6-6-2017

Bcl11b and combinatorial resolution of cell fate in the T-cell gene regulatory network.

William J R Longabaugh
Institute for Systems Biology, Seattle, WA 98109

Weihua Zeng

Jingli A Zhang

Hiroyuki Hosokawa

Camden S Jansen

See next page for additional authors

Follow this and additional works at: https://digitalcommons.psjhealth.org/publications

Part of the Genetics and Genomics Commons

Recommended Citation
https://digitalcommons.psjhealth.org/publications/482

This Article is brought to you for free and open access by Providence St. Joseph Health Digital Commons. It has been accepted for inclusion in Journal Articles and Abstracts by an authorized administrator of Providence St. Joseph Health Digital Commons. For more information, please contact digitalcommons@providence.org.
Bcl11b and combinatorial resolution of cell fate in the T-cell gene regulatory network

William J. R. Longabaugh1,2, Weihua Zeng1, Jingli A. Zhang1,3, Hiroyuki Hosokawa4, Camden S. Jansen5, Long Li4,5, Maile Romero-Wolf2, Pentao Liu6, Hao Yuan Kueh5,6, Ali Mortazavi7,8, and Ellen V. Rothenberg1,2

1Institute for Systems Biology, Seattle, WA 98109; 2Department of Developmental and Cell Biology, University of California, Irvine, CA 92697; 3Division of Biology & Biological Engineering, California Institute of Technology, Pasadena, CA 91125; and 4Wellcome Trust Medical Research Council, Cambridge Stem Cell Institute, University of Cambridge, Cambridge CB2 1QR, United Kingdom

Edited by Neil H. Shubin, The University of Chicago, Chicago, IL, and approved January 30, 2017 (received for review October 25, 2016)

T-cell development from hematopoietic progenitors depends on multiple transcription factors, mobilized and modulated by intrathymic Notch signaling. Key aspects of T-cell specification network architecture have been illuminated through recent reports defining roles of transcription factors PU.1, GATA-3, and E2A, their interactions with Notch signaling, and roles of Runx1, TCF-1, and Hes1, providing bases for a comprehensively updated model of the T-cell specification gene regulatory network presented herein. However, the role of lineage commitment factor Bcl11b has been unclear. We use self-organizing maps on 63 RNA-seq datasets from normal and perturbed T-cell development to identify functional targets of Bcl11b during commitment and relate them to other regulators. We show that both activation and repression target genes can be bound by Bcl11b in vivo, and that Bcl11b effects overlap with E2A-dependent effects. The newly clarified role of Bcl11b distinguishes discrete components of commitment, resolving how innate lymphoid, myeloid, and dendritic, and B-cell fate alternatives are excluded by different mechanisms.

Bcl11b | Notch-delta signaling | PU.1 | E2A | commitment

Lymphocyte development from blood stem cells depends on the regulatory inputs from a suite of transcription factors needed by all T-cell types, as well as a set of factors that subdivide mature T cells into different functional groups. Before functional subdivision, a common core gene regulatory network (GRN) guides multipotent precursors to generate committed pro-T cells that will ultimately serve as progenitors for all T-cell lineages. Commitment occurs in the thymus before the cells acquire specific T-cell receptors (TCR) for antigen, and is driven by a combination of extrinsic signals and intrinsic transcription factor activity changes. This early T-cell GRN not only turns on T-cell-specific genes but also coordinates early precursor proliferation with stepwise renunciation of alternative developmental potentials. We and others have found that the commitment process concludes with the onset of expression of a zinc finger transcription factor, Bcl11b (1–4). Here we identify the genes that are immediately sensitive to Bcl11b activity, and integrate the Bcl11b activation process into the context of the broader T-cell specification GRN.

The thymic environment instructively promotes T-cell differentiation of multipotent immigrant cells by presenting Notch ligand Delta-like 4 (DLL4), and providing supportive cytokines (Kit ligand, IL-7). The resulting Notch signaling drives precursors to proliferate through a canonical series of stages [double negative (DN) = CD4−CD8−; double positive (DP) = CD4+CD8+] from Kit+ DN1 (or early T-cell precursor, ETP), to DN2a, DN2b, and DN3a stage (reviewed in ref. 5). If the precursors can begin TCR expression successfully in DN3a, they continue through DN3b and DN4 to DP, when the cells finally acquire complete TCR recognition complexes (5, 6). Importantly, if individual T-cell precursors in the ETP or DN2a stage are removed from thymic Notch signaling complexes (5, 6), they can no longer do this unless genetically manipulated (1, 7). This transition defines “commitment.”

