suzuki Diabetes Foundation (to H.H.) is gratefully acknowledged. This work was supported by grants from the USPHS to E.V.R. (R01AI083514 and R01HD076915) and Grants-in-Aid for Advanced Research and Development Programs for Medical Innovation, the Takeda Science Foundation, and SENSHINE Medical Research Foundation (to T.T.). This work was partly performed in the Collaborative Research Project Program of the Medical Institute of Bioregulation, Kyushu University, and also received support from a Swedish Research Council fellowship (to J.U.), the California Institute of Regenerative Medicine Bridges to Stem Cell Research Program (Pasadena City College and Caltech, to M.R.-W.), the L. A. Garfinkle Memorial Laboratory Fund and the Al Sherman Foundation, special project funds from the Provost and Division of Biology & Biological Engineering of Caltech, and the Albert Billings Ruddock Professorship to E.V.R.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

REFERENCES

1. Wakabayashi Y et al. Bcl11b is required for differentiation and survival of αβ T lymphocytes. Nat Immunol 4, 533–539 (2003). [PubMed: 12717433]
2. Liu P, Li P & Burke S Critical roles of Bcl11b in T-cell development and maintenance of T-cell identity. *Immunol. Rev.* 238, 138–149 (2010). [PubMed: 20969590]
3. Shibata K et al. IFN-γ-producing and IL-17-producing γδ T cells differentiate at distinct developmental stages in murine fetal thymus. J Immunol 192, 2210–2218 (2014). [PubMed: 24489104]
4. Kueh HY et al. Asynchronous combinatorial action of four regulatory factors activates *Bcl11b* for T cell commitment. Nat Immunol 17, 956–965 (2016). [PubMed: 27376470]
5. Ikawa T et al. An essential developmental checkpoint for production of the T cell lineage. Science 329, 93–96 (2010). [PubMed: 20595615]
6. Li L, Leid M & Rothenberg EV An early T cell lineage commitment checkpoint dependent on the transcription factor Bcl11b. Science 329, 89–93 (2010). [PubMed: 20595614]
7. Longabaugh WJR et al. Bcl11b and combinatorial resolution of cell fate in the T-cell gene regulatory network. Proc Natl Acad Sci U S A 114, 5800–5807 (2017). [PubMed: 28584128]
8. Li P et al. Reprogramming of T cells to natural killer-like cells upon Bcl11b deletion. Science 329, 85–89 (2010). [PubMed: 20538915]
9. Inoue J et al. Expression of TCRαβ partly rescues developmental arrest and apoptosis of αβ T cells in Bcl11b$^{-/-}$ mice. J Immunol. 176, 5871–5879 (2006). [PubMed: 16670294]
10. 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). [PubMed: 20544728]
11. 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). [PubMed: 25422283]
12. 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). [PubMed: 25128552]
13. Kojo S et al. Priming of lineage-specifying genes by Bcl11b is required for lineage choice in post-selection thymocytes. Nat Commun 8, 702 (2017). [PubMed: 28951542]
14. Uddin MN et al. Transcription factor Bcl11b sustains iNKT1 and iNKT2 cell programs, restricts iNKT17 cell program, and governs iNKT cell survival. Proc Natl Acad Sci U S A 113, 7608–7613 (2016). [PubMed: 27330109]
15. Avram 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). [PubMed: 12196208]
16. Cismasiu VB et al. BCL11B functionally associates with the NuRD complex in T lymphocytes to repress targeted promoter. Oncogene 24, 6753–6764 (2005). [PubMed: 16091750]
17. 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). [PubMed: 12930829]
18. Hu G et al. Transformation of accessible chromatin and 3D nucleome underlies lineage commitment of early T cells. Immunity 48, 227–242 (2018). [PubMed: 29466755]
19. Califano D et al. Transcription factor Bcl11b controls identity and function of mature Type 2 innate lymphoid cells. Immunity 43, 354–368 (2015). [PubMed: 26231117]
20. Yu Y et al. Single-cell RNA-seq identifies a PD-1^{hi} ILC progenitor and defines its development pathway. Nature 539, 102–106 (2016). [PubMed: 27749818]
21. 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). [PubMed: 28514688]
22. Mowel WK 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, 435–449 (2017). [PubMed: 28930659]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

23. Delconte RB et al. The helix-loop-helix protein ID2 governs NK cell fate by tuning their sensitivity to Interleukin-15. Immunity 44, 103–115 (2016). [PubMed: 26795246]
24. Seillet C et al. Deciphering the Innate Lymphoid Cell transcriptional program. Cell Rep 17, 436–447 (2016). [PubMed: 27705792]
25. Rankin L & Belz GT Diverse roles of inhibitor of differentiation 2 in adaptive immunity. Clin Dev Immunol 2011, 281569 (2011). [PubMed: 21437223]
26. Boos MD, Yokota Y, Eberl G & Kee BL 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). [PubMed: 17452521]
27. Constantinides MG et al. PLZF expression maps the early stages of ILC1 lineage development. Proc Natl Acad Sci U S A 112, 5123–5128 (2015). [PubMed: 25838284]
28. Constantinides MG, McDonald BD, Verhoef PA & Bendelac A A committed precursor to innate lymphoid cells. Nature 508, 397–401 (2014). [PubMed: 24509713]
29. Raberger J et al. The transcriptional regulator PLZF induces the development of CD44 high memory phenotype T cells. Proc Natl Acad Sci U S A 105, 17919–17924 (2008). [PubMed: 19004789]
30. Savage AK et al. The transcription factor PLZF directs the effector program of the NKT cell lineage. Immunity 29, 391–403 (2008). [PubMed: 18703361]
31. 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). [PubMed: 18660811]
32. Miller JC et al. Deciphering the transcriptional network of the dendritic cell lineage. Nat Immunol 13, 888–899 (2012). [PubMed: 22797772]
33. Satpathy AT, Wu X, Albring JC & Murphy KM Re(de)fining the dendritic cell lineage. Nat Immunol 13, 1145–1154 (2012). [PubMed: 23160217]
34. Doulatov S et al. PLZF is a regulator of homeostatic and cytokine-induced myeloid development. Genes Dev. 23, 2076–2087 (2009). [PubMed: 19723763]
35. de Boer J et al. Transgenic mice with hematopoietic and lymphoid specific expression of Cre. Eur J Immunol 33, 314–325 (2003). [PubMed: 12548562]
36. Lee PP et al. A critical role for Dnmt1 and DNA methylation in T cell development, function, and survival. Immunity 15, 763–774 (2001). [PubMed: 11728338]
37. Liu N et al. Direct Promoter Repression by BCL11A Controls the Fetal to Adult Hemoglobin Switch. Cell 173, 430–442 (2018). [PubMed: 29606353]
38. Dionne CJ et al. Subversion of T lineage commitment by PU.1 in a clonal cell line system. Dev Biol 280, 448–466 (2005). [PubMed: 15882585]
39. Del Real MM & Rothenberg EV Architecture of a lymphomyeloid developmental switch controlled by PU.1, Notch and Gata3. Development 140, 1207–1219 (2013). [PubMed: 23444353]
40. Kadoch C & Crabtree GR Mammalian SWI/SNF chromatin remodeling complexes and cancer: Mechanistic insights gained from human genomics. Sci Adv 1, e1500447 (2015). [PubMed: 26601204]
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). [PubMed: 29070704]
42. Gao Z et al. An AUTS2-Polycomb complex activates gene expression in the CNS. Nature 516, 349–354 (2014). [PubMed: 25519132]
43. Hosokawa H et al. Functionally distinct Gata3/Chd4 complexes coordinately establish T helper 2 (Th2) cell identity. Proc Natl Acad Sci U S A 110, 4691–4696 (2013). [PubMed: 23471993]
44. Miccio A et al. NuRD mediates activating and repressive functions of GATA-1 and FOG-1 during blood development. EMBO J. 29, 442–456 (2010). [PubMed: 19927129]
45. Williams CJ et al. The chromatin remodeler Mi-2β is required for CD4 expression and T cell development. Immunity 20, 719–733 (2004). [PubMed: 15189737]
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 (2018). [PubMed: 29924977]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

