Transcriptional Control of Early T and B Cell Developmental Choices

Ellen V. Rothenberg
Division of Biology 156-29, California Institute of Technology, Pasadena, CA 91125

Abstract

T and B cells share a common somatic gene rearrangement mechanism for assembling the genes that code for their antigen receptors and developmental pathways with many parallels. Shared usage of basic helix-loop-helix E proteins as transcriptional drivers underlies these common features. However, the transcription factor networks in which these E proteins are embedded are different both in membership and in architecture for T and B cell gene regulatory programs. These differences permit lineage commitment decisions to be made in different hierarchical orders. Furthermore, in contrast to B-cell gene networks, the T-cell gene network architecture for effector differentiation is sufficiently modular so that E protein inputs can be removed. Complete “T-cell-like” effector differentiation can proceed without T-cell receptor rearrangement or selection when E proteins are neutralized, yielding natural killer and other innate lymphoid cells.

Introduction

From the first recognition of how somatic mutation, clonal receptor gene expression, and clonal selection operate in the adaptive immune system, T and B lymphocytes have appeared to be cell types with a self-evidently close relationship. Current evidence still indicates that they are the only cell types in the mammalian organism that use RAG1/RAG2-mediated programmed gene rearrangement in their development. The receptors that they use to recognize antigen are highly similar immunoglobulin superfamily structures which form the recognition surfaces for antigen when assembled into disulfide-bonded heterodimers. The development of the two lymphoid cell types presents even more striking parallels, as both pass through an ordered series of alternating proliferative phases, cell cycle arrest phases for gene rearrangement, and quality control checkpoints that operate to ensure a properly expanded population with a properly selected antigen recognition receptor repertoire. However, in development T- and B-cell precursors adopt strictly divergent paths at a surprisingly early stage of differentiation. Furthermore, recent evidence on the evolution of immune cell types indicates that the separation between T-cell-like and B-cell-like programs dates back more than 500 million years, before the use of immunoglobulin superfamily genes in antigen recognition (1). How can we understand the relationship between the shared and divergent features of these cell types? The answers lie in the use of distinct combinations of transcriptional regulatory network modules within the programs.
that generate these cell types, some of them mutually inhibitory, which this review will try
to bring into focus.

**Parallel, distinct, and more broadly shared developmental program elements**

**Parallel pathways for T and B cell precursor differentiation**

Major outlines of T and B cell development are well established and have been extensively reviewed as separate subjects (2–13). Figure 1 reviews the main pathways and stages for development of B cell and T cell precursors in mice, the system in which they have been most thoroughly dissected. Table 1 lists the markers by which successive stages are distinguished. Uncommitted hematopoietic precursors can develop into B cells in the bone marrow, primarily in the endosteal niche (14, 15), or in the fetal liver before birth. In contrast, uncommitted precursors must migrate first to the thymus in order to receive the signals that trigger T cell development, most importantly via ligands that activate the Notch signaling pathway. However, the two programs once under way are strikingly parallel, as shown in Figure 1, in which the program for B cells is compared with that for the major fraction of T cells that use TCRαβ-class receptors. From the earliest stages, the T and B cell programs display both shared and mutually exclusive characteristics.

Both T and B cell precursors begin their development with a period of transcriptional ground-breaking that turns on genes that contribute to antigen-receptor assembly and antigen receptor-dependent signaling, during the BLP, pre-pro-B and ProB stages of B-cell differentiation and the DN2a/2b and DN3a stages of T-cell differentiation. This early activation is important because antigen receptor gene rearrangement status and signaling competence must be repeatedly used at distinct checkpoints during development to determine the fate of the cells.

Immune receptor genes are assembled through tightly limited phases of somatic gene rearrangement, mostly at two specific stages that are equivalent in B and TCRαβ T cells (Fig. 1). For rearrangement of the immunoglobulin (Ig) and T-cell receptor (TCR) genes themselves, both T and B cells use the same gene products: RAG1 and RAG2 to recognize and cleave the recombination substrates, Terminal deoxynucleotidyl transferase (TdT=Dntt) for mutagenesis of the cleaved ends, and DNA-PK (Prkdc), Artemis, Ku70 and Ku80, and DNA ligase IV for resealing. Among these genes, expression of the Rag1, Rag2, and Dntt genes is specific to developing lymphocytes, and common to both T and B lineages.

In order to generate the signals that will enable cells to pass developmental checkpoints, the nascent antigen receptor gene products need to be assembled into signaling complexes. Thus, components that are not antigen receptors themselves must also be provided in early development. Such gene products include the CD3γ, δ, and ε and TCRζ chains (Cd3g, Cd3d, Cd3e, Cd247 respectively) for T cells and the Igα and Igβ chains (mb-1=Cd79a and B29=Cd79b, respectively) for B cells, and the transcriptional activation onsets for these genes represent useful landmarks for entry into each pathway. Also expressed transiently are immunoreceptor “surrogate chains”, pre-TCRα (Ptcra) for T cells and the combination of VpreB and A5 (Igll1) for B cells. In each lineage, these supply a lineage-specific
dimerization partner or partner complex to stabilize the first immune receptor chain to be expressed in each lineage, enabling a provisional immunoreceptor complex to be assembled at the cell surface even before the second chain of the immunoreceptor has been generated. Such antigen receptor assembly and signaling genes are turned on at parallel, corresponding stages of T and B cell development, i.e. the DN2a, 2b, 3a stages for T cells and the BLP, pre-proB, and proB stages for B cells. All of these events are respectively B or T lymphocyte specific.

Both T and B cell precursors begin their transcriptional priming for differentiation in coordination with a phase of stromal-cell dependent cell division. In mice, the activation of lineage-specific genes occurs in B and T cell precursors as they proliferate, initially under the influence of cytokines that include IL-7 (or TSLP) and Kit ligand or Flt3 ligand (16, 17). The involvement of the related receptor tyrosine kinases Kit and Flt3 in the progenitor cells and of IL-7R as the main growth factor receptor for early T cell and B cell precursors is another common link between the two programs. A role for CXCL12 (SDF) chemokine binding to the CXCR4 receptor is also likely to be crucial to support lymphoid precursors at early stages in the bone marrow (14, 15). Thus, similar growth factor receptor dependent signaling pathways begin to participate in both programs, long before antigen receptor-dependent signaling is possible. However, RAG1/2-mediated gene rearrangement is most efficient in the G1 stage of the cell cycle, and so proliferation needs to slow after initial expansion in order to enable antigen receptor gene rearrangement to proceed. Rearrangement in both B and T cell precursors focuses initially on genes that code for one of the heterodimer chains of the immune receptors: the TCRβ chain genes in TCRαβ T cell precursors or the Ig heavy chain (IgH) genes in B cell precursors, joining first one of the D to one of the J gene segments, then some V to the joined DJ segments. (A somewhat different path or set of paths is followed by T-cell precursors that rearrange TCRγ and TCRδ to form TCRγδ receptors. For reviews, see (18–22).) Many of the developing B cells or precursors of αβ T cells stop proliferating entirely to complete this round of gene rearrangement.

The ability of the cells to progress beyond this point now becomes tightly dependent on the success of the rearrangement event to generate a transcription unit that codes for a functionally competent TCRβ or IgH chain. The cells detect success by the assembly of the newly generated immune receptor chain with the pre-existing surrogate chain and signaling complex components. A new burst of proliferation can only be triggered by the ligand-independent signal generated when the new immune receptor gene product assembles successfully with the signaling complex components. This primary developmental checkpoint is the pro-B to pre-B transition in B cell development (pre-BCR checkpoint, in Fig. 1) and the β-selection checkpoint in T cell development (pre-TCR checkpoint).

Once IgH+ pre-B cells or T-cell precursors that have successfully passed through β-selection finish their proliferative burst, they are allowed to rearrange the gene segments coding for the second of the dimeric immune receptor chains, the Igκ or Igλ light chain genes for B cells and the TCRα chain for T cells. This process will create the cells' clonally distinct, mature antigen recognition structures, and only cells that generate successful mature immune receptors will be allowed to emerge as functional B or T cells. Success criteria for
B cells include expression of a light chain competent to dimerize with the IgH heavy chain, to form a receptor that is not grossly reactive with autologous structures within the bone marrow environment. Success criteria for T cells include not only correct dimerization and avoidance of gross autoreactivity, but also the ability to form low-affinity but significant interactions with major histocompatibility complex (MHC) cell-surface structures in the thymic microenvironment.

Limitation on parallels between T and B cell differentiation pathways

Although the outlines of the pathways of differentiation are similar, the signaling molecules that complex with T and B cell immunoreceptors are highly lineage specific and are expressed in mutually exclusive developmental patterns. They provide some of the clearest case studies for the transcriptional regulatory mechanisms that define T or B cell identity. Furthermore, the pathways of differentiation themselves have very different character because the T-cell developmental program branches to offer many more effector specialization options (CD4, CD8, NKT, Treg, and multiple classes of TCRγδ cells) than there are B-cell options.

Once a full TCRαβ heterodimer is expressed, at the CD4+ CD8+ “Double Positive” stage, the interaction with MHC triggers the “positive selection” phase of T-cell development (Fig. 1). In a departure from the pattern of B cell development, this selection event not only rescues cells from programmed cell death or developmental arrest, but also results in functional program divergence, as cells become programmed to a CD4, CD8, NKT (innate-like T cell), or Treg fate. The details of these branching intrathymic selection pathways are deeply studied and well understood (23–33), and although they continue beyond the scope of this review, we note here that they involve new roles for some of the same transcriptional regulatory factors that establish T-cell identity in the first place. Different sublineages of T cells, including different types of T cells that use TCRγδ receptors instead of TCRαβ receptors, may vary according to the degree to which they are polarized to separate functional pathways before they leave the thymus, by interaction with self-ligands, or in the periphery, by interaction with non-self ligands. However, T cells need to choose among various alternative sets of effector genes to poise for future TCR-dependent activation, and this choice is as much a feature of T-cell development as is the expression of a particular TCR heterodimer.

Further diversification of T-cell effector roles can continue, after the cells leave the thymus, as individual T cells continue to shape their transcriptional profiles in response to the conditions in which they encounter antigen. These stimulation-triggered changes involve choice of one of several distinct, coherent transcriptional response programs, each based on a different, highly coordinated pattern of cytokine gene expression, and each at least partly exclusive of the others. In addition to the cytotoxic T lymphocyte (CTL) killer program, which is mostly determined intrathymically, these programs include: a mucosal surface-defensive inflammatory program (Th17); a program that triggers antiviral responses and destruction of intracellular bacteria (Th1); a program that enhances expulsion of parasites (Th2); a partially overlapping program that enhances B-cell stimulation and Ig secretion (Tfh); and a program for damping the responses of other immune cells (Treg) which can be...
induced either within the thymus or in the periphery. The existence of these distinct functional branches within the T-cell program, and the need to coordinate expression of genes at numerous non-TCR loci within and between these branches [e.g. (34–40)], together give the T-cell program its distinctive transcriptional regulatory complexity.