The robust change in potential from DN2a to DN2b is also accompanied by dynamic transcription factor expression changes (8, 9). At least 20 regulatory genes have expression patterns that can be classed as “phase 1” (expressed in ETP and DN2a, then down-regulated) or “phase 2” (turned on or significantly up-regulated around commitment in DN2b) (3). To date, the most-studied regulators of the phase 1 to phase 2 transition have been Notch signaling, GATA-3, TCF-1, and E2A, and PU.1 as a natural, endogenous “opponent” of developmental progression. Of these, all are present in ETP and DN2a cells, and only Bcl11b is up-regulated de novo during this commitment transition itself.

The activity of Bcl11b in commitment has had two aspects poorly resolved to date. After commitment, Bcl11b seems to repress specific types of effector programs in T cells and innate lymphoid cells (ILC), blocking genes associated with natural killer (NK)-like or effector cytolytic T cells or with differentiation to specific IL-17-producing subsets of invariant NK T cells (reviewed in refs. 6 and 10, 11). This effector-subtype blockade is important to support memory CD8 cell function and regulatory T-cell differentiation, but it does not explain why Bcl11b is important for T lineage commitment as a whole. Cells losing Bcl11b acutely long after commitment activate an NK-cell like gene-expression profile (2, 12) or a precociously specialized T-cell profile (13), but not an immature gene-expression profile like the one extinguished during commitment.

To resolve how Bcl11b is working, we have focused specifically on the stages at the fulcrum of commitment to carry out a genome-wide RNA-seq analysis of the effects of Bcl11b deletion. We have identified the most strongly influenced target genes in a

This paper results from the Arthur M. Sackler Colloquium of the National Academy of Sciences, “Gene Regulatory Networks and Network Models in Development and Evolution,” held April 12–14, 2016, at the Arnold and Mabel Beckman Center of the National Academies of Sciences and Engineering in Irvine, CA. The complete program and video recordings of most presentations are available on the NAS website at www.nasonline.org/Gene_Regulatory_Networks.

Data deposition: The data reported in this paper have been deposited in the Gene Expression Omnibus (GEO) database, www.ncbi.nlm.nih.gov/geo (RNA-seq data accession no. GSE89198; ChIP-seq data accession no. GSE53572).

1W.J.R.L. and W.Z. contributed equally to this work.
2To whom correspondence may be addressed. E-mail: wlongabaughystemsbiology.org, ali.mortazavi@uci.edu, or evroth@its.caltech.edu.
3Present address: Department of Molecular Biology, Genentech, South San Francisco, CA 94080.
4Present address: Department of Cell Biology, Tianjin Medical University, 300070 Tianjin, China.
5Present address: Department of Immunology, Tianjin Medical University, 300070 Tianjin, China.
6Present address: Department of Bioengineering, University of Washington, Seattle, WA 98105.

This article contains supporting information online at www.pnas.orglookup/suppl/doi:10.1073/pnas.1610617114/DCSupplemental.
de novo analysis of 63 RNA-seq datasets of wild-type as well as perturbed developing T cells. We have resolved distinct subsets of Bcl11b targets as they relate to the regulomes of other developmentally important transcription factors in T-cell development, using self-organizing maps (SOM) (reviewed in ref. 14). These SOM have long been used for analysis of gene-expression (15, 16) and chromatin data (17) because they can resolve relationships in very large datasets with large numbers of clusters, and here they provide an elegant way to identify multiple, related modules of Bcl11b targets with distinct roles in development. In fact, many of the new Bcl11b target genes are unique to the peri-commitment developmental context. We place these findings in the context of an updated version of the T-cell specification GRN model (5, 18–20) enhanced with recent insights into connections between PU.1, Bcl11b, Notch signaling, and other crucial factors: that is, GATA-3, TCF-1, Runx1, and the basic helix-loop–helix factor E2A.

Results

T-Cell Specification GRN Transitions at Commitment: A Distinct Context for Bcl11b. A model for the GRN that guides T-cell commitment has been assembled over the last decade (18, 19) but new linkages that control aspects of commitment timing have now been illuminated. Recent reports have identified key mechanisms that make the transition from precommitment to commitment developmental context. We place these findings in the revised and updated GRN model shown in Fig. 1. The precommitment phase 1 period is defined by strong expression of the B-cell–, dendritic-cell–, and myeloid-cell–associated transcription factor PU.1 (encoded by Spi1, alias Spil1), and stem-cell–associated regulatory genes that are activated prethymically, including Erg, Myn, Hhex, Bcl11a, Gfi1b, Mef2c, Lmo2, and Lyl1 (reviewed in ref. 5) (Fig. LA). The most important growth factor receptor sustaining this phase is encoded by Kit. These genes all need to be down-regulated during "by CreERT2 activation Lyl1 Cxcr5 Gata3 from DN2b or DN3 cells Lmo2 Nfil3 was deleted before up-regulation depends on a stringently combinatorial Bcl11a was Sup (encoding TCF-1) Hhex cells spanning the DN2a phase during the DN1 phase, before actual tran-