47. Ishizuka IE 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). [PubMed: 26779601]
48. Di Santo JP Staying innate: transcription factor maintenance of innate lymphoid cell identity. Immunol Rev 261, 169–176 (2014). [PubMed: 25123284]
49. Xu W et al. E2A transcription factors limit expression of Gata3 to facilitate T lymphocyte lineage commitment. Blood 121, 1534–1542 (2013). [PubMed: 23297135]
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). [PubMed: 21857655]
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). [PubMed: 22669466]
52. McManus 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). [PubMed: 21552207]
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). [PubMed: 25123279]
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). [PubMed: 20451411]
55. Albu DI et al. BCL11B is required for positive selection and survival of double-positive thymocytes. J Exp Med 204, 3003–3015 (2007). [PubMed: 17998389]
56. Mingueneau M et al. The transcriptional landscape of αβ T cell differentiation. Nat Immunol 14, 619–632 (2013). [PubMed: 23644507]
57. Heng TSP, Painter MW & Consortium TIGP The Immunological Genome Project: networks of gene expression in immune cells. Nat Immunol. 9, 1091–1094 (2008). [PubMed: 18800157]
58. Strasser A, Harris AW & Cory S *bcl-2* transgene inhibits T cell death and perturbs thymic self-censorship. Cell 67, 889–899 (1991). [PubMed: 1959134]
59. Platt RJ et al. CRISPR-Cas9 knockin mice for genome editing and cancer modeling. Cell 159, 440–455 (2014). [PubMed: 25263330]
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). [PubMed: 11299042]
61. Scripture-Adams DD et al. GATA-3 dose-dependent checkpoints in early T cell commitment. J Immunol 193, 3470–3491 (2014). [PubMed: 25172496]
62. Higashi AY et al. Direct hematological toxicity and illegitimate chromosomal recombination caused by the systemic activation of CreERT2. J.Immunol. 182, 5633–5640 (2009). [PubMed: 19380810]
63. Champhekar 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). [PubMed: 25846797]
64. Speir ML et al. The UCSC Genome Browser database: 2016 update. Nucleic Acids Res 44, D717–D725 (2016). [PubMed: 26590259]
65. Dobin A et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21 (2013). [PubMed: 23104886]
66. Li B & Dewey CN RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics 12, 323 (2011). [PubMed: 21816040]
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). [PubMed: 20513432]
68. Robinson MD, McCarthy DJ & Smyth GK edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26, 139–140 (2010). [PubMed: 19910308]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

69. Li H et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2078–2079 (2009). [PubMed: 19505943]

70. Quinlan AR & Hall IM BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26, 841–842 (2010). [PubMed: 20110278]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

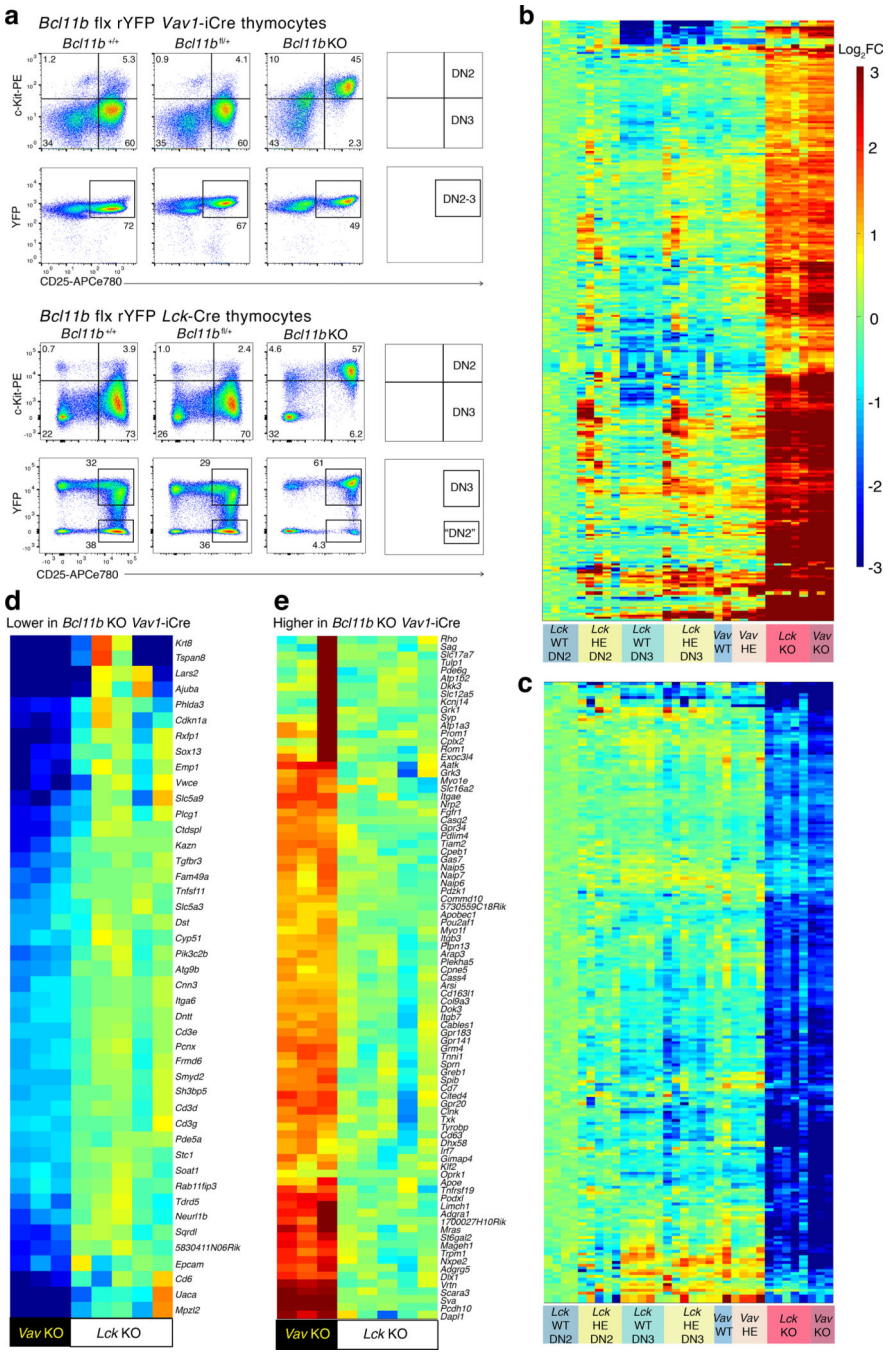


Figure 1: Cellular and molecular phenotypes of *in vivo Bcl11b* deletion by *Vav1*-iCre or *Lck*-Cre.