Developmental paths and choices

Shared cellular and regulatory histories for T and B cell precursors

There is good evidence for strong common program elements underlying the competence to make T and B cells as well as other types of lymphocytes. Three large subdivisions of hematopoietic differentiation are recognized: erythromegakaryocytic development, myeloid development, and lymphoid development. Stem cell differentiation spawns partially-restricted precursors as intermediates that have been channeled toward (or away from) one of these three general directions through cell-intrinsic mechanisms. From stem cells through most multipotent precursor cells, B and T cell developmental potentials initially cosegregate (Figure 2; HSC, MPP, LMPP, ELP, “ALP” stages). They also cosegregate with potential to give rise to non-B, non-T innate lymphocytes (ILC), including Natural Killer (NK) cells. Interestingly, lymphoid precursors generally reveal an ability to give rise to dendritic cells as well, until they become fully committed to a particular B, T, or Innate Lymphocyte fate. Cells that have lost all access to erythroid and megakaryocytic development (LMPP, ELP) remain capable of giving rise to all these cell types. In contrast, many if not all pan-lymphoid precursors still possess conditional myeloid potential (41–48), although there are clear differences between the physiological conditions that favor generating all types of lymphocytes and those that favor generation of myeloid cells from the same precursors (49, 50). This suggests that lymphocytes emerge through a jointly lymphoid-myeloid biased progenitor pathway before specializing to different lymphocyte fates.

The most restricted precursors that can still generate both T and B cells are Lymphoid-Primed Multipotent Precursors (LMPP) and Common Lymphoid Precursors (CLP; those rigorously defined to be pan-lymphoid precursors are also termed “ALP”)(Figure 2). Both of these are distinguished from stem cells by their expression of transcription factors Ikaros (Ikzf1) and PU.1 (Sfpi1 in mouse, SPI1 in human), and the growth factor receptor protein tyrosine kinase Flt3 (Flk2). Whereas both have lost the ability to generate erythroid and megakaryocytic cells, LMPP are much better as myeloid precursors than CLP in vivo, whereas LMPP are slower to generate lymphoid progeny than CLP (45, 47, 51–54). Despite these consistent findings, several questions are unresolved. It is uncertain whether transition from an LMPP-like state through a CLP-like intermediate is an absolute prerequisite for lymphoid development (55). For example, Ikaros mutants that cannot generate CLP or early B cells at all can still enter the T-cell developmental pathway (56). It is also not completely clear yet what the regulatory difference is that explains these differences in myeloid potential, as lymphomyeloid bifunctional regulators like PU.1 and Flt3 (see below) seem to be expressed similarly in both. To some extent, the ability of cells to manifest myeloid potential is a function of the signaling environment, as Toll-like receptor signaling or alterations in G-protein coupled receptor signals can drive even the most restricted CLP to a nonlymphoid fate (57, 58). An obvious difference between LMPP and CLP has to do with
the relative levels of growth factor receptors IL-7R and Kit they express on their surfaces. Both are defined by high expression of the growth factor receptor Flt3, but LMPP express more Kit and less IL-7R, whereas CLP express less Kit and more IL-7R. IL-7R signaling can defer or antagonize myeloid development in several contexts as well as supporting lymphoid development (59, 60). However, expression levels of these growth factor receptors substantially overlap, and the fractions of both kinds of progenitors that have T-cell potential appear to share expression of the chemokine receptor Ccr9 as well (46, 61–64). In accord with their overlapping lymphoid potentials, then, LMPP and CLP type cells thus probably share responsiveness to common signals.

Intrinsically, B and T cell potential depends on a common, indispensable set of transcription factors. Precursors for both lineages are intensely dependent on the Ets-family transcription factor PU.1 for their generation and survival, and they also require the zinc finger transcription factors Ikaros (Ikart1) and Bcl11a and the basic helix-loop-helix (bHLH) “E protein” family member E2A (encoded by a gene known as Tcfe2a, bHLHb21, or Tcf3). PU.1 has other major roles in myeloid and dendritic cell development, but it is targeted in lymphoid precursors to distinctive genomic sites (65, 66). Ikaros, Bcl11a, and E2A are also expressed in stem or multipotent progenitors as well as lymphoid precursors (2, 44, 67–71), but they are most important for the lymphocyte developmental programs in general and absolutely required for B-cell development. In T cell development, where paralogs of Ikaros, Bcl11a, and E2A are also programmed for expression together with these three factors themselves, some compensation is possible for mutations in any one of these genes. Nevertheless, in Ikaros, Bcl11a, or E2A mutants the breakthrough T-cell development is suboptimal, abnormal, and prone to leukemic transformation (72–74).

Phenotypically recognizable CLPs only emerge in cells with intact function of Ikaros, Bcl11a, and E2A, and both B and T cell precursor activity is acutely dependent on E2A even under manipulations by which the precursors can be restored. Generation of both T and B cells also depends on a second group of factors which are important not only for producing lymphoid precursors but also for a broad range of other hematopoietic programs, including stem-cell, erythroid, neutrophil, and megakaryocyte development. In addition to the lymphomyeloid factor PU.1, already mentioned, these include the essential definitive hematopoiesis factor Myb (75–79), the zinc-finger repressive factor Gfi1 (80–85), the cytokine receptor signal transducing factor pair STAT5a/STAT5b (86, 87), and heterodimers of Runx family transcription factors with their common partner, CBFβ (88–92). Recent evidence suggests that a role for Hoxa9 may be important as well (93), while another broadly expressed zinc finger factor, Miz-1 (Zbtb17), appears important to potentiate responsiveness to IL-7 receptor signaling in early lymphoid precursors (94, 95). The requirements for programming of the CCR9+ LMPP and CLP precursors establish the ground state upon which both T and B cell gene regulatory networks are built.

**Timing of lineage choices for T and B-cell precursors**

In the bone marrow there is a continuum of B cell development from LMPP through CLP to EBF+, Pax5+ pro-B cells, and several combinations of surface markers and transgenic markers have helped to establish prospectively the developmental stages involved (Fig. 2,
Table 1). Activation of several genes in otherwise LMPP-like cells act as harbingers of lymphoid bias. These include the first expression of \textit{Rag1} as well as IL-7R (96). Cells with clear IL-7R expression are defined as CLP (52, 97) (Figs. 1, 2). At the CLP stage, the cells do not yet express any B-lineage restricted surface markers, but they are primed to begin lymphocyte development. Assay systems based on transplantation in vivo or clonal growth in culture without stroma indicate that they are greatly diminished in ability to generate myeloid cells, although under other conditions with supportive stroma they do reveal significant latent myeloid potential (49, 50). Many CLPs still retain considerable T, NK, and/or DC potential; these are also Flt3\textsuperscript{high} and Ly-6d-negative (“ALP”), whereas CLPs that have turned on Ly-6d expression are strongly biased to the B-cell fate (“BLP”) (98–100), (101). Progress towards B-cell commitment within these BLPs also correlates with downregulation of Kit. Complete commitment to the B-cell fate is accompanied by increased \textit{Rag1} expression and ability to express an Ig\textsubscript{\alpha} transgene, even before expression of classic ProB cell surface markers (102, 103). The cells are then termed pre-pro-B cells when they turn on B220, and pro-B cells when they activate CD19, when CD27 expression is also extinguished.

Myeloid potential of all IL7R\textsuperscript{+} CLP is reduced in comparison with that of IL7R\textsuperscript{−} LMPP. It is ambiguous whether the B-cell biased CLP have lost the last vestiges of myeloid potential any more absolutely than the Ly6d\textsuperscript{−} CLPs (“ALP”), but they differ from ALP in showing a greatly reduced ability to generate T cells or NK cells (99). Thus, from a B-cell perspective, T- and NK-lineage fate exclusion is the most prominent aspect of lineage commitment, and closely linked with the initiation of the B-cell differentiation program.

For T lineage cells, the impact of Notch signaling causes lineage choices to be made in a different order. The events that cause loss of access to the B-cell fate are not the rate limiting events of T-cell commitment, but rather are the earliest responses that the cells make under the influence of thymic signals (Fig. 2, “thymus settling” precursors vs. ETP). In contrast, the final stages of T-cell lineage commitment clearly involve mechanisms that exclude fates other than the B cell fate. Myeloid and dendritic-cell potential is readily demonstrated in early T cell precursors from fetal or adult thymus, at stages many cell divisions after B-cell potential is lost. The T-cell precursors most related to bone marrow or fetal liver “ALPs” are the Early T-cell Precursors (ETP; Kit-high DN1 cells)(41, 56). In postnatal animals, evidence from lineage tracing reporters, intrathymic Notch signaling mutations, and actual cell purification and culture assays has converged to imply that many cells can enter the thymus with B-cell potential intact (61, 64, 104–108), but they lose this B-cell potential early during the expansion of the ETP population. In fetal thymus during the first wave of T-cell development, ETPs appear to enter the thymus devoid of B-cell potential already (109, 110). Thus, even in the postnatal thymus, the great majority of ETPs at any given time will already have become intrinsically unable to adopt the B cell fate, even when switched to a favorable B-cell promoting environment (61, 107, 108). Under the same conditions, in contrast, they retain readily demonstrable ability to generate macrophages, dendritic cells, granulocytes, and NK cells among other cell types (111–120). This natural plasticity coincides with the sustained relatively high expression of the myeloid and progenitor-cell factor \textit{PU.1} in the T-cell precursors throughout these cell cycles, while B-cell specific
regulatory gene expression is strictly prohibited. Access to the dendritic cell pathway then decreases after the ETP stage, but the ability to generate innate-type lymphocytes and nonlymphoid cells alike persists for multiple additional generations, into the DN2 stage, and is only lost as the T-cell precursors transition from DN2a to DN2b (see Fig. 1). Loss of B-lineage access thus not only precedes loss of myeloid and dendritic potential; it also precedes the upregulation of recombinase factors Rag1, Rag2, and Dntt in T-lineage cells, which occurs in the DN3 and later stages of T cell development, as discussed below (for review, see (121)). Thus, even the features that are most shared by B and T cells are controlled through developmental programs that in fact differ sharply between B- and T-cell precursors.

Creating a B-cell specific regulatory state

Essential regulators of the B cell program

B cell development depends on a vital combination of transcription factors which are activated or upregulated early in B cell development and sustained in activity long into mature B cell immunological responses. The threshold to begin B cell development is crossed when the transcription factors EBF1 and Pax5 are turned on de novo, creating a self-stabilized and unique regulatory state (2, 4, 122). EBF1 expression, together with E2A, is needed to establish a recognizable B-lineage gene expression profile. Pax5 expression is needed to complete and stabilize B-cell lineage commitment, to open the full range of Ig heavy-chain V regions for rearrangement, and to complete the signaling complexes that enable the cells to receive B-cell receptor signals (5).

Both EBF1 and Pax5 are expressed only in B lineage cells among all hematopoietic cell types, even though they are representatives of transcription factor families with ancestral roles in completely different tissues including adipose tissue, stroma, kidney, liver, and the central nervous system. Because there are no paralogs for EBF1 and Pax5 used in other hematopoietic cells, the advent of these factors with their unique DNA binding specificities establishes a regulatory state that is clearly disjunct from those in any other hematopoietic cells. In combination with the pan-hematopoietic legacy of Myb, Runx, Ikaros and other factors, the addition of EBF1 and Pax5 gives early B-cell precursors a repertoire of DNA binding factors that is probably unique in the whole organism. Many if not all B-lineage specific genes appear to get direct regulatory contributions from EBF1 and/or Pax5. Only in the very terminal stages of B-cell immune responses, when the cell switches off most B-cell genes and transdifferentiates into a constitutively antibody-secreting plasmacell, do Ebf1 and Pax5 become silenced.