eloid program via Hes1 against Myb, and GATA-3 (21, 23, 24), on Notch blockade of the my-

eloid genes, including Gata3 and Tcf7 (encoding TCF-1) starting in the DN1 (ETP) stage, but these are not sufficient to impose commitment. In contrast, Bcl11b, a phase 2 regulatory gene, is transcriptionally silent until the late DN2a stage, and then turns on abruptly. It is then expressed throughout the rest of T-cell development and in mature T cells thereafter (Fig. IB). Both silencing of PU.1 and activation of Bcl11b are operationally important for commitment (2–5, 21).

This updated model incorporates recent data on functional interactions E2A (22), on PU.1 and its interactions with Notch, Myb, and GATA-3 (21, 23, 24), on Notch blockade of the myeloid program via Hes1 against Cebpa (25, 26), on GATA-3 and its mutual antagonism with the B-cell program (27–30), and on positive regulators of Bcl11b itself (1, 31). Detailed in discussion in SI Appendix, Supplementary Text, these results reveal two major switch circuits. First, a mutual inhibition circuit between PU.1 and Notch, without mutual repression, buffers the intensity of Notch signaling but restricts the lineage-diversification activity of PU.1 if Notch signaling is present, while letting PU.1 support genes linked to multipotentiality (21, 23) (Fig. L1). This balance defines the phase 1 pro–T-cell state, which is surprisingly per-

sistent in pro-T cells through days of proliferation. Second, Bcl11b up-regulation depends on a stringently combinatorial action of GATA-3, TCF-1 (encoded by Tcf7), Runx1, and Notch signaling. Here, the most collaborative events appear necessary to “prime” the locus during the DN1 phase, before actual tran-

scriptional activation in late DN2a phase (1). New results on E2A targets are considered further below. Detailed review of the operation of the network as a whole and evidence for specific links are provided in SI Appendix, Supplementary Text.

The commitment event is tightly coupled with Bcl11b up-regulation (Fig. 2A) (1), but the exact role of Bcl11b in this process depends on the unique regulatory context, which is not duplicated in any of the later T-cell contexts where Bcl11b works (reviewed in ref. 10) nor in type 2 ILC, where it also has a role (11, 32, 33). Until commitment (Fig. L4), endogenous PU.1, possibly other phase 1 genes, and even GATA-3 still represent regulatory bridges to distinct alternative programs. Notch sig-

naling selectively intervenes in their actions to prevent alternative lineage specification (21, 26, 34, 35), but still allows PU.1 to sustain lymphoid-compatible phase 1 regulatory genes like Bcl11a, Lmo2, and Mef2c (21) in phase 1 pro-T cells. Gene regu-

lation at this stage can also reflect balances between different factors competing to regulate common targets: for example, PU.1 and GATA-3 may compete, respectively, to repress or activate Zpml1, and Ets1, and to activate or repress Bcl11a (21, 23, 27). This balance is tipped during commitment, with the silencing of the phase 1 regulatory genes and down-regulation of Kit, but only limited aspects of this process are understood. GATA-3, Runx1, and TCF-1 (or its relative LEF-1) eventually play roles in silencing expression of phase 1 regulatory genes encoding PU.1 and Bcl11a during commitment, as demonstrated by gain- and loss-of-

function data (27, 36–39) (Fig. 1B). However, all of these factors are present at comparable levels for multiple cell cycles before commitment (examples shown in Fig. 2B). Thus, although the precommitment and postcommitment states are clearly distin-

guished from each other, the exact role of rate-limiting compo-

nents like Bcl11b in triggering the transition between them has not been fully defined.

Determination of Bcl11b Targets During Commitment. Cells failing to turn on Bcl11b during commitment are arrested in vivo (13, 40) but proliferate well in vitro with strong Notch and cytokine signals. In key respects the proliferating cells resemble DN2a cells (3, 4) (DN2a* in Fig. 2A). They tend to differentiate to NK-like cells if Notch signal intensity is reduced (2, 3). One prediction is that Bcl11b should itself repress regulatory genes like Spi1 (PU.1) or Lyl1 that are sharply down-regulated in committed (Fig. 2B). Another prediction is that, if Bcl11b is important to block access to the NK cell fate, then the NK-promoting genes it promotes turn on only in late T-cell precursors until the time that Bcl11b is turned on. However, the actual results differ from these predictions.