(a), Representative flow cytometric analysis of DN thymocytes from *Bcl11b*$^{+/+}$, *Bcl11b*$^{fl/+}$, and *Bcl11b*$^{fl/fl}$ rYFP *Vav1*-iCre or *Lck*-Cre mice, showing gates used for defining DN subpopulations (top) and sorting strategy used for purifying YFP^{-}CD25^{+} and YFP^{+}CD25^{+} thymocytes (bottom). n ≥ 5 mice of each genotype.

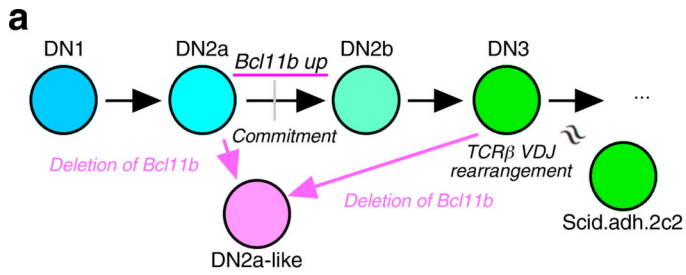
(b, c), Heatmaps generated from transcriptomic analysis of sorted DN thymocytes using RNA-seq. Columns within each category represent cells from individual mice. (b, c) Hierarchical clustering analysis of Bcl11b-repressed (b), and Bcl11b-dependent (c) genes

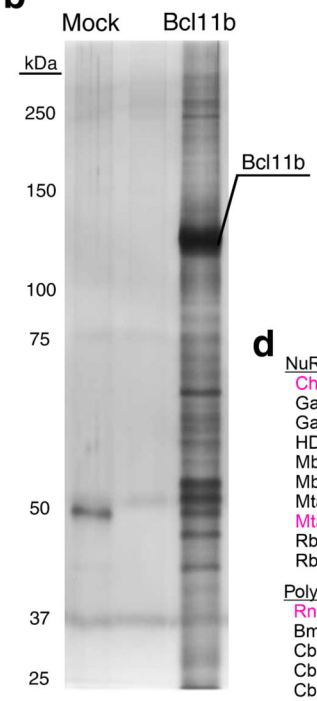
Author Manuscript Author Manuscript Author Manuscript Author Manuscript

differentially expressed in both *Vav1*-iCre and *Lck*-Cre deletions of *Bcl11b*. *Vav1*-iCre DEG are identified by comparing the average of *Vav1-iCre;Bcl11b*$^{+/+}$ (WT), and *Vav1-iCre;Bcl11b*$^{fl/+}$ (HE) together to *Vav1-iCre;Bcl11b*$^{fl/fl}$ (KO). *Lck*-Cre DEG are identified by comparing the average of DN3 WT to *Lck-Cre;Bcl11b*$^{fl/fl}$ KO cells. Color scale shows fold change relative to average of WT DN2 samples. For gene names, see Supplementary Table 1.

(d, e), Identification of subsets of Bcl11b DEGs that are expressed at lower (d) or higher (e) levels when *Bcl11b* is deleted with *Vav1*-iCre than when it is deleted with *Lck*-Cre. Hierarchical clustering analysis shows expression differences of significantly differentially expressed genes ($|Log_2FC|>1$, FDR<0.05) between *Vav1*-iCre-deleted and *Lck*-Cre deleted *Bcl11b*-deficient CD25^{+} thymocytes, with color scale representing fold change relative to average for *Lck*-Cre KO samples.

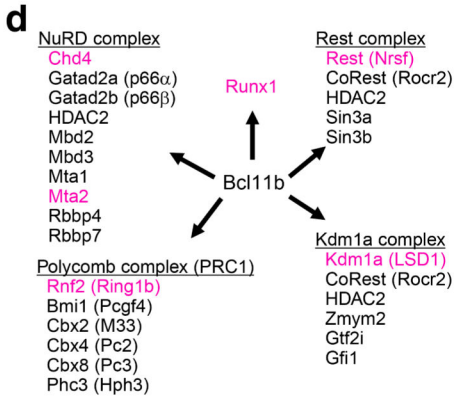
Author Manuscript Author Manuscript Author Manuscript Author Manuscript





c

GO term	Fold enrichment	*P*-value	FDR
Regulation of transcription	3.536086051	1.10E-53	1.77E-50
Transcription	3.863131883	5.67E-49	9.08E-46
Chromatin modification	10.03222907	3.99E-32	6.39E-29
Chromosome organization	7.007013201	2.31E-30	3.71E-27
Chromatin organization	7.67960558	1.52E-27	2.44E-24
Negative regulation of nucleic acid metabolic process	6.611976185	7.11E-27	1.14E-23
Negative regulation of nitrogen compound metabolic process	6.54602131	1.14E-26	1.82E-23
Negative regulation of macromolecule metabolic process	5.696251048	2.49E-26	3.99E-23
Negative regulation of gene expression	6.40232816	3.20E-26	5.13E-23
Negative regulation of transcription	6.641251222	2.52E-25	4.03E-22



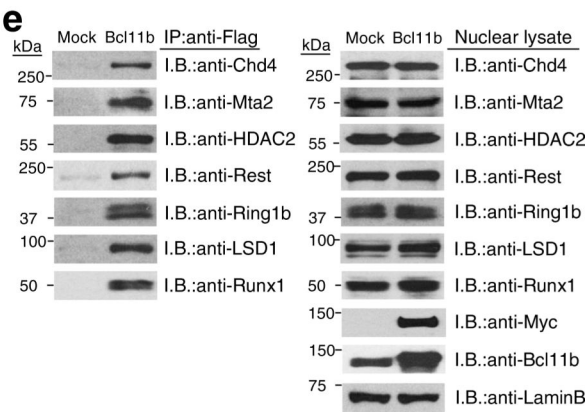


Figure 2: Identification of Bcl11b interacting molecules in early T cells

(a), Schematic representation of phenotype of *Bcl11b*-deficient pro-T cells. *Bcl11b* is turned on during T-lineage commitment. *Bcl11b*-deficient cells retain a DN2a-like phenotype, when the *Bcl11b* gene is deleted either before or after commitment. Scid.adh.2c2 cell is a DN3-like cell line.

(b), Identification of Bcl11b complexes. Total extracts from Myc-Flag-Bcl11b-expressing Scid.adh.2c2 cells and mock-transduced controls were subjected to two-step affinity purification followed by SDS-PAGE and silver staining. All of the visible bands from both samples were subjected to liquid chromatography-tandem mass spectrometry (LC-MS/MS) analysis.

(c), Bcl11b associates with molecules involving positive and negative regulation of transcription. Gene ontology (GO) analysis for Bcl11b interacting molecules was performed. Top ten GO terms are shown.