Starting the B-cell program

Activation of Ebf1 and Pax5 is triggered by a combination of E2A activity, FoxO1 activity, and Runx/CFBβ activity, under permissive conditions that depend on IL-7R signaling in the earliest CLP. The cells must surpass a certain minimum net level of E2A activity before this can occur. The rise in activity from LMPP to CLP is not fully explained, but many uncommitted early lymphoid precursors already detectably express E2A target genes such as Rag1, Rag2 and Dntt at low levels (124–126), and expression of the diagnostic receptor
genes Flt3 and Il7r itself appears to depend on E2A as well (98, 127). E2A then appears to upregulate one of the additional factors with which it will collaborate for EBF1 activation, namely FoxO1, creating a feed-forward network circuit (128, 129). Runx/CFBβ activity at this point is required independently (89), and as a prerequisite for direct positive regulation of EBF1 as well (88). Once EBF1 is present, it collaborates potently with E2A to turn on multiple early B-cell genes directly (124, 130, 131) and to activate Pax5 expression, with assistance from PU.1 working at another Pax5 enhancer element (132). This switch to turn on EBF1 and Pax5 also draws upon the background contributions of Ikaros (44, 133–136), Bcl11a (67, 137), Myb (76, 77, 138), and Gfi1 (81, 82) expression, for additional regulatory and supportive inputs.

Activation of EBF1 and Pax5 is not easy. Even if E2A, Runx1/CFBβ and FoxO1 are present, these B-cell specific regulatory genes can be kept silent if the myeloid transcription factor C/EBPα is induced (139–141). Excessive signals from a variety of cytokine receptors are sufficient to accomplish this in uncommitted lymphoid precursors (139, 142, 143). Forced ectopic imposition of C/EBPα rapidly turns EBF1 and Pax5 off, together with their target genes, not only in ALPs but even in later-stage cells that had already committed to a B-cell fate (140, 141), as the cells are reprogrammed to a myeloid fate. This may be facilitated by the presence of the myeloid-supportive factor PU.1 which is also required by these B-cell precursors. As already noted, Toll-like receptor signaling in the CLP may also abort B-cell development (57), and this may occur by redirecting the effects of PU.1 to activation of myeloid or dendritic-cell genes instead. In normal multilineage precursors, even if present, E2A activity can also be antagonized by signals that induce the E protein antagonists Id1, 2, or 3. Id1 and Id2 especially are well expressed in a variety of other hematopoietic progenitors and cell types (98, 144), and initially E2A itself can induce Id2 expression in a negative autoregulatory feedback loop (145). Thus, while the outlines of B-cell development can be understood once a critical threshold of E2A activity is crossed, there is competition for the fate of the cells. The SNAG-domain repressor, Gfi1, appears to be important not only to moderate levels of PU.1 in early precursors of B cells (82), but also to keep Id1 and Id2 expression in check (81, 85). Ikaros can directly limit PU.1 expression once cells have entered the B lineage pathway (146), and may also contribute to crossing the E protein:Id protein threshold, by helping to activate Gfi1 (82, 133).

Managing IL-7R signals may also require special regulation, to accommodate the cells' peculiar use of the global cell cycle arrest factor FoxO1 as a part of their cell-type specification machinery. IL-7R signals are crucial for early murine B cell development whether or not they directly activate Ebf1 (94, 123, 147, 148). Both Myb and another required transcription factor, Miz-1 (Zbtb17), act to promote IL-7R expression and sensitivity (76, 77, 94). However, signaling through a growth factor receptor like IL-7R would generally activate not only STAT5, which either activates or supports expression of EBF1, but also the phosphatidylinositol (PI) 3-kinase pathway. PI 3-kinase activation

1Although binding sites for STAT5 have been identified in the Ebf1 cis-regulatory regions, the role of IL-7 receptor signaling may be more to support viability of the cells or to enhance Ebf1 expression indirectly than to participate directly in Ebf1 activation (123). However, this dependence has been evaluated in cells that have already turned on substantial Rag1 expression, which may reflect primarily “BLP” and later cells.
triggers the protein kinase Akt, which stimulates proliferation through a pathway that powerfully inhibits FoxO1. Yet FoxO1 is important not only to activate EBF1 but also for direct regulation of Rag genes, Dntt, and Il7r itself, in collaboration with E2A (149–151). Thus, in principle, if IL-7R strongly activated PI 3-kinase at a pre-specification stage it could retard activation of Ebf1 and a key group of B-cell genes. B lineage cells appear to be able to avert this problem by regulating the coupling of IL-7R signaling to PI 3-kinase as opposed to STAT5 activating pathways (149, 152). Later in pro-B cells, when cell cycle arrest is needed to permit efficient Ig gene rearrangement, IL-7R-mediated activation of STAT5 itself is curtailed (153), but shifts between FoxO1 protection and de-protection punctuate the later phases of B cell development as well (149, 154).

Mechanistic insights into transcription factor activity to define cell type in early B cells

As soon as Ebf1 and Pax5 are activated, these B-cell regulatory factors become entrained in a direct mutual positive feedback circuit which amplifies and locks down the B-lineage identity of the cells. Analyses of transcription factor mutants, epistasis tests, and direct analysis of transcription factor binding across the genome in pro-B and pre-B cells have provided a very strong picture of this circuit (75, 124, 128, 132, 155–158). These results show that E2A has direct positive inputs into Ebf1 and Pax5 loci, EBF1 and PU.1 also have direct positive roles in regulating Pax5, EBF1 and E2A directly sustain FoxO1 expression, while Pax5 and FoxO1 positively regulate Ebf1 as well. Furthermore, although E2A is already broadly expressed in hematopoietic precursors, it is highly upregulated in B-cell precursors. Besides a high level of transcription of the gene coding for E2A, the effectiveness of the E2A protein is sharply increased once EBF1 is present, by the fact that EBF1 powerfully keeps Id2 repressed (157, 159). This inhibition of an inhibitor makes B-lineage specification irreversible, while the mutual positive feedbacks between EBF1 and Pax5 help to lock in a stably specified state.

Not only E2A and FoxO1, but also Runx1 and PU.1, can be drawn into the enterprise of activating B-cell specific differentiation genes together with EBF1 and Pax5, through corecruitment to sites of joint binding with lineage-specific factors (66, 89, 124, 132, 160). PU.1 binding in particular shows a very different distribution in pre-pro B cells and B cells than in myeloid cells, shifting in the B-lineage cells to a pattern dominated by linkage to E2A or EBF1 sites (66). In addition to their activating roles, moreover, EBF1, Pax5, and even E2A and FoxO1 acquire specific repressive roles in the B-cell context (75, 124, 128, 155, 157, 159, 161–163). The ability of the same factors to silence non-B cell genes as to activate B-cell genes gives the B-cell program its tight coherence.

B-cell gene networks begin to operate in the critical stages through which a CLP becomes a pro-B cell, using the upregulation of the marker Ly6d and the downregulation of Flt3 (Flk2) to dissect these early events as described above (99–101, 122, 164). Before initiation of the B-cell program, many E2A binding sites are associated with likely Runx target sites (124). The nature of the dimers that contain E2A in the earliest progenitors may in fact be different from the E-protein type homodimers in slightly later B-lineage precursor cells, since the total E2A pool in the progenitors likely includes a substantial fraction of E2A in heterodimers with SCL (Tal1) or especially Ly11, class B bHLH factors which contribute
their own fine-scale binding preferences. Ly1 continues to be expressed throughout B cell development. (165, 166). In whatever dimer form it is found, however, E2A positively regulates FoxO1 in these early precursors, and then works with FoxO1 to turn on expression of EBF1 as already noted (128, 129). Individually, many cells activate EBF1 before they express Pax5, and this initial activation corresponds with the “ALP to BLP” transition (102). The presence of EBF1 enables E2A to redistribute to new sites: about half of the EBF1 sites are sites of co-binding with E2A, many of them sites where E2A was not bound before (124). FoxO1 collaborates intimately with both E2A and EBF1 in regulating genes that define the B-cell identity (128). As Pax5 is induced, it in turn creates new sites of recruitment for the other factors.

Like PU.1 and E2A (66, 124, 167), both Pax5 and EBF1 can play rate-limiting roles to recruit chromatin-modifying enzymes in the context of gene regulation (155, 157). Both EBF1 and Pax5 are clearly bifunctional. As is often seen by comparison of ChIP-seq with gene expression analysis of mutant cells, both transcription factors are found engaged at many-fold more high-quality binding sites than the number of genes that they functionally affect. However, their binding is implicated directly in both positive and negative gene regulatory responses.

Pax5 has a larger number of known binding sites across the genomes of developing B cells than EBF1, and these shift significantly as B-cell development proceeds, even between successive immature B-cell states. It binds to a large fraction of all active cis-regulatory elements at given stages of B-cell development, even though only a small percentage of the bound genes show Pax5-dependent activity patterns. Pax5 binding has been shown to result in the appearance of new DNase-hypersensitive sites in the genes that it activates, whereas it causes loss of DNase-hypersensitivity at sites in genes that it represses (164). In the case of EBF1, a comparison of binding, local histone modification, and gene expression across successive B-cell developmental stages shows that EBF1 binding can also poise a B-lineage gene’s cis-regulatory site for future activation, long before the full requirements have been met for transcriptional activation. This activity is correlated with local histone modification to a histone H3K4me2+ state, which can persist through immature developmental stages until the EBF1-bound regulatory element is ready for full activation in the mature B cell (157). However, EBF1 cannot freely open chromatin at all of its own cognate sites across the genome. Its binding in early B cells is shaped by the hematopoietic context, and ectopic expression of EBF1 only turns on B-lineage target genes in a hematopoietic context. In fact, it binds a completely different spectrum of genomic sites in another context where it plays an equally important developmental role, in adipose cells, where it also interacts with a different set of regulatory partners. Even in mature B cells, EBF1 shifts to occupy a different spectrum of sites (75), where the rules for its collaborations with partners such as FoxO1 also appear to change. Thus, despite the continuing importance of this factor in mature B cells, its precise deployments vary.

One highly B-cell specific locus that has been studied in depth, the Cd79a locus (mb-1, Igα), has provided a detailed example of the mechanism through which these B-cell specific and broadly utilized factors can collaborate (2, 160). EBF1 supplies the critical rate-limiting element needed to complete a DNA-binding complex with E2A and Runx1/CFBβ, which
can then initiate local DNA demethylation of the Cd79a promoter. As demethylation propagates from the site of EBF1, E2A, and Runx1 binding, it unmasks a key Pax/Ets composite site which can then cooperatively bind Pax5 with a pan-lymphoid factor, Ets1, to lock the promoter into an active state. Versions of this intricate transcription factor collaboration are likely to be occurring at many other sharply activated B-cell genes.

Another country: the leap into T-cell development

Three ambiguous regulators and an “outside agitator”

T-cell precursors begin with the same intrinsic regulatory factor complex as B-cell precursors, inherited from multipotent progenitor ancestors, and they initially express very similar cytokine and chemokine receptors to guide and mediate their interactions with environmental cues. However, the T-cell developmental program depends on the migration of the cells into a foreign environment, namely the thymus. The special feature of the thymus is its dense presentation of potent Delta-class Notch ligands, primarily Delta-like 4 (DLL4), which bombard the cells with signals that ultimately render most non-T-cell fates inaccessible (7, 13, 168). Notch signaling prohibits activation of the B-lineage specific factors EBF1 and Pax5 and sets in train a regulatory cascade that ultimately makes their silence irreversible. It also antagonizes myeloid development and blocks access to several other fates through mechanisms described below. But in order to forge a T-cell identity for the progenitors, it also needs to turn on T-cell specific regulatory factors. The most T-lineage specific of these factors in fact comprise a very short list: they are TCF-1 (encoded by the Tcf7 gene) (169–172), GATA-3 (173–177), and Bcl11b (178–182).