Conditional knockout (KO) Bcl11b DN2 cells and control Bcl11b+ cells spanning the DN2a–DN3a interval plus two control DP samples were generated for RNA-seq comparisons (Fig. 2C) using two protocols (Fig. 2 D1 and D2 and SI Appendix, Supplementary Methods). In most samples, Bcl11b was deleted before its normal onset of expression, by introducing Cre into condi-

tional KO hematopoietic progenitor cells or wild-type controls before T-cell development (protocol 1) (Fig. 2D1). We used fetal liver-derived hematopoietic precursors and cultured control and conditional KO cells in parallel in a well-characterized in vitro T-cell differentiation system, with cytokines and OP9-DL1 or OP9-DL4 stromal cells (3, 41) (SI Appendix, Fig. SL4). For protocol II (Fig. 2D2), we deleted Bcl11b from DN2b or DN3 cells after commitment, usually deleting Bcl11b by CreERT2 activation in vitro (2) in freshly isolated DN3 thymocytes. One other protocol II sample pair used Lck-Cre deletion in vivo (SI Appendix, Supplementary Methods). Both protocols generated characteristically large, DN2a-like Bcl11b KO cells with high expression of Kit (SI Appendix, Fig. SL4, days 9 and 12 and B, day 4). However, fetal-

derived cells that had lost Bcl11b before commitment could pro-

derive extensively with high viability in OP9-DL4 coculture (3), whereas adult-derived cells that had lost Bcl11b after commitment grew poorly in the same conditions.

For RNA-seq analysis, multiple matched control vs. KO pairs, additional Bcl11b KO samples, and additional controls were sorted from in vitro cultures at corresponding developmental stages, as well as three sets of adult samples with Bcl11b deleted after commitment (Fig. 2C). Only cells with a DN2 or DN3 phenotype were sorted (SI Appendix, Fig. SLA, Inset) to exclude any cells already transformed to NK cells. RNA-seq analysis (SI Appendix, Fig. S2) showed that the targeted exon of Bcl11b was efficiently deleted, and genes including Tmn1, Itg2b, Tyrobp, Fcεr1g, Cxcr5, Zbtb16, Nfil3, Id2, and I2rb were up-regulated in
the KO cells, whereas, e.g., Gbp4 and the Cd3 gene cluster were down-regulated. Because of the dynamic developmental context (Fig. 2D), there were variances in developmental progression overall in different experiments (SI Appendix, Fig. S3A), as shown by a panel of 90 indicator regulatory genes (SI Appendix, Fig. S3B). Thus, although some gene-expression changes were
seen in all Bcl11b KO samples, some were significant only in particular types of samples: adult vs. fetal, protocol I vs. protocol II, or where controls were more DN2b-like vs. more DN3a-like (e.g., SI Appendix, Fig. S4: Kit, Cd7, Bcl11a). However, overall, the expression changes observed were reproducible among multiple independent KO-control sample pairs (e.g., Kit, Zbtb16, Cpa3, Tmn1, Cd3g, Cd3d, Cd3e, all P < 10^-6, n = 10) (SI Appendix, Fig. S3C). To identify a “gold-standard” list of differentially expressed genes, we grouped the paired-sample comparisons into three pools and selected genes that showed significant up or down-regulation (EdgeR, false-discovery rate < 0.05) in at least two of the three separate pools (Dataset S1A; component lists in Dataset S1B and C; all genes differentially expressed in ≥one pool in Dataset S1D). Of differentially expressed transcripts with known protein-coding potential (omitting TCR loci, Gm and miR transcripts), 77 were up-regulated significantly when Bcl11b was knocked out and were thus presumably repressed by Bcl11b normally, whereas 22 lost expression and thus were Bcl11b-dependent knocked out and were thus presumably repressed by Bcl11b normally (SI Appendix, Fig. S3C). Normal expression patterns of all these genes are shown in SI Appendix, Fig. S5.

The high-confidence Bcl11b-dependent genes were more likely to increase expression through normal commitment, whereas Bcl11b-repressed genes tended to be down-regulated (SI Appendix, Fig. S5A), consistent with the expression of Bcl11b itself (Fig. 2B). Several genes associated with initial TCR assembly were notably Bcl11b-dependent in this developmental context, although this has not been seen in later stages; that is, genes coding for CD3 signaling components and for Dntt, the terminal deoxynucleotidyl transferase that helps in generating a highly diverse TCR repertoire (SI Appendix, Figs. S2 and S5C). We excluded Tcrb and Terc loci from the gold-standard gene set because of annotation complexity, but Dataset S1B suggests that Bcl11b may support these differentiation-associated transcripts as well. In contrast, genes that were repressed by Bcl11b (i.e., up-regulated in the KOs) fell into several differently regulated groups, only a subset of them phase 1-restricted. Purified DN3 cells from Bcl11bKO mice that did not express significant levels of Bcl11b-repressed targets in vivo could be induced to express these genes by acute deletion of Bcl11b in vitro, as confirmed by quantitative PCR (qPCR) analyses (SI Appendix, Fig. S5B).