(d), Schematic representation of Bcl11b–repressor complexes detected by LC-MS/MS. Magenta indicates cofactors that will be discussed in detail.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

(e), Bcl11b interacts with several repressor complexes. Total extracts from Myc-Flag-Bcl11b-expressing Scid.adh.2c2 cells and mock-transduced controls were subjected to immunoprecipitation (IP) with anti-Flag mAb followed by immunoblotting (IB) (left). Nuclear lysates were also subjected to IB in parallel (right). Gels are cropped to focus on protein species migrating near the mobilities of the indicated size markers.
Data are representative of two independent experiments (b, e).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

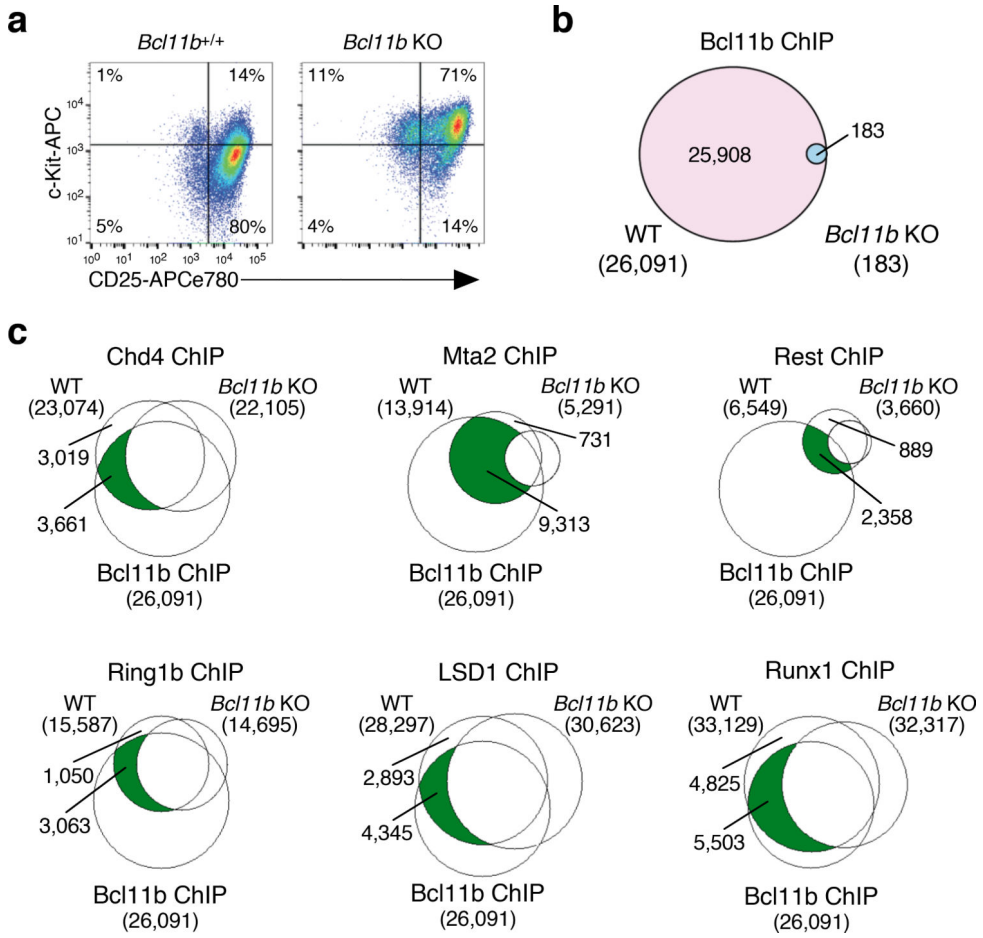


Figure 3: Identification of Bcl11b-dependent cofactor peaks in DN3 cells

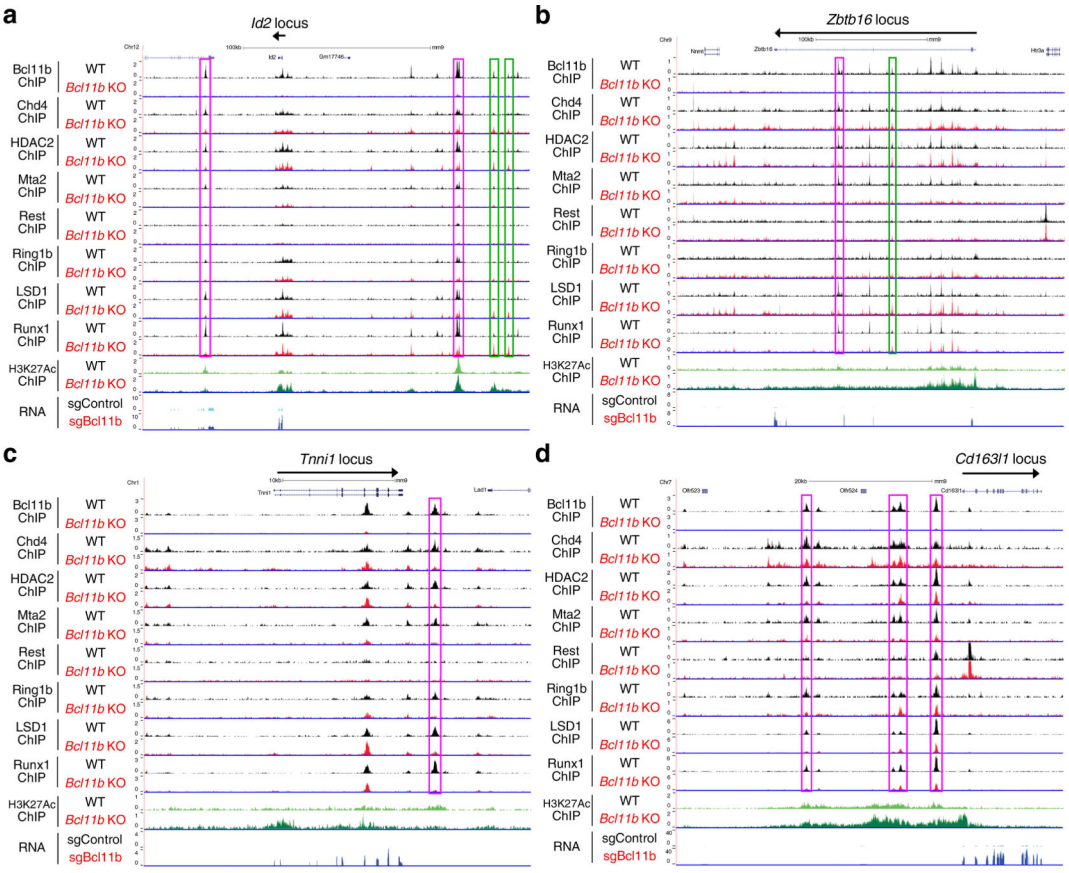
(a), Flow cytometric analysis of BM-derived precursors after 14 days of OP9-DL1 co-culture were performed. Representative c-Kit/CD25 profiles in Lin$^-$ and CD45$^+$ cells are shown with the percentages of cells in each quadrant.

(b), Number of Bcl11b ChIP peaks in DN3 cells. Bcl11b ChIP-seq analyses were performed using wildtype or *Bcl11b*-deficient cells depicted in (a). Venn diagrams show the number of ChIP peaks in each sample. In parentheses, total number of peaks in each sample is indicated.

(c), Bcl11b-dependent cofactor peaks. ChIP-seq analyses for cofactors were performed. Venn diagrams, approximately to scale, show the number of ChIP peaks in each sample with Bcl11b peaks in wildtype DN3 cells. Bcl11b-dependent cofactor peaks are indicated in green along with the number of peaks.