There are a number of central differences between the T-cell program governed by Notch, TCF-1, GATA-3, and Bcl11b and the B-cell program governed by EBF1 and Pax5.

First is that the Notch signaling is extrinsically controlled and non-cell-autonomous. Even so, it is not simply a transient trigger to activate the T-cell specific factors. Instead, it remains a sustained regulatory requirement throughout multiple cell cycles of the T-cell specification program, before and even for some time after commitment. Thus even after GATA-3, TCF-1, and Bcl11b are turned on, many T-cell differentiation genes continue to receive direct regulatory input from an environmental signal-dependent transcription factor and cannot continue to be expressed if the contact with the environment is interrupted (183–185).

Second, the essential T-cell specific transcription factors themselves are mostly confined to developing T cells, with instructive exceptions discussed below; but their paralogs are abundantly expressed and prominent in other hematopoietic lineages including hematopoietic stem and progenitor cells. GATA-3 is a close relative of the stem-cell and mast-cell factor GATA-2, and of the erythromegakaryocytic factor GATA-1. Bcl11b is a close relative of the pan-lymphoid required factor Bcl11a. TCF-1 itself is related to TCF-4 (Tcf7l2 gene product), which is also expressed in stem and progenitor cells, although the level to which TCF-1 rises in T-cell precursors is much higher. The progenitor factors are turned off in different patterns, but mostly at about the same time that their T-cell relatives are turned on (11). The close biochemical similarities between T-cell specific factors and...
their relatives expressed in multipotent progenitors should enable them to bind to similar or identical target sites, all other things being equal. This makes it more challenging than in B cells to understand how new cis-regulatory elements are activated in early T-cell development.

Third, many of the other transcription factors expressed by T-cell precursors are coexpressed along with their paralogs, providing considerable regulatory redundancy. Examples include Ikaros, which is coexpressed with 1–3 other family members during T-cell specification; E2A, which is coexpressed with two isoforms of its close relative HEB (Tcf12); all three Runx family members; more than 10 Ets family members, and even TCF-1, which after the initial stages is coexpressed with its close relative LEF-1 (186–188). The patterns of expression of these family members are not the same, and there are some examples of different family members that antagonistically cross-regulate each other in early T cells: e.g. Gfi1 and Gfi1b (189), PU.1 and Ets1 (190, 191), and particular isoforms of TCF-1 against LEF-1 (192, 193). Nevertheless, the cases of Ikaros/Helios and E2A/HEB make it clear that considerable redundancy is real. This systematic lack of parsimony in the cells' regulatory behavior may indicate that evolutionary selection has buttressed these regulatory functions so that they are not vulnerable to mutation in a single gene. Ironically, though, the weak single-gene knockout phenotypes that result may be interpreted to suggest the reverse.

Fourth, unlike the B-cell specification network, the T-cell specification gene regulatory network does not appear to be dominated by positive feedback relationships. Instead, at least three different network modules within the program can be discerned, and the interactions between them are rich with cross-inhibition, autoinhibition, and incoherent feed-forward circuit relationships.

**Stages in progression toward T-lineage commitment**

Because T-cell development occurs in a separate anatomical compartment from the rest of hematopoiesis, it is easier to define phenotypic stages in the process even before a distinctive set of T-lineage marker genes is turned on. Progressively finer stages have been distinguished by phenotype and placed in sequence by the development of increasingly powerful developmental assay systems. Whereas precursor-product relationships in the early pioneering studies had to be assayed by injection of precursor cells into the thymus of a recipient mouse, the cells could later be manipulated by reconstitution into fetal thymus organ cultures. Elegant modifications of these systems revealed a wealth of information not only about early development but also about later TCR-dependent selection events (194–196). The field of early T-cell development has taken a great step forward with the development of completely open monolayer coculture systems in which the thymus is replaced by OP9 (or Tst4) stromal cells transfected to express DLL1 or DLL4 (197–199). The ability to track, modify, interrupt and restart the T-cell development process at will has made the contingencies of progression and the single-cell developmental potentials at each stage in this system extremely well defined.

The cells that enter the thymus initially (thymus-settling precursors) are not committed to the T-cell lineage. Both in terms of regulatory genes and in terms of their targets, these cells appear very similar to lymphoid-primed multipotent precursors. As they expand in the
thymus under the influence of Notch-DLL4 interaction and local Kit ligand expression, they initially express growth factor receptor genes, for example, which are most associated with myeloid and dendritic cell fate (Flt3, Kit, Csf2rb, Csf1r), and signal transduction kinases (Lyn, Btk) that are more commonly associated with B-cell function. However, interestingly, even when they are derived from IL-7R+ CLP (50) and/or Rag1+ Early Lymphocyte precursors (96), these ETP cells have no detectable IL-7R and virtually no Rag1 expression. Conceivably, therefore, their intrathymic development begins with an actual reversal of certain lymphoid-priming gene expression steps that had already occurred in their prethymic antecedents (see below).

More in parallel with B cell precursors, ETP then go on to downregulate Flt3 midway through this first phase. Initial repression of Il-7R and Flt3 downregulation in thymus-settling precursors may be important for the process leading to T-lineage restriction (108, 200) as described in a later section. The survival of these precursors thus becomes dependent on Kit/Kit ligand interaction (201–203), and the high Kit level maintained on ETPs and DN2a cells is one of their defining characteristics. As Flt3 is downregulated, some progenitor-associated transcription factors begin to be turned off (Meis1, Mef2c, Lmo2, and early steps in the downregulation of Bcl11a and Erg) (rev. in (11)), and the pro-myeloid Cebpa regulatory gene begins to be silenced (65, 204). As Notch-DLL4 signaling turns on Hes1, it also causes the cells to activate Gata3 and Tcf7 expression (205, 206). Tcf7, at least, is clearly a direct target of the Notch-RBPJ (CBF1/Suppressor of Hairless/Lag1) transcription factor complex (169, 170). Thus, by the time the first definitive cell-surface markers of entry into the T-cell pathway are expressed during the transition from ETP/DN1 to DN2a, the cells have already modified their progenitor inheritance in important ways.

The transition into DN2a is indicated by upregulation of several other genes that are at least partially driven by Notch signaling. Notch directly induces sharp de novo activation of cell-surface marker CD25 (IL2 receptor α, Il2ra) and the HEBalt promoter isoform of Tcf12 (187). This is accompanied by increases in levels of already-expressed Runx1, Gfi1, Ets1, and the canonical isoform of Tcf12, as well as further increases in Gata3 and Tcf7 expression. These factors promote the first increases in expression of numerous T-cell marker and functional genes including Thy1, Cd3g, Cd3d, and Cd3e, that will increase their expression more dramatically in the next stage. Now Rag1 is (re)activated and Il7r is also strongly upregulated, and the expansion of the cells becomes highly sensitive to IL-7 levels. Even at this early point, the T-cell program begins to offer cells various possible branches of specialization within the T-cell fate. It is at the DN2a/2b stages that the precursors may have their highest propensity to rearrange TCRγ and TCRδ genes instead of TCRβ (207, 208) (Fig. 1), and this too is known to be dependent on IL-7 (209, 210).

Despite their broad expression of T-lineage genes, however, DN2a stage cells remain uncommitted to the T-cell fate (Fig. 2). They have already lost the ability to make B cells, and this ability cannot be restored by removing them from the DLL-rich environment. However, upon transfer to recipient animals or to DLL-free cultures with appropriate cytokines, they can still develop into NK cells, dendritic cells, macrophages, or granulocytes (rev. in (121)). The DN2 population also includes rare cells with mast-cell potential (114). A fraction of them can even respond to the addition of cytokines IL-33 + IL-7 in the presence

Anna Rev Immunol. Author manuscript; available in PMC 2015 March 21.
of Notch signaling, to become an innate lymphocyte type called ILC2 ("nuocyte") (211).

Note that these are not "reprogramming" experiments in which the regulatory state of the cell needs to be altered by transfection to force new transcription factors to be expressed. Instead, these assays explore the flexibility of the cell's intrinsic developmental gene network at that stage, simply providing the intact cell with supportive environmental conditions for a variety of options. The cells remain uncommitted through multiple rounds of cell division under the influence of Notch signaling (212), until they cross into the "DN2b" stage. They then become committed, unable to develop into anything but T cells, no matter what the environment.

Commitment was first detected as linked to activation of a particular Lck-GFP transgene (213), and is functionally dependent on onset of expression of the transcription factor Bcl11b (178, 179). Bcl11b is sharply upregulated in late DN2a stage and then strongly expressed in virtually all T cells from the DN2b stage onward (182, 214). Its advent causes the cytokine receptor Kit to be downregulated, while the CD3 cluster, Thy1, and Rag1 genes are now strongly upregulated. The surrogate TCR chain gene Ptcra, the surrogate partner for a newly rearranged Tcrb gene product to make the pre-TCR, also begins its steep rise in expression which will peak just before β-selection. As additional T-cell signaling genes such as Cd247, Prkcb1 and Itk are turned on, and Zap70 begins to be expressed, the prerequisites for TCR signaling approach completion (65, 215).

The drivers of these climax events in T-cell specification may come from a variety of regulatory inputs, working independently or jointly. From the DN2b stage to the DN3a stage, the cells appear to experience stronger Notch signaling, since the Notch signal-dependent target genes Dtx1 and Notch3 as well as Ptcra all shoot to their highest levels during this transition. Many of the genes activated at this stage are also thought to be E2A/HEB targets (216, 217). In fact, the DN3a phenotype depends on E2A and/or HEB maintenance (218), implying that the net activity of these bHLH factors reaches a maximum at this point as well. The E2A/HEB-dependent genes include the recombinase genes (Rag1, Dntt) that now reach a peak of expression. Meanwhile, GATA-3 reaches its highest permitted level, and TCF-1 is joined by its paralog, LEF-1, from DN2b stage onward. There is a sharp increase in expression of the pan-lymphoid factor Ets1, which among diverse other functions collaborates with Runx1 to activate the enhancer of the TCRβ gene (219). With these major actors on the stage, the regulatory requirements to undergo β-selection are complete.

**Molecular switches in the T-cell lineage commitment transition**

The exclusion of T-cell fate alternatives is caused at least in part by extinguishing expression of the "non-T" regulatory factors with which the cells began their intrathymic development. This is easy to understand for the factors that would otherwise provide access to myeloid development. C/EBPα is strongly downregulated starting within the ETP stage, and its coding locus Cebpa becomes progressively more heavily modified by repressive histone marks throughout commitment (65). PU.1 itself, although required for the generation of T-cell precursors, is an antagonist of the later T-cell program (190, 191, 220–222). It becomes silenced between the DN2a and DN3a stages (111, 222) through a mechanism...
dependent on Runx factors and possibly GATA-3 (114, 223–225). With the shutoff of PU.1, the cells lose a crucial prerequisite for both myeloid and dendritic-cell development.

Silenced in approximately the same interval are the stem- and progenitor-cell factor genes Tal1 (coding for SCL), Gfi1b, and Hhex; Lyl1, the paralog of Tal1, is repressed just a stage later, between DN2b and DN3a. Bcl11a, Tsc22d1, and the progenitor-cell-specific Ets-family gene Erg are also downregulated throughout this process, although their final extinction is not until DN3 (65, 215)(rev. in (11)).