Context Dependence of Bcl11b Targets. Previous studies have examined Bcl11b deletion in later-stage thymocytes, CD4+ CD8+ DP cells (2, 12, 13), where an NK-like gene signature emerged. Indeed, the Kyoto Encyclopedia of Genes and Genomes pathway most highly enriched in our Bcl11b-repressed gold-standard gene set was NK cell-mediated cytotoxicity (SI Appendix, Fig. S6A), and multiple genes were shared with gene sets that distinguish ILC from T cells (42, 43). However, the majority (53 of 77) of Bcl11b-repressed gold-standard genes from DN2/3 cells were not identified in DP cell analyses (e.g., Cd7, Cpa3, Kit, Tmn1, and Zbtb16) (compare SI Appendix, Figs. S2 and S4). Moreover, only 2 of 22 Bcl11b-dependent genes in DN2-3 cells were affected in DP cells (SI Appendix, Fig. S6C and Dataset S1F). Thus, the genes that Bcl11b must keep silent are more consistent across developmental stages than those that it may help to activate. When examining many hematopoietic cell samples from the ImmGen consortium (9, 44), we found that the gold-standard Bcl11b-dependent gene set was NK cell-mediated cytotoxicity (SI Appendix, Fig. S6A), which was expressed more strongly in NK cells than in other hematopoietic cell types (SI Appendix, Fig. S7A). Many Bcl11b-repressed genes (Tmn1, Iilr2, Zbtb16, Id2, Cxcr5) did not
simply continue expression from uncommitted phase 1 progenitors, but were either up-regulated from ETP–DN2a levels, or else activated dramatically de novo if Bcl11b were absent.

**Bcl11b Binding to Target Genes in Vivo.** ChIP and sequencing (ChIP-seq) showed clear Bcl11b binding to sites around positive and negative regulation targets in DN3 cells (Fig. 3A and SI Appendix, Fig. S8 A–G). Bcl11b specifically occupied sites at most if not all gold-standard loci, including Zbtb16, Tnnt1, Dntt, and Cdsd (ChIP-PCR in Fig. 3B). Sites were often in open chromatin at positively regulated loci but also at closed or closing sites in many repressed targets (SI Appendix, Fig. S8 A–C). Such sites are candidates for mediators of direct regulation.

**Global Picture of Bcl11b Effects in the Context of Early T-cell Development.** To relate the genes that Bcl11b controls to the global context of T-cell development, we used all of the RNAseq data in our 63 samples to generate a SOM, clustering gene-expression patterns in all our samples into a large toroidal 40 × 60-unit map. This is a format that is particularly well-suited for clustering high-dimensional datasets first into units of groups of genes that have nearly identical expression profiles among all samples (clusters), then placing similarly regulated units near each other on the map geographically in >one dimension (metaclusters). Importantly, the 2D mapping relates clusters by more than one criterion of similarity. SOM analysis thus enabled us to group the whole transciptome into 300 metacusters based on fine-grained similarities of regulation not only in development but also under perturbation (17) (Fig. 3C and SI Appendix, Supplementary Methods and Fig. S9A).

The gene-expression patterns defined a complex landscape (SI Appendix, Fig. S9A) in which changes in expression of different groups of genes were shown by coordinated increases or decreases in signal between stages from particular metacuster regions (Fig. 3 C–G and SI Appendix, Fig. S9B and Dataset S2). Phase 1-specific genes were concentrated in a subset of metacuster regions (blue in Fig. 3C and region B in SI Appendix, Fig. S9B), while later-expressed genes could be mapped to another subset of regions (red in Fig. 3C and region A in SI Appendix, Fig. S9B). Patterns of expression of the eigengenes for specific metacusters, across all our samples, are shown in SI Appendix, Fig. S10. For example, PU.1 (Spu1, Spii) is in metacuster 112 (region B, in SI Appendix, Fig. S9B), one whose expression is significantly down-regulated in DN2b and later samples (SI Appendix, Fig. S10A). Other phase 1-specific regulatory genes, such as Bcl11a, map nearby (in metacuster 212) (region B in SI Appendix, Fig. S9B and expression in SI Appendix, Fig. S10B). In contrast, the phase 2-specific genes, including Bcl11b itself (in metacuster 26) (SI Appendix, Fig. S10D), are in the upper-right cluster region up-regulated in DN2b T cells (region A in SI Appendix, Fig. S9B).