Data are representative of four independent experiments (a) or are based on reproducible ChIP-seq peaks in two replicate samples (b, c).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript



e Genes bound by Bcl11b-dependent cofactor peaks in DEG and non-DEG

Cofactor	Repressed genes (bound by Bcl11b) ($Log_2FC>1$, FDR<0.05, RPKM>1)	*P*-value (vs non-DEG) (Fisher's exact test)	Dependent genes (bound by Bcl11b) ($Log_2FC<-1$, FDR<0.05, RPKM>1)	*P*-value (vs non-DEG) (Fisher's exact test)	Non-DEG (bound by Bcl11b) ($|Log_2FC|<0.05$, RPKM>3)
Chd4	73 / 327 (22.3%)	5.57E-4	131 / 289 (45.3%)	2.2E-16	178 / 1374 (13.0%)
Mta2	180 / 327 (55.0%)	0.187	194 / 289 (67.1%)	1.33E-3	658 / 1374 (47.9%)
Rest	23 / 327 (7.00%)	1.35E-13	36 / 289 (12.5%)	1.51E-6	395 / 1374 (28.7%)
Ring1b	79 / 327 (24.1%)	7.69E-5	113 / 289 (39.1%)	1.62E-14	182 / 1374 (13.2%)
LSD1	83 / 327 (25.3%)	5.13E-5	151 / 289 (52.2%)	2.2E-16	189 / 1374 (13.8%)
Runx1	120 / 327 (36.6%)	3.97E-10	166 / 289 (57.4%)	2.2E-16	219 / 1374 (15.9%)
H3K27Ac	27 / 327 (8.26%)	4.80E-2	76 / 289 (26.3%)	2.2E-16	70 / 1374 (5.17%)

f Genes bound by new cofactor peaks in DEG and non-DEG

Cofactor	Repressed genes (bound by Bcl11b) ($Log_2FC>1$, FDR<0.05, RPKM>1)	*P*-value (vs non-DEG) (Fisher's exact test)	Dependent genes (bound by Bcl11b) ($Log_2FC<-1$, FDR<0.05, RPKM>1)	*P*-value (vs non-DEG) (Fisher's exact test)	Non-DEG (bound by Bcl11b) ($|Log_2FC|<0.05$, RPKM>3)
Chd4	121 / 327 (37.0%)	2.2E-16	64 / 289 (22.1%)	6.29E-5	154 / 1374 (11.2%)
Mta2	26 / 327 (7.95%)	0.011	17 / 289 (5.88%)	0.21	57 / 1374 (4.15%)
Rest	14 / 327 (4.28%)	0.015	4 / 289 (1.38%)	0.28	10 / 1374 (0.73%)
Ring1b	52 / 327 (15.9%)	2.46E-9	41 / 289 (14.2%)	5.44E-7	65 / 1374 (4.73%)
LSD1	166 / 327 (50.8%)	2.2E-16	107 / 289 (37.0%)	5.72E-7	257 / 1374 (18.7%)
Runx1	179 / 327 (54.7%)	2.2E-16	118 / 289 (40.8%)	2.61E-8	269 / 1374 (19.6%)
H3K27Ac	174 / 327 (53.2%)	2.2E-16	64 / 289 (22.1%)	0.26	255 / 1374 (18.6%)

Figure 4: Bcl11b-dependent cofactor peaks around the major Bcl11b target genes

(a–d), Binding patterns of Bcl11b, Chd4, Hdac2, Mta2, Rest, Ring1b, LSD1 and Runx1 in wildtype DN3 cells and *Bcl11b*-deficient $CD25^+$ cells, with H3K27Ac ChIP-seq and RNA-seq tracks of control and *Bcl11b*-deficient cells. Data are representative of two independent ChIP experiments. The *Id2* (a), *Zbtb16* (b), *Tnni1* (c) and *Cd163l1* (d) loci are shown. Magenta rectangles: Bcl11b-dependent cofactor peaks. Green rectangles: cofactor peaks specifically detected in *Bcl11b*-deficient cells.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

(e, f), Differentially expressed genes in *Bcl11b*-deficient cells are bound by Bcl11b-dependent cofactor peaks. DEGs were from 4 independent pairs of control and *Bcl11b*-deleted samples; $|\log_2 FC|>1$, FDR < 0.05, RPKM >1 (Supplementary Table 2). Numbers of Bcl11b-repressed, Bcl11b-dependent and non-DEGs bound by Bcl11b-dependent cofactor and H3K27Ac peaks (e), or newly generated cofactor and H3K27Ac peaks in *Bcl11b*-deficient cells (f) are shown. *P* values were determined by two-sided Fisher's exact test. Calculations were based on ChIP-seq peaks scored as reproducible in two independent replicate samples.