The exclusion of the NK cell alternative is a little different, since well-known NK-cell regulatory genes by default begin T-cell development silent in a repressive chromatin context: i.e. Tbx21, Eomes, Zfp105 (65). However, they can be activated readily when the cells are exposed to supportive cytokines in reduced Notch signaling conditions (228). The pathway may involve STAT5 and its potential to induce Id2 (229), assisted by Ets1 activity (230), which is well established in the earliest T-cell precursors. Id2 itself is not a transcriptional activator, but Hhex, one of the progenitor-associated transcription factors expressed normally through the DN2a stage, has recently been found to participate in the NK cell program (231). Another NK program activator, Nfil3 (E4BP4) (232, 233), is also weakly expressed but still available for activation in early T cells, even beyond the ETP and DN2a stages. Conceivably these factors provide the thin end of a wedge for activating the other NK-cell regulators.

The window of opportunity for switching to NK lineage fate is most likely limited by the advent of Bcl11b expression (compare Fig. 1, Fig. 2). Once it is turned on, Bcl11b is a powerful, nonredundant repressor of NK differentiation and NK cell potential. During and even after commitment (179, 180), it is required for the ongoing repression of three regulatory genes involved in NK cell development (231), i.e. Hhex, Id2, and Nfil3. Loss of Bcl11b enables T-lineage cells rapidly to upregulate these genes, followed by Tbx21 and Eomes and their target differentiation genes, as well as other regulators stringently restricted to the NK pathway (180, 181, 234). Thus, whereas other fates are excluded by silencing their main activators, the NK cell fate may be excluded by the activation of a constitutive repressor.

In contrast to factors supporting the myeloid, DC and NK fates, the crucial B-cell factors which distinguish the B-cell pathway from the T-cell pathway, EBF1 and Pax5, are by default both silent and inaccessible in repressive chromatin in ETP populations (65). The mechanism that renders and keeps them silent therefore needs to be explained in terms of the earliest events occurring when cells experience the thymic microenvironment. One factor that is needed for B lineage exclusion is GATA-3 (235), as discussed below.
T-cell transcription factors and T-lineage precursor survival

Molecular mechanisms operating in the T-cell developmental program are not as well understood as those of the B-cell program, but it is clear that specification depends on positive regulatory inputs from GATA-3, TCF-1, and E proteins, under the sustained influence of Notch signals. The roles of these factors are all required, to the best of current knowledge, from the ETP stage through β-selection. During T-lineage specification, these requirements are enforced by linkage of GATA-3, TCF-1, E proteins, together with Notch signals not only to particular differentiation genes as targets, but also to proliferation and survival pathways.

Notch signaling is important for proliferation, and becomes increasingly important for viability as the cells progress through commitment and lose their non-T developmental options. Perhaps the maximum dependence on Notch signals for viability is seen in the committed DN3a cells as they await a successful gene rearrangement for β-selection (236). These cells are no longer proliferating and can only be saved from death by a successful TCRβ gene rearrangement. Yet during this interval they draw support from Notch signals for basic survival until they can qualify for β-selection.

After β-selection, Notch signaling becomes disabled and T-cell identity becomes Notch-independent, presumably sustained because the other factors and Bcl11b continue to play crucial roles.

For GATA-3 and TCF-1, as early as the ETP stage there are catastrophic declines in viability if these factors are removed (170, 237). This is striking because the immediate precursors of ETP, before thymic entry, do not express either factor yet. This newly acquired “addiction” emphasizes how much impact intrathymic signaling has on the programming of cells even before they progress beyond the ETP stage. However, it also means that some gene regulatory effects of these factors are hard to measure, since the cells in which loss of function has altered gene expression may not be viable.

E protein function is important also from the prethymic stages onward (71), and many of the T cell differentiation genes, e.g. Ptcra, the HEBalt promoter of Tcf12, and Rag1, are apparently directly activated by E proteins or E proteins in collaboration with Notch (187, 216, 217, 238–240). Despite partial redundancy with HEB (241), loss of E2A alone causes sharp population decreases in the ETP and DN2a compartments, as well as a severe impact on the regulation of the β-selection checkpoint for cells that manage to get that far (74, 242, 243). HEB (Tcf12) is also crucial for T-cell development, with a special positive role during the DN2-DN3 stages for the “HEBalt” variant isoform of HEB that is induced in response to Notch signaling (187, 244).

Balancing act: restrictive cross-regulation of T-cell transcription factors

These equally essential T-cell factors might be predicted to work in a positive regulatory feedback circuit similar to the E2A, FoxO1, EBF1, and Pax5 circuit in early B cells. Indeed, there is evidence for a positive feedback between E proteins and Notch within the T-cell context (227, 245). An “exception that proves the rule” is the case of TCRγδ cells, which are weaned from their Notch-dependence precociously (207, 246) when, or because, their
assembly of TCRγδ complexes triggers many of them to upregulate the E protein antagonist, Id3 (247).

E2A itself helps to turn on Notch1 expression, as one key contribution to T cell development (71, 216). This may seem strange in view of E2A’s crucial role in B cell development, which is aborted by Notch signals. However, the Notch1 transmembrane molecule is not itself a transcriptional regulator until it is cleaved after contact with a ligand (e.g. DLL4 or DLL1), and cells can express Notch1 very strongly without activating Notch target genes if the environment does not provide these ligands. Therefore, as long as cells are in a DLL-poor environment, there is no contradiction between E2A turning on Notch1 itself and E2A triggering the B cell pathway. Interestingly, in the context of T cell development, Notch1 signaling also may reciprocate to protect E protein activity. E protein antagonists Id2 and Id3 are inducible in T-cell precursors by cytokine or TCR signaling (248). However, Notch signaling in some T-lineage contexts can restrain the activation of these E protein antagonists (227). Thus, as T-cell specification is under way, an environment-dependent positive feedback can be established, and Notch and E proteins are able to work together.

However, the roles of TCF-1 and GATA-3 appear to fall into a separate regulatory circuit module or modules from both Notch and E proteins. TCF-1 and GATA-3 are normally activated in parallel by Notch signaling (Figure 3), and they generally operate together. Tcf7 can also be positively regulated by E proteins, as shown in mature T cells (249). Both TCF-1 and GATA-3 play repeated roles in modulation of various T-cell functions during development and later function, usually coexpressed and often required in parallel. However, they are not obligate partners at the target gene level, but rather operate on distinct sets of target genes and provide inputs to the T cell pathway different from those provided by Notch signaling or E protein action. TCF-1 overexpression in precursors can bypass the need for Notch signals to activate a surprisingly large fraction of the T-cell developmental program (169). In multipotent hematopoietic precursor cells forced to express TCF-1 in the absence of Notch ligands, there is substantial activation of Gata3, endogenous Tcf7 and Lef1, and even Bcl11b, along with Il2ra and some Cd3g expression (Fig. 3; note that inputs from Notch and TCF-1 follow “OR” logic). Some of the target gene effects can be due to collaboration of the exogenous TCF-1 with endogenously activated GATA-3 and E proteins. However, the directly Notch-dependent targets Ptcra and Dtx1 remain completely silent, implying that these TCF-1 effects can be divorced from Notch activation pathways.

GATA-3 expression above a certain threshold level is important not only for ETP viability, but also to furnish DN2 cells with the capacity to advance to DN3 stage (225, 243). In complete knockouts, although a small number of cells still manage to reach a DN2-like stage, GATA-3 loss severely reduces their ability to turn on Bcl11b or Cd3e (235). However, GATA-3 loss does not inhibit activation of Notch target genes such as Ptcra or Il2ra (i.e. the CD25 marker that defines cells as “DN2”), yielding an abnormal phenotype with some genes at DN3-like levels while others are ETP-like (235). Thus, GATA-3 is not needed to sustain the T-cell genes activated in direct responses to Notch signals, nor to transduce Notch signaling itself. Furthermore, although GATA-3 binds to a conserved, active enhancer of the Tcf7 gene in early T cells (65), it is not needed for Tcf7 activation. In
this case, although many TCF-1 targets may also require GATA-3, the Tcf7 gene itself may be activated by Notch (169, 170) without GATA-3 (Fig. 3).

GATA-3 can, however, act as an antagonist of T-cell differentiation, rather than as a driver. Evidence actually emerged several years ago, when GATA-3 was tested in gain-of-function scenarios similar to one used successfully to show the positive effects of TCF-1 (114, 250). When higher levels of GATA-3 were introduced to prethymic cells or early ETP/DN2 stage cells, T-cell development was catastrophically inhibited, not promoted. Overexpression did not simply kill the cells, but actually changed their developmental status to drive them out of the T-cell pathway. Survival became possible only if the cells were removed from Notch signaling; then, high-level GATA-3 synergized with the loss of Notch signals to repress Il7r and extinguish the T-cell gene expression profile. It repressed the Tcf7 expression that is normally sustained when Notch signals are withdrawn, suggesting that TCF-1 positive autoregulation might be disabled by high GATA-3 (Fig. 3, dash-dot line). If the cells were kept alive, high-level GATA-3 could convert them into mast cells (114). Thus, T cell identity depends on GATA-3 but also on the preservation of strict limits to GATA-3 expression.

If Gata3 is being activated by Notch signals initially, and then by positive inputs from TCF-1 as well, how can its expression be so carefully restrained during normal T cell development? The answer seems to be by antagonistic feedback from two other essential T-cell regulators: Bcl11b, and E2A (Fig. 3). Deletion of either of these factors causes Gata3 expression to rise (L. Li, J. A. Zhang, and E. V. R., unpublished). Although Bcl11b is activated in a GATA-3 dependent process (235), it also feeds back to limit Gata3 expression (L. Li, J. A. Zhang, and E. V. R., unpublished). In E2A knockout cells, the GATA-3 increase can be enough to disrupt T-cell development: surprisingly, a large part of the E2A-deletion defect can be corrected by shRNA against GATA-3 (243). Thus, E proteins have at least two qualitatively distinct roles in regulating GATA-3 activity. First, they help to turn it on: they may support its expression directly (216, 240), and they also make possible the Notch signaling that turns Gata3 on in the first place and initially protects it from being inhibited by PU.1 (227). Second, they enforce a damping repressive limit on Gata3 expression to prevent it from blocking the T cell pathway (243). Both the positive and the negative roles are critical for T-cell specification.

**Special status of E proteins and Notch in the T and B lymphocyte developmental programs and their alternatives**

E protein activity is the most important common strand that relates the T-cell and B-cell programs at the transcriptional level. Directly attributable to E protein positive regulation are expression of Rag1/Rag2 genes for immunoreceptor gene rearrangement in both lineages. However, in important ways the linkages of E proteins to other regulatory factors in gene regulatory networks are profoundly different, and this is crucial to explain how the two programs can diverge.

In the B cell program, E2A activity participates directly in the feed-forward regulatory network that confers and sustains B-cell identity (124), including the direct regulation of
FoxO, EBF1, and Pax5. It is crucial throughout the bone marrow developmental stages of B-cell development with continuing roles in the periphery (251). Although the strict requirement for E2A in B-cell specification can be bypassed in part by forced expression of EBF1, this is only due to the presence of lower, but potentially functional, levels of the related bHLH factor HEB in E2A-knockout cells (129, 252). The case is different for T cells. In the T-cell program, E protein activity can be dissected from other parts of the specification network circuitry, working against as well as along with the other crucial T-cell factor, GATA-3. Even certain genes like Il7r that receive positive input from E protein activity at early stages can switch to be negatively regulated by E proteins at later stages (239). Working across the decisions mediated by lineage-selection factors like GATA-3 and TCF-1 in T-cell precursors and mature T cells alike, E protein inhibition and release will be used repeatedly as the switch between activated and resting states (239, 248, 253, 254). In particular, while E protein levels will be high during resting and memory states, Id activation to neutralize E proteins will be a broad feature of antigen-triggered T-cell effector activation (255). Thus, E protein dominance need not be a constant in T cell programming. As discussed below, this reversibility of E protein usage gives T cells a particularly close relationship to NK cells and other innate lymphocytes.