The SOM analysis clearly showed that Bcl11b deletion up-regulated expression of genes in two distinct metacuster neighborhoods with different relationships to normal development (Fig. 3D and SI Appendix, Fig. S9B, regions C1 and C2). When the DN2b vs. DN1 fold-change map is compared with the
Bcl11b KO vs. control fold-change map (Fig. 3 C and D), one region of genes enhanced in the Bcl11b KO overlapped with genes with phase 1-biased expression (Fig. 3 C and D; region C1 substantially overlapping region B in SI Appendix, Fig. S9B). These metaclusters include or adjoin the clusters containing most of the signature phase 1 regulatory genes (Dataset S2A), identifying the specific immature properties that are perpetuated in Bcl11b KO cells. However, the other salient region (region C2 in SI Appendix, Fig. S9B) consisted of genes up-regulated de novo in Bcl11b-deficient DN2-like cells. This includes genes associated with NK function and ILC/NK fate, like Nfil3 and Zbtb16 themselves (metacluster 168) (SI Appendix, Fig. S10C). Thus, through distinct sets of genes Bcl11b loss retards aspects of T-cell differentiation and promotes acquisition of a new fate.

Explaining Bcl11b Impact on Development.

No global epistasis with PU.1. Although Bcl11b KO cells perpetuated some aspects of the phase 1 state, including high Kit expression (DN2a* in Fig. 2A), few phase 1 signature transcription factor genes themselves were specifically up-regulated in Bcl11b KOs (although Bcl11a down-regulation tended to be delayed) (SI Appendix, Fig. S4). Bcl11b KO cells also seemed normally competent to express T-cell–specifying regulatory genes (Dataset S2B). In general, they often up-regulated Gata3 (metaclass 2A), while specifically activating the distal promoter isoform S9 of Runx3 (metaclass 188) (browser view shown in SI Appendix, Fig. S11). Bcl11b KO effects were not mimicked by Notch inhibition, as measured by short-term γ-secretase inhibition (SI Appendix, Figs. S2, S3) in DN2a or DN2b cells, shown by the lack of impact on similar metaclusters in this SOM analysis (SI Appendix, Fig. S12A). Therefore, we tested two other ways that Bcl11b might normally advance the T-cell gene regulatory program: by suppression of phase 1 regulator effectiveness or enhancement of E protein activity.

During commitment, the trajectories of Bcl11b and Sp1 (PU.1) expression cross, one up-regulated as the other is down-regulated (1.45) (Fig. 2B). To test whether these act as mutual antagonists, we directly compared Bcl11b KO effects with those of an acutely acting PU.1 obligate repressor in DN2a cells (PU.1–ENG), which we had used as a dominant negative to identify genes that depend on wild-type PU.1 (23). SOM analysis identified a supercluster of genes that were severely down-regulated by the obligate PU.1 repressor (Fig. S5, lower right; near regions B and C1 in SI Appendix, Fig. S9B), a SOM placement consistent with expression in ETP-DN2a, as expected if driven by high PU.1 activity (compare with Fig. 3C). To see if these PU.1-dependent genes were up-regulated by Bcl11b deletion, we used trait-enrichment heat maps to compare these perturbations on each SOM gene metacluster (Fig. 3H; complete version in SI Appendix, Fig. S12A). In fact, only rare metaclusters showed opposite responses to the PU.1 antagonist and to Bcl11b deletion (Fig. 3H; e.g., metacluster 212). In the gold-standard set, the overlaps were limited and extraordinarily symmetrical, showing no bias toward cooperative or antagonistic relationships between Bcl11b and PU.1 (SI Appendix, Fig. S12B and Dataset S3). Interestingly, some genes under dual control of Bcl11b and PU.1 were repressed by both (e.g., Il2rb and Zbtb16, and more weakly, Kit) (23). Thus, neither Bcl11b nor PU.1 is globally epistatic to the other. Instead, they provide mostly orthogonal inputs to the pro–T-cell program.

Strong network intersection with E2A. In contrast, Bcl11b deletion effects had strong overlap with reported effects of E2A deletion in DN2-stage pro-T cells (22) (SI Appendix, Fig. S12C). The gold-standard gene responses to Bcl11b loss (Dataset S4A) overlapped greatly with the effects of E2A deletion (SI Appendix, Table S4B), highly biased to respond in the same direction (Fig. 3 F and G; compare with Fig. 3D; numbers given in SI Appendix, Fig. S12C) ($P = 3E-04$ by $\chi^2$ test). Most of the overlap was between genes up-regulated in both Bcl11b KOs and E2A KOs (Dataset S4C). There was less overlap between Bcl11b-dependent genes and genes affected in any way by E2A KO, but the overlap again was in genes dependent on both (Dataset S4D). Interestingly, this linked Bcl11b with the repressive effects of E2A even more than with E2A’s activating effects (46–48). Thus, Bcl11b appears to cooperate with E2A in its effects to promote developmental progression and to suppress innate-cell like gene expression, as summarized in the model shown in Fig. 4.