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

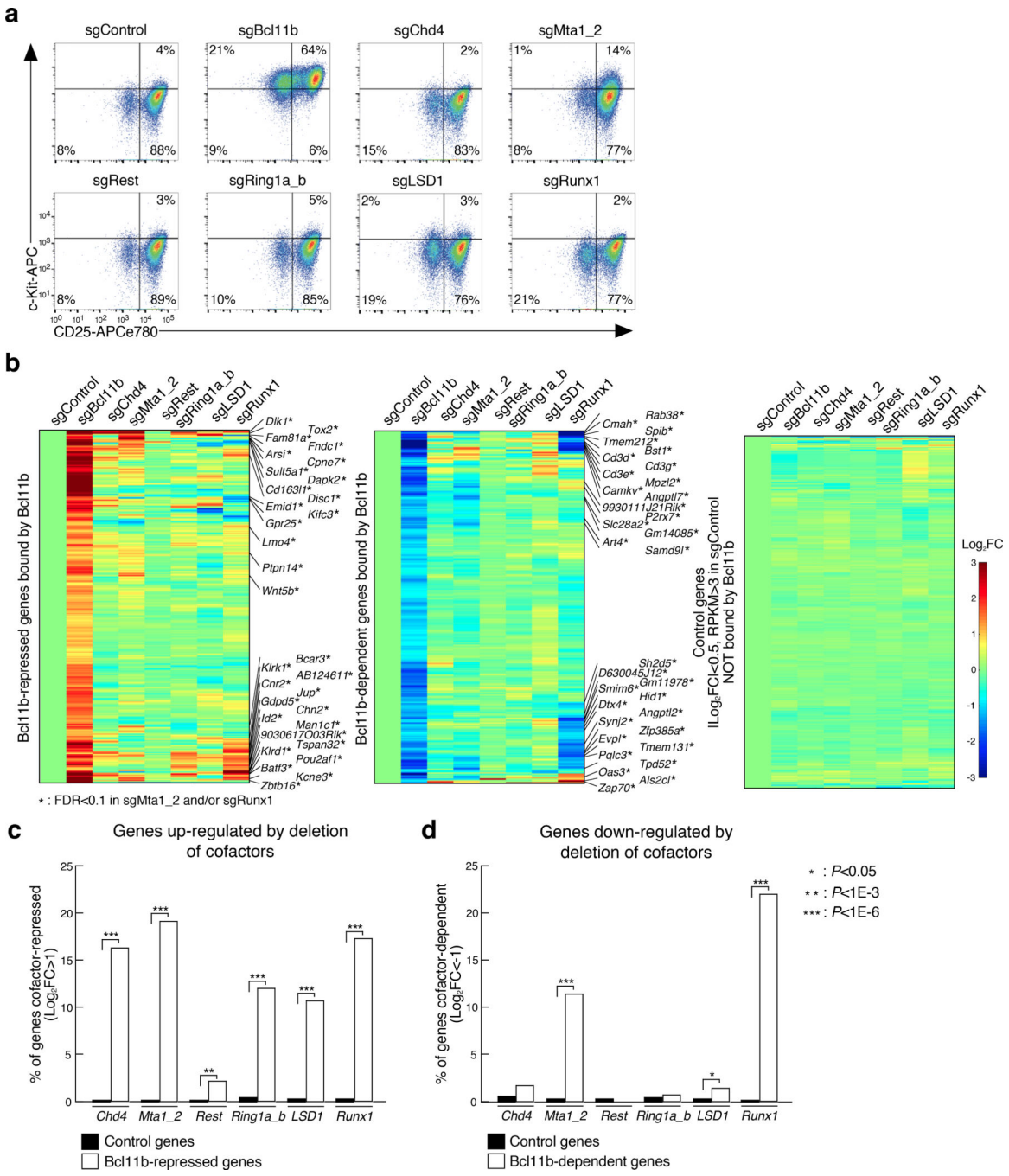


Figure 5: Effect of cofactor deletion on the expression of Bcl11b target genes

(a), CRISPR/Cas9-mediated deletion of *Bcl11b*, *Chd4*, *Mta1_2*, *Rest*, *Ring1a_b*, *LSD1* or *Runx1* gene in primary DN cells. Flow cytometric analysis of sgRNA transduced BM-derived precursors from Cas9 mice after 14 days of OP9-DL1 co-culture in $Lin^{-}CD45^{+}CFP^{+}$ cells are shown.

(b), Distinct repressor complexes coordinately control Bcl11b-mediated gene regulation. Heatmaps show hierarchical clustering analyses of the expression of the Bcl11b-repressed (left) and Bcl11b-dependent (middle) genes, which are directly bound by Bcl11b

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

(Supplementary Fig. 4a), in *Bcl11b*- or cofactor-deficient DN cells. Expression in these samples of control genes, which are not bound by Bcl11b and do not change expression upon *Bcl11b* deletion, is also shown for comparison (right). Names are indicated for Bcl11b-regulated genes that also show significant change (FDR<0.1) in response to disruption of Mta1_2 or Runx1 alone.

(c, d), Deletion of cofactors has strong effect on gene expression for Bcl11b-regulated genes, but minimal effect for genes expressed but not regulated by Bcl11b in DN3 cells. The percentages of cofactor-dependent genes ($|Log_2FC|>1$) among the Bcl11b-repressed (Fig. 5b, left) and control (Fig. 5b, right) genes (c), and Bcl11b-dependent (Fig. 5b, middle) and control (Fig. 5b, right) genes (d) are shown, tallied from Supplementary Table 6. *P* values are determined by two-sided Fisher's exact test.

Data are based on reproducible ChIP-seq peaks in two replicate samples and two replicate RNA-seq results (b, c, d), or representative of three independent experiments (a).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

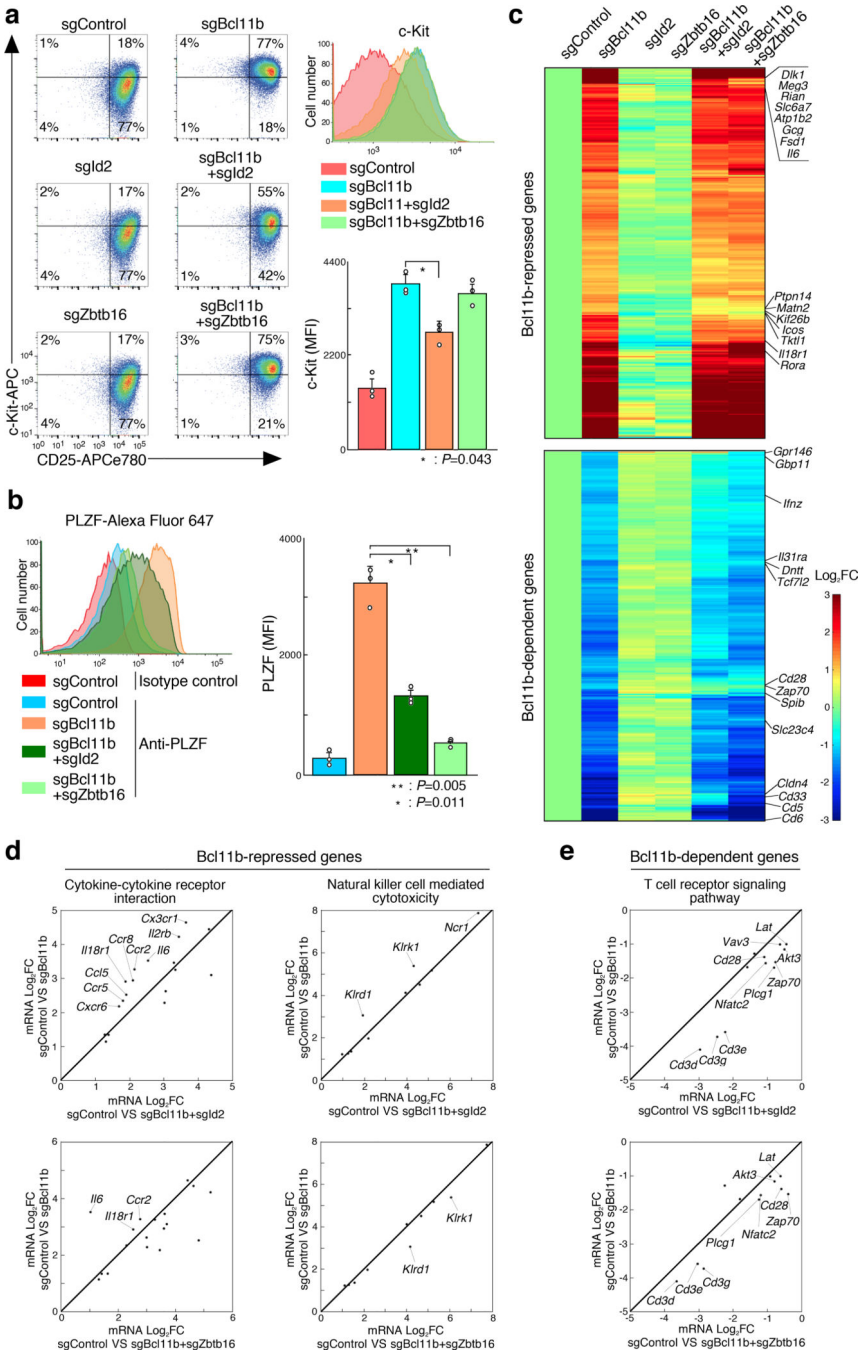


Figure 6: Id2 is involved in establishment of phenotypes of *Bcl11b*-deficient cells

(a), Up-regulation of c-Kit expression in *Bcl11b*-deficient cells is partially dependent on Id2. Flow cytometric analysis of sgRNA transduced BM-derived precursors co-cultured with OP9-DL1 (Supplementary Fig. 3a) was performed. Representative c-Kit/CD25 profiles in $Lin^{-}CD45^{+}CFP^{+}$ cells are shown with the percentages of cells in each quadrant. A representative histogram and a summary plot of the mean fluorescent intensity (MFI) of c-Kit expression of each group is also shown. $*P < 0.05$ (two-sided Student's t-test).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

(b), De-repression of *Zbtb16* expression is partially dependent on Id2. PLZF protein expression was determined by intracellular staining. A representative histogram and a summary plot of the MFIs of PLZF expression is also shown. $^{**}P<0.01$, $^{*}P<0.05$ by two-sided Student's t-test.