**T vs. B lineage decision and its alternatives**

**Double locking mechanism for B cell exclusion of the T cell option**

To block access to the T cell program in B-cell precursors, B cell regulatory factors must target the functions that distinguish the T-cell network from the B cell program. They cannot attack E proteins, since these are needed by both developmental programs. Other points of vulnerability in the T-cell program must be targeted for silencing by EBF1 and Pax5, during the transition to Ly-6d+ “BLP” and pro-B cell stage, as cells make the decision to exclude the T-cell fate (Fig. 4).

Pax5 has long been of interest for its negative roles. The concept of commitment as an important stage in hematopoietic development was galvanized by the demonstration that Pax5-deficient pro-B cells could proliferate in vitro for months or years, yet never fully commit to the B cell fate (256, 257). The cells maintain a clearly B-lineage gene expression profile and even undergo significant Ig gene rearrangement as long as IL-7 is supplied. However, upon IL-7 removal, the cells reveal that they still have access to a variety of myeloid and dendritic cell fates (macrophages, granulocytes, dendritic cells, osteoclasts), and if exposed to Notch ligands, they efficiently switch to the T-cell fate. Thus, although it is turned on too late to explain the initial decrease in myeloid potential from LMPP to CLP, Pax5 is involved in the gene network that confirms loss of access to both the T and myeloid developmental pathways.

Although Pax5-knockout pro-B cells have a ready, spontaneous ability to develop into myeloid cells when IL-7 is removed, the fate which is most severely limited by Pax5 is the T cell fate. Ectopic Pax5 expression in hematopoietic precursors has a particularly strong repressive influence on the ability of multipotent precursors to give rise to T cells, and as B-cell development is amplified it is at the expense of T cell development (258, 259). This can be explained by the finding that Pax5 acts as a potent repressor of Notch1 expression itself.
In this respect, from the moment a B-cell precursor achieves an adequate level of Pax5, it should become the equivalent of a Notch1-deficient precursor, selectively blocked in T-cell potential. However, biasing against T-cell potential in Ly6d+ BLPs begins even before the cells fully express Pax5 (102, 103). Furthermore, even Pax5-deficient B-cell precursors can be completely blocked from T-cell development, provided that their EBF1 is sufficiently activated (159). If Ebf1 is deleted, pro-B cells with the Pax5 gene intact regain the ability to enter the T-cell pathway (260, 261). Thus EBF1 as well as Pax5 plays a role in blocking access to the T-cell program.

Recent evidence shows that EBF1 can downregulate both Gata3 and Tcf7 (260, 261). It binds directly to sites in the Tcf7 gene and at least two sites bracketing the transcriptional start site of Gata3, one or both of which are important for repression function. When EBF1 binding is blocked at these sites, Gata3 expression surges again (260). Notably, EBF1 appears to be repressing here by acting as an “anti-insulator”. The position of the upstream EBF1 site lies at the boundary between the open chromatin of the Gata3 gene and a large block of repressed chromatin (strongly marked with H3K27me3 modification), which constitutively covers the Gata3 “1a” promoter in precursors (65). In T-lineage cells, the main Gata3 “1b” promoter is clear to drive expression. EBF1 binding, however, allows the Gata3 1b promoter and first intron to become buried in repressive marks as well (260), i.e., presumably by allowing the repressive chromatin to spread. Gata3 and Tcf7 repression condemns to death any cell that may be embarking on the T cell program in response to Notch signals. Thus, EBF1 and Pax5 complementarily block access to the T cell fate, first by preventing Notch signaling from turning on the T-cell regulatory program even in precursors that still express substantial Notch1, and then by removing Notch1 itself (Fig. 4).

Divergence of T-cell precursors from CLP: undoing a B-cell promoting state

To avoid these repressive events, T-cell precursors themselves must apparently forego using the early B-lineage events broadly as a “head start”, even in cases where later they will turn on the same genes. It is likely that the initial expression of Ets1, Ikaros, Runx1, Myb, and even factors like PU.1 in LMPPs and CLPs may be equally efficient as a start for both. However, in specific ways the T-cell precursors must start by weakening the B-lineage biases that may already be imposed by expression of IL-7R and E proteins, and revealed by expression of Rag1.

Even strong IL-7R expression from the earliest stage can be a problem for T-cell precursors. Indeed, forced expression of either IL-7R or constitutively activated STAT5 can redirect ETP differentiation toward the B-cell fate even within the thymus (262). Interestingly, as noted above, both Il7r and Rag1/Rag2 are activated in early B-lineage cells in a FoxO1-dependent way. Thus, one way to downregulate both genes might be simply to switch growth factor receptor dependence so as to allow Akt activation to inhibit FoxO1. PI 3-kinase activation by signals from the growth factor receptor Kit, which is vital for early T-cell development (263), could play a role in establishing this condition (264), until the B-cell pathway is fully inaccessible.

The relationship of E proteins to Notch signals must be altered in order to allow the T lineage to develop. In the B cell program, Notch signaling itself has the power to antagonize...
E2A itself, by inducing degradation of the E2A protein (265). This mechanism has been invoked as one of the barriers to B cell development in the thymus. Indeed, Notch signaling applied to CLP may initially prime the cells for development into innate type lymphocytes, such as NK cells and lymphoid tissue inducer cells (discussed further below), i.e., differentiation fates which depend on antagonism or loss of E protein function (266, 267). Why does this not occur in developing T-cell precursors, then, or how do they escape from the late ETP stage on? One part of the answer is the regulation of the components that work downstream of Notch to target E2A for degradation. The critical step in Notch-triggered E2A degradation is the phosphorylation of E2A in a conserved domain by a MAP kinase (265). B cell precursors are richly endowed with MAP kinase activity. Peculiarly, though, T cell precursors maintain unusually low levels of MAP kinase throughout the specification stages (268). Thus E2A protein becomes much less sensitive to degradation by Notch at an early stage of the T-cell program, enabling it to establish its positive collaborative circuit with Notch described above (Fig. 3).

**T cell exclusion of the B cell option**

The crucial E2A target gene for activation of the B cell program is EBF1, as we have seen. Both FoxO1 and STAT5a/b have been invoked as collaborators in this process, yet neither one is excluded from the T-cell program. STAT5a and b are well expressed and activated by IL-7R signaling in the DN2a/2b stages particularly; and although FoxO1 expression is higher in B-lineage precursors, T-lineage cells compensate with high expression of its close relative, FoxO3a. Thus, all the expected requirements for E2A activation of EBF1 are there. However, despite co-activation of these potential collaborators with E2A in the early T-cell program, EBF1 and Pax5 are never activated from the closed state in chromatin in which they reside in multipotent precursors and stem cells (65, 269). One possibility is that the E2A levels simply fail to climb high enough in the early T-cell program to activate EBF1; it is true that Tcf3 (or Tcfe2a, codes for E2A) transcript levels are notably stable throughout T-cell development, ~3× fold lower than their peaks in pro-B and pre-B cells. This might suffice if EBF1 activation by E2A were strictly dose dependent, more so than other target genes. There are, however, variant CLP with particularly low E2A:Id2 ratios that are reported to be able to generate B cells and NK cells but not T cells (98). Extrapolating from the behavior of these cells, the E protein-dependent functions in T cell development would require even higher doses of E protein activity than those in B-cell development, not lower ones. A more satisfying explanation may therefore be that one of the early-activated genes in the T cell program actively interferes with the ability of E2A to activate EBF1, or directly singles EBF1 out for silencing.

The timing of the loss of B-cell potential in developing T cells implies that one of the earliest Notch-induced regulatory changes in these cells creates an intrinsic barrier to the B-cell program, long before the cells make the transition from ETP to DN2a (Fig. 4). Three T-lineage regulatory candidates that are induced strongly in this period are Hes1, TCF-1, and GATA-3. Hes1 is required for T-cell specification and population survival, but Hes1 conditional knockout cells do not expand the population of B cells in the thymus, showing that the inhibition of B-cell development by Notch signaling remains intact (270); and Hes1 is actually expressed in early B lineage precursors. Tcf7 (encoding TCF-1) knockout cells
are also severely blocked in T cell development, but they do not display increased B or myeloid potential (169, 170). The later lineage commitment factor Bcl11b is expressed too late to account for B-cell exclusion (111, 182), and does not contribute to this aspect of commitment (225). It is GATA-3, rather, that appears to have an indispensable activity as an antagonist of B-cell development (235). The extreme impacts of Gata3 deletion on early T-lineage survival and proliferation have made it necessary to exclude alternative interpretations assiduously (237), but the evidence seems strong. The cells that can give rise to B cells may proceed as far as turning on Tcf7 and Notch target genes, showing that they are not deficient at entering the T-cell pathway (235). They become arrested in a DN2a-like state, as their expression of Bcl11b and the Cd3 genes remains low. However, it is the loss of Gata3 and not the loss of Bcl11b that restores access to the B-lineage program, as shown by analysis of T-cell precursors from single and double conditional deletion mutants (225).

The exact mechanism through which GATA-3 is working to prohibit B-lineage specification is still under investigation, but its timing of expression and functional impact imply that it is the crucial mediator.

**Common vs. exclusive transcription factor modules: linkages to “innate” lymphoid alternatives**

The parallel aspects of the T and B cell programs reflect their shared use of gene rearrangement and selection in a particular order to enable the cells to acquire recognition specificities, and as already described in both lineages these shared features are regulated at least in part by E protein activity. Some of the most discordant aspects have to do with function. Programming for B-cell function is mostly concerned with making the cells responsive to Ig-based signals that trigger cellular decisions to proliferate, survive or die, and to express different variants of the cell’s immunoglobulin genes (some of which require additional rounds of somatic mutagenesis). The same genes that confer B cells’ main recognition specificity are the genes that confer most of their effector function. In contrast, the programming for T-cell function requires transcriptional poising and regulation of a considerably broader set of effector-related gene loci in addition to the genes that regulate recognition, proliferation, survival, and apoptosis decisions.

Importantly, these effector functional elements are not unique to T cells, but they are not shared with B cells; instead, they are shared with the newly appreciated group of non-T, non-B cells called innate lymphocytes (ILC), which do not use RAG-based gene rearrangement in their developmental programs. This creates a disjunction between recognition-based and effector-based criteria for defining the closest relatives of T cells, and this poses a challenge to any uniquely intimate relationship between T and B cell programs.

The ILC lineages of lymphocytes depend not on E protein activity but on the neutralization of E proteins by Id2 in order to develop (Fig. 4). ILCs include not only natural killer (NK) cells, similar to cytotoxic T lymphocytes, but also a range of other cytokine-producing innate lymphocyte subsets corresponding roughly to Th2 and Th17 T cells (271, 272). None of these correspond to “innate B cells” – it appears that the B cell program has too direct and intimate an involvement of E proteins in every aspect of cellular function to possess an “innate” equivalent. However, the T cell effector programs are not E protein dependent in
the same way. They depend on GATA3, TCF-1/LEF-1, Runx factors, T-bet (Tbx21) or Eomesodermin (Eomes), Tox, the nuclear receptor Rorγt (Rorc), Maf, Batf, Ahr, and other factors in different combinations. To a large extent, these are the same factors which the innate-type lymphocytes deploy for their own corresponding functional programs. These effector transcriptional programs are thus separable from the developmental use of E proteins, which play their main T-cell developmental roles in setting up TCR gene rearrangement and selection checkpoints.