GRN Roles of Bcl11b as a Regulator of Cell Fate. Bcl11b showed surprisingly few specific cross-regulatory effects on other transcription factor-coding genes previously implicated in the progression from phase 1 to phase 2 in T-cell development. Metaclusters up-regulated in Bcl11b KO cells (Dataset S2) contained some phase 1 regulatory genes that we verified to respond consistently to Bcl11b, although not meeting gold-standard criteria, including Bcl11a, Nnap, and Jap (SI Appendix, Figs. S3C and S4), and distal promoter transcripts of Runx3 (SI Appendix, Fig. S11). Bcl11a is associated with progenitor and B-cell fates, whereas Nnap and distal Runx3 are expressed in NK cells. Metaclusters up-regulated in Bcl11b KO pro-T cells (Dataset S2) also contained later-acting T-cell regulators Gata3 and Lmo4, also verified by inspection to be frequently but weakly up-regulated (SI Appendix, Fig. S3C). These regulators are expressed not only in later T cells but also highly in ILC2 cells (Gata3) and in invariant NK T cells, ILC2, and ILC3 cells (Lmo4), and are also up-regulated in E2A KO samples (22).

Thus, Bcl11b during commitment could be damping the non-T lineage roles of these factors. The most striking effects of Bcl11b on other regulators were to maintain silence of genes seemingly extrinsic to the T-cell program. These include regulatory factor and cytokine receptor genes Zbtb16, Nfil3, Pou2af1, and Il2rb. These are genes without obvious drivers within the rest of the known program, and are normally poorly expressed, if at all, before Bcl11b is expressed (compare with SI Appendix, Fig. S2, DN1). Whereas most of these genes have at least some baseline expression, some sources of positive regulation are logically required even for targets that
normally have undetectable expression (e.g., *Tnni1*). The *Bcl11b* KO effects thus logically imply a cryptic positive regulator for them (Fig. 4).

**Discussion**

*Bcl11b* is a critical component of the T-lineage commitment machinery, but the target genes it regulates during the commitment process itself have not been globally defined until now. Here we have shown that *Bcl11b* exerts highly stage-specific positive effects on genes involved in TCR complex assembly and more complex negative effects on a range of target genes, some only during commitment and others at later stages as well, involved in effector response and cellular identity. *Bcl11b* has many binding sites around both positive and negative regulatory targets, which can be extensively mined in future detailed characterization for insights into its activating and repressive modes of action.

*Bcl11b* plays a key repressive role during commitment that affects two large groups of genes under distinct baseline patterns of regulation. The first group comprises genes, including the key growth factor receptor gene *Kit*, that are most expressed during phase 1 and turned off after *Bcl11b* is induced (Fig. 4, phase 1-specific expression). Some, though not all of these genes are only open to *Bcl11b* regulation during the DN2 to DN3 transition, and later may be permanently silenced. The second group consists of genes that are normally not activated to substantial levels at all during the phase 1 stages, then are highly enriched for genes used in NK cells, ILCs, and innate-like T cells, including the powerful regulatory genes *Id2*, *Zbtb16*, and *Nfil3* (6, 42, 49–55). Many of these genes, especially NK-associated genes, require continual *Bcl11b* action to keep them silent later (Fig. 4, zero or low normal expression). Many other genes are expressed earlier and silenced during commitment. Why are these genes not expressed before *Bcl11b* is activated? The fact that they are silent before *Bcl11b* turns on could imply that they need another positive driver that gains in net activity during T-cell specification in parallel with *Bcl11b* being activated. Is this truly a new function? Some of these targets, like *Zbtb16* and *Il2rb*, could have positive regulators that are masked before *Bcl11b* is turned on, as phase 1 regulator PU.1 also represses them (23) (Fig. 1I, L4). It could thus be loss of PU.1-dependent restraint in phase 2 that creates a “need” for *Bcl11b* to “take over” the repressive role. Further analysis will be required to relate individual *Bcl11b* occupancy sites at these loci to specific activating or repressive functions, which could ultimately indicate how *Bcl11b* works to antagonize their activators. However, the fact that *Bcl11b* is not needed to repress these genes until *Bcl11b* itself is activated suggests that *Bcl11b* acts on these genes in classic incoherent feed-forward circuit architecture.