(c), Expression levels of a subset of Bcl11b-repressed and -dependent genes regulated by Id2 and Zbtb16. $CD25^{+}$ sgRNA transduced BM-derived precursors in (a) were purified and subjected to RNA-seq analysis. Hierarchical clustering analyses of the expression of the Bcl11b-repressed (upper) and -dependent (lower) genes in *Bcl11b*, *Id2*, *Zbtb16*, *Bcl11b* and *Id2* or *Bcl11b* and *Zbtb16*-deficient DN cells are shown.

(d), Id2 is involved in Bcl11b-mediated repression of 'Cytokine-cytokine receptor interaction' pathway. In KEGG pathway analysis, 'Cytokine-cytokine receptor interaction' (*P* value: 1.15E-05, n=18) and 'Natural killer cell mediated cytotoxicity' (*P* value: 3.09E-04, n=10) pathways were enriched among Bcl11b-repressed genes (n=410). Genes in 'Cytokine-cytokine receptor interaction' and 'Natural killer cell mediated cytotoxicity' pathways are shown. Dots represent the mRNA log2 FC values for sgControl vs. sgBcl11b (x-axis) and sgControl vs. sgBcl11b+sgId2 (left) or sgControl vs. sgBcl11b+sgZbtb16 (right) (y-axis).

(e), Id2 and Zbtb16 are involved in Bcl11b-mediated activation of 'T cell receptor signaling' pathway. 'T cell receptor signaling' pathway (*P* value; 5.32E-07, n=12) was most enriched among Bcl11b-dependent genes (n=384). Genes in 'T cell receptor signaling' pathway were shown. Dots represent the mRNA $\log_2$FC values for sgControl vs. sgBcl11b (x-axis) and sgControl vs. sgBcl11b+sgId2 (left) or Control vs. sgBcl11b+sgZbtb16 (right) (y-axis).

Data are representative of three independent experiments (a, b; three biological replicates with mean ±s.d.) or are pooled from two experiments (c, d, e)

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

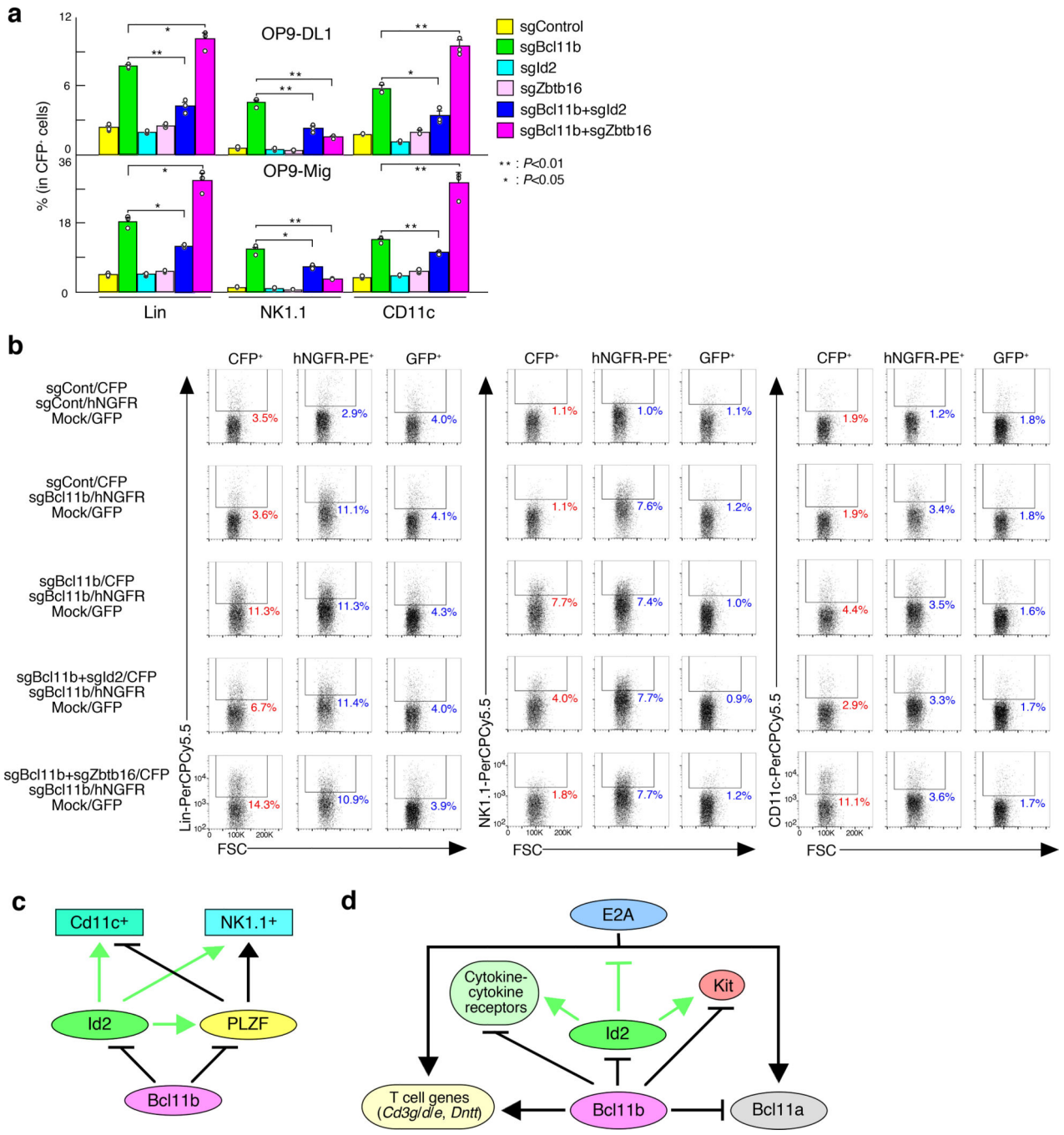


Figure 7: Id2 and Zbtb16 play key roles in Bcl11b-mediated exclusion of alternative fates

(a), Id2 and Zbtb16 are involved in alternative fate exclusion by Bcl11b. The percentages of Lin^+, $NK1.1^+$ and $CD11c^+$ in CFP^+ sgRNA transduced cells (Supplementary Fig. 11d) are indicated. $^{**}P<0.01$, $^{*}P<0.05$ by two-sided Student's t-test.

(b), Generation of alternative lineages in *Bcl11b*-deficient cells is cell-intrinsic. BM-derived precursors were cultured on OP9-DL1 for 7 days, then they were split to three aliquots. One aliquot was transduced with sgRNA against Bcl11b alone (hNGFR marker); another with the indicated combination of sgRNAs and CFP marker; and the third with pMxs-GFP empty

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

vector, separately. Three days after infection, they were pooled as shown, transferred to OP9-Mig and cultured 4 more days. Representative profiles of Forward Scatter (FSC) vs. Lin, NK1.1 or CD11c in CFP^+, $hNGFR^+$ or GFP^+ cells are shown with the percentages of cells in rectangles.