The common thread linking all the ILCs is the fact that they all depend on Id2 (Fig. 4). Because of their neutralization of E protein activity, ILCs do not express Rag genes. However, this does not give them a profoundly different heritage from T and B precursors. In fact, lineage tracing has shown that most NK cells descend from Rag1+ precursors (96). Thus, it is important to note that Id2 is not properly a lineage-specific identifier. It is readily inducible in common lymphoid precursors and early T-cell precursors until the activation of definitive commitment factors EBF1 and Bcl11b, respectively. Even after T-lineage commitment, it is regularly, albeit transiently, upregulated in mature T cell function (as Bcl11b levels temporarily drop) in the innate-like “NKT pathway and when an effector response is induced. Here, the modularity of the T-cell program is a great help to understand its relationship to ILC developmental lineages. As late as the DN2a stage, signals from IL-33 can redirect the effects of factors like GATA-3 and TCF-1 to work in the ILC2 program (Fig. 2), via induction of Id2 and the orphan nuclear receptor RORA (211, 273–275). Thus, while repression of PU.1 helps to explain the loss of dendritic cell or myeloid access during commitment, it is easiest to explain the loss of access to the innate lymphocyte fates by forcible imposition and protection of E protein activity, at a minimum in two phases to last just long enough for TCR gene rearrangement and TCR-based selection to be complete.

The connection between T and innate cell developmental programs is thus arguably much closer than that between the T cell program and the B cell program. Three different innate cell developmental programs share major aspects of their regulatory programs with three different branches of the T-cell development pathway. Nevertheless, the very bridge between the T-cell and innate cell programs is based on eliminating exactly those regulatory functions, the E protein dependent expression of Rag genes, that T and B cells most uniquely share. Thus, the regulatory factors and regulatory circuit modules that determine lymphocyte identity do not act in a simple hierarchy of relatedness, but rather in a matrix of regulatory interactions, which crosses valuable cellular features to generate a richly diversified immune functional repertoire.

**Concluding Remarks**

The T-cell and B-cell programs are linked by their shared use of E protein-dependent genes which include the functions they share for immunoreceptor gene rearrangement as well as lineage-specific functions. Their common progenitors share the need for IL-7 receptor signaling, which provides unique support functions, and the need for E protein activation and protection from inhibitors. Both the T-cell program and the B cell program use repressors of the E-protein antagonist Id2, Bcl11b and EBF1 respectively, at pivotal stages
in their lineage commitment processes. However, the architectures of the gene networks in which these lineage commitment factors work are distinct. In the case of the B-cell program, E protein use is also tightly interlocked with lineage-specific factors in positive feedback and feed-forward circuits that control direct specification functions. In the case of the T-cell program, E protein use is modular and capable of being unlinked from the circuitry that maintains T-cell identity factors such as GATA-3 and TCF-1. Transcription factor deployment generates cell type identity, but transcription factors can be used in different combinations and may be expressed in cycles of repression and reactivation. There is much still to be learned to clarify the mechanisms involved, particularly to connect factor expression to profiles of genome-wide binding to transcriptional function, and to reveal the ways that environmental signals impinge on these relationships. The results already emphasize that single factors must be viewed not as “master regulators” but as participants in a richly endowed and interlinked regulatory context. To understand the relationship between lymphocyte developmental programs, it is important to understand the operating rules for the transcriptional regulatory networks at their hearts as well as the components from which they are built.

Acknowledgments

I am grateful to many colleagues for inspiring this paper, and I especially wish to thank the attendees of the FASEB Conferences on Molecular Mechanisms of Lymphocyte Development and Function and the Aegean Conferences on Gene Regulation in Lymphocyte Development for discussions, clarifications, and insights. I also thank members of my group for letting me refer to their own work before publication. Development of these ideas was supported by NIH grants CA148278, CA90233, HL89123, AI083514 and AI95943 and by the Albert Billings Ruddock Professorship.

REFERENCES


Anna Rev Immunol. Author manuscript; available in PMC 2015 March 21.


105. Weerkamp F, Baert MR, Brugman MH, Dik WA, de Haas EF, Visser TP, de Groot CJ, Wagemaker G, van Dongen JJ, Staal FJ. Human thymus contains multipotent progenitors with...


Annu Rev Immunol. Author manuscript; available in PMC 2015 March 21.


264. Moller C, Alfredsson J, Engstrom M, Wootz H, Xiang Z, Lennartsson J, Jonsson JI, Nilsson G. Stem cell factor promotes mast cell survival via inactivation of FOXO3a-mediated transcriptional...


FIGURE 1.

Schematic of major stages of B and T cell development. Consult Table 1 for definition of stage phenotypes. The figure introduces key stages and emphasizes the parallelism between B cell development stages and αβ T lineage cell stages in terms of immunoreceptor gene rearrangement timing, proliferative bursts and major developmental checkpoints. Specific regulatory genes important for lineage specification are turned on during the intervals shown. Stages are defined by ability to discriminate phenotypes and do not represent uniform lengths of time or numbers of cell cycles. Note that the T cell program unlike the B cell program generates at least five distinct types of T cells within the thymus (in fact the TCRγδ cells are further subdivided, not shown). “Thymus settling” = thymus settling precursors, which are thought to be derived from “ALP” type common lymphoid precursors and/or from certain LMPP or other ELP type cells in vivo. All of these, and even myeloid specified cells, can respond when introduced into the thymus by developing into T cells.
FIGURE 2.
Major alternative fate branchpoints for B and T cell precursors. The figure explains the timing of commitment, by showing the demonstrated alternative fates that can still be adopted by B and T-cell precursors until the latest stages shown. Heavy line arrows: major pathways. Light solid line arrows: strong pathways, high precursor frequencies for the indicated fate alternative. Light dashed line arrows: variant pathways, not a default in vivo but readily demonstrable at high frequency under experimental conditions. Dotted line arrows: measurable experimentally but distinctly reduced precursor frequency for the indicated path at this stage as compared to immediate precursor. Mac, DC: macrophage and dendritic cells. In addition, granulocyte fates are also accessible to pre-commitment early T cells. For simplicity, within the T-cell lineage, fates that are still robustly accessible at the DN2a stage are not also shown for the antecedent thymus settling and ETP stages. BLP and DN2b cells are largely if not completely committed.
Partial gene regulatory network for early T-cell specification. Horizontal lines: genes coding for regulatory factors. Bent arrows: transcription and translation of gene leading to product (itself a regulatory factor). Genes and their products have the same color code. Arrows: positive regulatory effects of factors on indicated genes. Blocked-end lines: negative regulatory effects of factors on indicated genes. Filled circle: ability of E2A to form a functionally active dimer: this assembly is blocked by Id2. Star: activation of Notch transmembrane protein by binding to its ligand Delta-like 4 (DLL4) on a neighboring cell; the effect is to cleave the Notch intracellular domain and allow it to be transported to the nucleus where it functions as a transcriptional coactivator on the indicated target genes. Note that none of these regulators acts in a strictly all-or-nothing way. Although many genes require Notch input for activation, those shown here do not require its continuation for maintenance of expression. The repressors indicated here modulate rather than silence the indicated target genes. Solid lines: relationships with strong molecular and functional support within the context of T-lineage specification. Dashed lines: relationships seen in related cells but not directly demonstrated in DN T-cell precursors. Long dash-dot lines: inferred effect of GATA-3 on Tcf7 when GATA-3 is overexpressed. Details and sources are given in the text.
FIGURE 4.

Key regulatory targets for program maintenance and alternative program exclusion in development of T, B, and innate lymphocytes. The framework is a simplified version of Fig. 2. Green: stages of B-cell development in which EBF1 and Pax5 expression is dominant; i.e. the EBF1 and Pax5 “territory”. Light blue: stages of T-cell development in which GATA-3 and TCF-1 are dominant, i.e. the GATA3 and TCF-1 “territory”. Violet: Id2 “territory” of programs that generate innate cells. Note that the influence of Notch signaling (dark blue frame) does not extend throughout T-cell development but does extend to some ILCs. GATA-3 and TCF-1 also help generating certain ILCs, including a major type of NK cells. The figure depicts the selective cross-repression between the components of B and T-cell programs that uniquely distinguish them from each other. In contrast, E proteins including E2A are components of both.
## TABLE 1