The role of *Bcl11b* is mostly orthogonal to the role of PU.1, and it does not appear to be a direct repressor of most phase 1 transcription factor-coding genes. Many of its phase 1-associated targets encode cell surface markers, signaling, and cell biology components. However, the convergence of *Bcl11b* actions with E2A (46) or with the phase 1-specific Notch-negative feedback regulator, *Nrarp*, which is also highly expressed in NK cells. These genes are not all up-regulated uniformly in all *Bcl11b* KO cases, however. For example, *Zbtb16* was up-regulated in our DN2–DN3 KO samples made by both protocols I and II (*SI Appendix, Fig. S2 and Dataset S1A–C*), but more strongly in fetally derived cells and not in DP KO cells (2, 13; but see ref. 59). In contrast to *Zbtb16*, *Nfil3* was more strongly up-regulated in protocol II and in DP cells, again suggesting underlying differences in their needs for positive regulators that could promote alternatives to T-cell fate.

The fate alternatives restricted by *Bcl11b* have further differences from those controlled by PU.1. Pro-T* cells in phase 1 overtly express multiple myeloid-associated genes regulated by PU.1 (23), even as Notch signaling restrains others. However, these are silenced during commitment and most will never be expressed by T cells again. In contrast, many of the ILC response genes that *Bcl11b* represses are normally expressed much later by mature T cells during postthymic immune effector responses, despite the cells’ committed status (*SI Appendix, Fig. S7*). Thus, it may be physiological for these *Bcl11b*-repressed genes to be intrinsically primed for activation during T-cell specification. If so, then another way to see *Bcl11b*’s role at the DN2b stage is that it imposes a temporal delay between the time that these genes could be initially specified for activity (in DN2b pro-T* cells) and the time that mature TCR-αβ T cells (their much later descendants) will be allowed to deploy them. As we and others have discussed previously (60, 61), the timing of deployment of the effector response gene subnetwork distinguishes αβ-T* T cells, γδ-T* T cells, and ILCs more profoundly than the actual effector response networks themselves. Some of the targets we see up-regulated in *Bcl11b* KO samples in fact include genes specific for certain γδ lineages (e.g., *Cd163II, Tnai1*, and certain *Terg* transcritps). Thus, *Bcl11b* appears to control the conditionality of access to effector activation gene subnetworks even more than progression within a single canonical T-cell pathway itself.

The role of *Bcl11b* in commitment thus reflects a different mechanism of innate/adaptive divergence than the control of access to myeloid fates. Unlike PU.1, *Zbtb16* and *Nfil3* are not normally present in phase 1 regulatory state. However, considerable overlap between T-cell and ILC programs (6), including use of Runx3 and GATA-3, suggests ways that the T-cell specification process itself can generate mechanisms to prime these genes for expression. Loss of *Bcl11b* can allow levels of both *Gata3* and *Runx3* to rise. Maintaining repression then appears to depend on E protein, directly or indirectly, as well as *Bcl11b*. Thus, in the T-cell GRN, the most dynamically regulated T-cell factor collaborates with the most unchanging one to establish and preserve the committed state.

**Methods**

See *SI Appendix*, Supplementary Methods for details.

**Animals and Cell Preparations.** *Bcl11b* conditional KO mice, C57BL/6 (B6) controls, and PLBD ([68,129], *Bcl11b*+/–;ROS26-CreERT2 (2)) mice with or without a ROSA262E-yFP Cre-reporter gene were bred and used as sources of control and *Bcl11b*–deficient T-cell precursors as previously described (3, 27) (*SI Appendix, Supplementary Methods and Dataset S5*). B6.Lck-Cre mice were obtained from Taconic Laboratories. Both KO and control cells were Cre-treated and usually sorted based on Cre-induced *yFP* phenotype as well as DN2αβ phenotype. RNA was prepared and sequenced as described previously (8). Animals were bred and maintained under specific pathogen-free conditions in our colony at the California Institute of Technology under protocols approved by the Institutional Animal Care and Use Committee.

**Data Analysis.** RNA-seq and ChIP-seq fastq files were aligned and processed, and the data matrix submitted to SOM analysis (17), as described in detail in *SI Appendix, Supplementary Methods*. The gold-standard gene set was defined by EdgeR comparisons of eight KO and control sample pairs (*SI Appendix, Supplementary Methods and Datasets S1 B and C and S5*).
locus is differentially regulated at the level of expression by the Notch target and transcriptional repressor Hes1.

thymocytes.

We thank Georgi Marinov and Barbara Wold for Gene Network Construction.