(c, d), Summary of important roles of Id2 and PLZF (encoded by *Zbtb16*) in establishment of phenotypes of *Bcl11b*-deficient pro-T cells. (c) Effects on alternative lineage markers. (d) Effects on specific genes. Bcl11b directly represses *Id2*, *Zbtb16* and *c-Kit* expression, and Id2 is partially involved in Bcl11b-mediated down-regulation of *Zbtb16* and *c-Kit* expression. Up-regulated *Zbtb16* in *Bcl11b*-deficient cells supports generation of $NK1.1^+$ cells and represses $CD11c^+$ cells. On the other hand, *Id2* positively contributes to generation of $NK1.1^+$ and $CD11c^+$ cells (c), activation of genes in 'Cytokine-cytokine receptor interaction' pathway and antagonism of E-protein-mediated activation of T cell genes (d). The rare Bcl11b repression of an E-protein dependent gene, *Bcl11a*, shows gene specificity of effects.

Data are individual values of three biological replicates with mean $\pm$s.d. (a), or representative of three independent experiments (b).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 1:

HIGHEST CONFIDENCE BCL11B-REGULATED TARGET GENES

A. Genes upregulated in Bcl11b knockout DN2/3 cells: Bcl11b repression targets					
1700024P16Rik	*1700112E06Rik*	*2900026A02Rik*	*41340*	*5730559C18Rik*	*9030617O03Rik*
9130019P16Rik	*Aatk*	*AB124611*	*Acot11*	*Adamtsl3*	*Aifm2*
Ak4	*Akap6*	*Aoah*	*Apobr*	*Appl2*	*Arap3*
Arhgap20	*Arl10*	*Arl6*	*Arsi*	*Art2a-ps*	*Art2b*
Arvcf	*Asph*	*Bace1*	*Batf3*	*Bcar3*	*Bcas1os2*
Calm1	*Camk2b*	*Camkmt*	*Card11*	*Cass4*	*Ccdc157*
Ccr2	*Cd160*	*Cd163l1*	*Cd244*	*Cd63*	*Cd7*
Cd72	*Cd9*	*Ceacam1*	*Cers4*	*Chad*	*Chn2*
Chpf	*Cisd3*	*Cited4*	*Cldn2*	*Clnk*	*Clvs1*
Cnbd2	*Cnksr3*	*Cnnm2*	*Cnr2*	*Col9a3*	*Colq*
Coprs	*Cpd*	*Cpeb2*	*Cpne7*	*Cpt1a*	*Ctbp2*
Cx3cr1	*Cxcr5*	*Dab2ip*	*Dapl1*	*Dennd3*	*Disc1*
Dlx1	*Dlx1as*	*Dok3*	*Drc7*	*Dscam*	*Dsp*
Dyrk3	*Eci1*	*Eea1*	*Egln3*	*Fam129a*	*Fam151a*
Fam179a	*Fam46a*	*Fam71b*	*Fbxw8*	*Fcer1g*	*Fdx1*
Fes	*Fgf3*	*Flna*	*Flt3*	*Frmd4b*	*Gas7*
Gimap4	*Gimap6*	*Gimap7*	*Gimap8*	*Glis2*	*Golm1*
Gpr141	*Gpr183*	*Gstm1*	*Gstm3*	*Gzmc*	*Hey2*
Heyl	*Hip1r*	*Ica1l*	*Id2*	*Il2rb*	*Iqgap2*
Irak3	*Itgb6*	*Itgb7*	*Ivns1abp*	*Kank2*	*Kcne3*
Kif13a	*Kif5a*	*Kifc3*	*c-Kit*	*Klhdc2*	*Klrd1*
Lag3	*Layn*	*Ltk*	*Lyn*	*Man1c1*	*Micall1*
Myo1e	*Myo1f*	*Myo7a*	*Nat8l*	*Nav2*	*Nectin1*
Neurl3	*Nfil3*	*Npffr1*	*Nr2f6*	*Nrgn*	*Nt5e*
Osbpl5	*Oscp1*	*Papss2*	*Pcyox1l*	*Pde2a*	*Pde4a*
Pear1	*Pik3r2*	*Pkig*	*Plekhg5*	*Polm*	*Pou2af1*
Prtg	*Ptpn14*	*Ptpn21*	*Rab19*	*Rapgef2*	*Rassf4*
Rgs3	*Rhobtb1*	*Rnh1*	*Rora*	*S100a10*	*S1pr3*
Sccpdh	*Scn5a*	*Sema3c*	*Sema4c*	*Serpinb9*	*Sh2d1b1*
Sh3bgrl2	*Siae*	*Slc22a15*	*Slc22a23*	*Slc2a6*	*Slc35f5*
Slc45a3	*Spa17*	*Spaca9*	*St6galnac6*	*Sult5a1*	*Tas1r1*
Tgfb3	*Tiam2*	*Timp2*	*Tirap*	*Tjp2*	*Tmem126a*
Tmem198	*Tmem231*	*Tmem67*	*Tnfrsf25*	*Tnni1*	*Trf*
Trim2	*Trnp1*	*Trpm1*	*Tspan32*	*Tyrobp*	*Ublcp1*
Vsig2	*Wipi1*	*Wnt5b*	*Xcl1*	*Zbtb16*	*Zbtb7b*
Zcchc18	*Zfp105*	*Zfp296*	*Zfp316*	*Zfp518b*	*Zfp568*
Zfp768					

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

A. Genes upregulated in Bcl11b knockout DN2/3 cells: Bcl11b repression targets					
B. Genes downregulated in Bcl11b knockout DN2/3 cells: Bcl11b-dependent genes					
1300002E11Rik	*Abtb2*	*Actn1*	*Acy3*	*Bbof1*	*BC025920*
Bst1	*Camkv*	*Ccdc153*	*Ccdc18*	*Cd3d*	*Cd3g*
Cd5	*Cd6*	*Cldn4*	*Comp*	*Dapk1*	*Dcp1b*
Def8	*Dgke*	*Dgkeos*	*Dgkg*	*Ehd3*	*Emp1*
Evpl	*F730043M19Rik*	*Fam117a*	*Fer1l5*	*Frmd4a*	*Frmd6*
Gbp4	*Gbp8*	*Glyctk*	*Gm15708*	*Gm26839*	*Gmpr*
Grasp	*Hid1*	*Hmgcs2*	*Id3*	*Ifngr2*	*Il10ra*
Il21r	*Kif7*	*Llgl1*	*Map4k2*	*Matk*	*Mpzl2*
Mzb1	*Nfatc2*	*Nfkb2*	*Pisd-ps1*	*Pisd-ps2*	*Pisd-ps3*
Plcg1	*Plekha7*	*Plxnd1*	*Pqlc3*	*Rab11fip4*	*Rasal1*
Rgs10	*Rhbdf2*	*Sestd1*	*Sirt5*	*Slc37a1*	*Slc5a9*
Smim5	*Smox*	*Smyd2*	*Sox13*	*Spib*	*Steap3*
Synj2	*Tbxa2r*	*Tmem131*	*Tmem221*	*Tmprss4*	*Tnfrsf1a*
Ttc38	*Zeb2os*				

Table lists genes that show significant differential expression from wildtype controls in all three of the following conditions: in *Lck*-Cre deletion of *Bcl11b in vivo*, in *Vav1*-iCre deletion of *Bcl11b in vivo*, and in Cas9-mediated deletion *Bcl11b* in DN2b/3 cells differentiating *in vitro* from bone-marrow derived precursors. See Supplementary Table 3 for complete lists of genes identified in each of the individual comparisons. Genes are listed in alphabetical order; see Supplementary Tables 1,2 for expression values.