### Phenotypic definitions of lymphoid precursor subsets

<table>
<thead>
<tr>
<th>Cell type</th>
<th>Name (if different)</th>
<th>Major markers</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>HSC</td>
<td>Hematopoietic stem cell</td>
<td>Lin- Kith&lt;sup&gt;high&lt;/sup&gt; Sca-1&lt;sup&gt;high&lt;/sup&gt; Flt3&lt;sup&gt;−&lt;/sup&gt; CD34&lt;sup&gt;int&lt;/sup&gt;- CD150&lt;sup&gt;−&lt;/sup&gt; CD48&lt;sup&gt;−&lt;/sup&gt;</td>
<td>These cells still have erythroid and megakaryocyte potential</td>
</tr>
<tr>
<td>MPP</td>
<td>Multipotent precursor</td>
<td>Lin- Kith&lt;sup&gt;high&lt;/sup&gt; Sca-1&lt;sup&gt;high&lt;/sup&gt; Flt3&lt;sup&gt;−&lt;/sup&gt; CD34&lt;sup&gt;−&lt;/sup&gt; and/or CD150&lt;sup&gt;−&lt;/sup&gt; CD48&lt;sup&gt;−&lt;/sup&gt;</td>
<td>For lymphoid precursor activity, Vcam&lt;sup&gt;−&lt;/sup&gt; Flt3&lt;sup&gt;very high&lt;/sup&gt; are important</td>
</tr>
<tr>
<td>LMPP</td>
<td>Lymphoid-primed multipotent precursor</td>
<td>Lin- Kith&lt;sup&gt;high&lt;/sup&gt; Sca-1&lt;sup&gt;high&lt;/sup&gt; Flt3&lt;sup&gt;very high&lt;/sup&gt; CD27&lt;sup&gt;+&lt;/sup&gt; Vcam&lt;sup&gt;−&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>ELP</td>
<td>Early lymphoid precursor</td>
<td>Like LMPP–ALP, Terminal transferase&lt;sup&gt;a&lt;/sup&gt;, defined by Rag1 reporter transgene</td>
<td>Rag1 expression in LMPP/ALP type cells is low or intermittent</td>
</tr>
<tr>
<td>CLP or “ALP”</td>
<td>Common lymphoid precursor “A” type</td>
<td>Lin- Kith&lt;sup&gt;low&lt;/sup&gt; Sca-1&lt;sup&gt;low&lt;/sup&gt; Flt3&lt;sup&gt;very high&lt;/sup&gt; CD27&lt;sup&gt;+&lt;/sup&gt; IL7R&lt;sup&gt;+&lt;/sup&gt; Ly6d&lt;sup&gt;−&lt;/sup&gt;</td>
<td>These are the true “common” lymphoid precursors</td>
</tr>
<tr>
<td>BLP</td>
<td>B-cell biased common lymphoid precursor</td>
<td>Lin- Kith&lt;sup&gt;low&lt;/sup&gt; Sca-1&lt;sup&gt;low&lt;/sup&gt; Flt3&lt;sup&gt;int&lt;/sup&gt; CD27&lt;sup&gt;+&lt;/sup&gt; IL7R&lt;sup&gt;+&lt;/sup&gt; Ly6d&lt;sup&gt;−&lt;/sup&gt;</td>
<td>Within this stage, Rag1 is upregulated, a λ5 transgene begins to be activated, and EBF1 is turned on</td>
</tr>
<tr>
<td>Pre-Pro B</td>
<td>B220&lt;sup&gt;−&lt;/sup&gt; CD43&lt;sup&gt;−&lt;/sup&gt; CD19&lt;sup&gt;−&lt;/sup&gt; Kith&lt;sup&gt;low&lt;/sup&gt; Flt3&lt;sup&gt;low&lt;/sup&gt; IL7R&lt;sup&gt;+&lt;/sup&gt; Ly6d&lt;sup&gt;−&lt;/sup&gt;</td>
<td>Continuum from late BLP marked by B220 upregulation and EBF1 and Pax5 up; caution that without Ly6d marker, NK cells have similar phenotype</td>
<td></td>
</tr>
<tr>
<td>Pro-B</td>
<td>B220&lt;sup&gt;−&lt;/sup&gt; CD43&lt;sup&gt;−&lt;/sup&gt; CD19&lt;sup&gt;−&lt;/sup&gt; Kith&lt;sup&gt;low&lt;/sup&gt; Flt3&lt;sup&gt;low&lt;/sup&gt; IL7R&lt;sup&gt;+&lt;/sup&gt; Ly6d&lt;sup&gt;−&lt;/sup&gt;</td>
<td>Pax5 signature; rearranging Ig heavy chain; Flt3 and CD27 off</td>
<td></td>
</tr>
<tr>
<td>Large pre-B</td>
<td>B220&lt;sup&gt;−&lt;/sup&gt; CD19&lt;sup&gt;−&lt;/sup&gt; CD43&lt;sup&gt;−&lt;/sup&gt; internal IgH&lt;sup&gt;+&lt;/sup&gt;, surface IgM&lt;sup&gt;−&lt;/sup&gt;</td>
<td>Kit turns off; cells with good IgH rearrangement cycle</td>
<td></td>
</tr>
<tr>
<td>Small pre-B</td>
<td>B220&lt;sup&gt;−&lt;/sup&gt; CD19&lt;sup&gt;−&lt;/sup&gt; CD43&lt;sup&gt;−&lt;/sup&gt;, internal IgH&lt;sup&gt;−&lt;/sup&gt;, Surface IgM&lt;sup&gt;−&lt;/sup&gt;</td>
<td>Cells stop proliferating, rearrange Igκ and λ</td>
<td></td>
</tr>
<tr>
<td>B cell</td>
<td>B220&lt;sup&gt;−&lt;/sup&gt; CD19&lt;sup&gt;−&lt;/sup&gt; CD43&lt;sup&gt;−&lt;/sup&gt; Surface IgM&lt;sup&gt;+&lt;/sup&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thymus settling</td>
<td>Thymus settling cell (not well defined)</td>
<td>CD4&lt;sup&gt;α&lt;sub&gt;β&lt;/sub&gt;&lt;/sup&gt; C&lt;sub&gt;D&lt;/sub&gt;8&lt;sup&gt;−&lt;/sup&gt; C&lt;sub&gt;D&lt;/sub&gt;24&lt;sup&gt;−&lt;/sup&gt; C&lt;sub&gt;D&lt;/sub&gt;25&lt;sup&gt;−&lt;/sup&gt; C&lt;sub&gt;D&lt;/sub&gt;4&lt;sup&gt;high&lt;/sup&gt;, Kith&lt;sup&gt;very high&lt;/sup&gt;, Ccr9&lt;sup&gt;−&lt;/sup&gt;, C&lt;sub&gt;D&lt;/sub&gt;27&lt;sup&gt;high&lt;/sup&gt;, Flt3&lt;sup&gt;−&lt;/sup&gt;</td>
<td>Inferential phenotype</td>
</tr>
<tr>
<td>ETP</td>
<td>Early T-cell precursor</td>
<td>CD4&lt;sup&gt;α&lt;sub&gt;β&lt;/sub&gt;&lt;/sup&gt; C&lt;sub&gt;D&lt;/sub&gt;8&lt;sup&gt;−&lt;/sup&gt; C&lt;sub&gt;D&lt;/sub&gt;24&lt;sup&gt;−&lt;/sup&gt; C&lt;sub&gt;D&lt;/sub&gt;25&lt;sup&gt;−&lt;/sup&gt; C&lt;sub&gt;D&lt;/sub&gt;4&lt;sup&gt;high&lt;/sup&gt;, Kith&lt;sup&gt;very high&lt;/sup&gt;, Ccr9&lt;sup&gt;−&lt;/sup&gt;, C&lt;sub&gt;D&lt;/sub&gt;27&lt;sup&gt;high&lt;/sup&gt;</td>
<td>Not yet committed; good TCRγδ and TCRαβ potential</td>
</tr>
<tr>
<td>DN2a</td>
<td>C&lt;sub&gt;D&lt;/sub&gt;8&lt;sup&gt;−&lt;/sup&gt; C&lt;sub&gt;D&lt;/sub&gt;24&lt;sup&gt;−&lt;/sup&gt; C&lt;sub&gt;D&lt;/sub&gt;25&lt;sup&gt;−&lt;/sup&gt; C&lt;sub&gt;D&lt;/sub&gt;4&lt;sup&gt;high&lt;/sup&gt;, Kith&lt;sup&gt;very high&lt;/sup&gt;, Ccr9&lt;sup&gt;−&lt;/sup&gt;, C&lt;sub&gt;D&lt;/sub&gt;27&lt;sup&gt;high&lt;/sup&gt;}</td>
<td>T lineage committed; good TCRγδ and TCRαβ potential</td>
<td></td>
</tr>
<tr>
<td>DN2b</td>
<td>C&lt;sub&gt;D&lt;/sub&gt;8&lt;sup&gt;−&lt;/sup&gt; C&lt;sub&gt;D&lt;/sub&gt;24&lt;sup&gt;−&lt;/sup&gt; C&lt;sub&gt;D&lt;/sub&gt;25&lt;sup&gt;−&lt;/sup&gt; C&lt;sub&gt;D&lt;/sub&gt;4&lt;sup&gt;high&lt;/sup&gt;, Kith&lt;sup&gt;very high&lt;/sup&gt;, Ccr9&lt;sup&gt;−&lt;/sup&gt;, C&lt;sub&gt;D&lt;/sub&gt;27&lt;sup&gt;high&lt;/sup&gt;}</td>
<td>TCRβ rearranging; cells biased against TCRγδ</td>
<td></td>
</tr>
<tr>
<td>DN3a</td>
<td>C&lt;sub&gt;D&lt;/sub&gt;8&lt;sup&gt;−&lt;/sup&gt; C&lt;sub&gt;D&lt;/sub&gt;24&lt;sup&gt;−&lt;/sup&gt; C&lt;sub&gt;D&lt;/sub&gt;25&lt;sup&gt;−&lt;/sup&gt; C&lt;sub&gt;D&lt;/sub&gt;4&lt;sup&gt;high&lt;/sup&gt;, Kith&lt;sup&gt;very high&lt;/sup&gt;, Ccr9&lt;sup&gt;−&lt;/sup&gt;, C&lt;sub&gt;D&lt;/sub&gt;27&lt;sup&gt;high&lt;/sup&gt;}</td>
<td>Large proliferating transitional intermediate following success in TCRβ rearrangement</td>
<td></td>
</tr>
<tr>
<td>DN3b</td>
<td>C&lt;sub&gt;D&lt;/sub&gt;8&lt;sup&gt;−&lt;/sup&gt; C&lt;sub&gt;D&lt;/sub&gt;24&lt;sup&gt;−&lt;/sup&gt; C&lt;sub&gt;D&lt;/sub&gt;25&lt;sup&gt;−&lt;/sup&gt; C&lt;sub&gt;D&lt;/sub&gt;4&lt;sup&gt;high&lt;/sup&gt;, Kith&lt;sup&gt;very high&lt;/sup&gt;, Ccr9&lt;sup&gt;−&lt;/sup&gt;, C&lt;sub&gt;D&lt;/sub&gt;27&lt;sup&gt;high&lt;/sup&gt;}</td>
<td>Large proliferating transitional intermediate</td>
<td></td>
</tr>
<tr>
<td>DN4</td>
<td>C&lt;sub&gt;D&lt;/sub&gt;8&lt;sup&gt;−&lt;/sup&gt; C&lt;sub&gt;D&lt;/sub&gt;24&lt;sup&gt;−&lt;/sup&gt; C&lt;sub&gt;D&lt;/sub&gt;25&lt;sup&gt;−&lt;/sup&gt; C&lt;sub&gt;D&lt;/sub&gt;4&lt;sup&gt;high&lt;/sup&gt;, Kith&lt;sup&gt;very high&lt;/sup&gt;, Ccr9&lt;sup&gt;−&lt;/sup&gt;, C&lt;sub&gt;D&lt;/sub&gt;27&lt;sup&gt;high&lt;/sup&gt;}</td>
<td>Large proliferating transitional intermediate</td>
<td></td>
</tr>
<tr>
<td>ISP</td>
<td>Immature single positive</td>
<td>CD6&lt;sup&gt;−&lt;/sup&gt; or CD4&lt;sup&gt;−&lt;/sup&gt;, CD24&lt;sup&gt;−&lt;/sup&gt;, CD28&lt;sup&gt;−&lt;/sup&gt;, surface TCR&lt;sup&gt;−&lt;/sup&gt;, internal TCRβ&lt;sup&gt;−&lt;/sup&gt;</td>
<td>Large proliferating transitional intermediate</td>
</tr>
<tr>
<td>DP</td>
<td>Double positive</td>
<td>CD4&lt;sup&gt;+&lt;/sup&gt; CD8&lt;sup&gt;−&lt;/sup&gt; TCRβ&lt;sup&gt;−&lt;/sup&gt;, CD24&lt;sup&gt;−&lt;/sup&gt;</td>
<td>These cells acquire TCRα chains and cell-surface mature TCR complexes during the DP stage</td>
</tr>
<tr>
<td>Cell type</td>
<td>Name (if different)</td>
<td>Major markers</td>
<td>Comments</td>
</tr>
<tr>
<td>-----------</td>
<td>---------------------</td>
<td>---------------</td>
<td>----------</td>
</tr>
<tr>
<td>CD4 SP</td>
<td>CD4 single-positive</td>
<td>CD4+ CD8⁻ TCRαβ⁺ CD24⁻ to negative</td>
<td></td>
</tr>
<tr>
<td>CD8 SP</td>
<td>CD8 single positive</td>
<td>CD4⁻ CD8⁺ TCRαβ⁺ CD24⁻ to negative</td>
<td></td>
</tr>
<tr>
<td>NKT</td>
<td>Invariant NK-like T cell</td>
<td>CD4⁺⁻, CD8⁻ TCRαβlow, CD24 negative, CD44⁺, NK1.1⁺</td>
<td>Invariant TCR used; NK marker expression depends on allele</td>
</tr>
<tr>
<td>Treg</td>
<td>Natural (thymus-derived) regulatory T cell</td>
<td>Effectively identified with Foxp3 reporter gene</td>
<td></td>
</tr>
</tbody>
</table>