Check for updates

How transcription factors drive choice of the T cell fate

Hiroyuki Hosokawa[1,2] *and Ellen V. Rothenberg*[2] ✉

Abstract | Recent evidence has elucidated how multipotent blood progenitors transform their identities in the thymus and undergo commitment to become T cells. Together with environmental signals, a core group of transcription factors have essential roles in this process by directly activating and repressing specific genes. Many of these transcription factors also function in later T cell development, but control different genes. Here, we review how these transcription factors work to change the activities of specific genomic loci during early intrathymic development to establish T cell lineage identity. We introduce the key regulators and highlight newly emergent insights into the rules that govern their actions. Whole-genome deep sequencing-based analysis has revealed unexpectedly rich relationships between inherited epigenetic states, transcription factor–DNA binding affinity thresholds and influences of given transcription factors on the activities of other factors in the same cells. Together, these mechanisms determine T cell identity and make the lineage choice irreversible.

Notch pathway
Notch designates a cell-surface receptor (NOTCH1–NOTCH4 family in mammals) that interacts with cell-bound ligands of the Delta (Delta-like in mammals) and Serrate (Jagged in mammals) families. Originally discovered through its potent role in fruit fly development, Notch signalling controls important embryological switch points for the generation of various cell types in organisms of all kinds.

[1]*Department of Immunology, Tokai University School of Medicine, Isehara, Kanagawa, Japan.*

[2]*Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA, USA.*

✉*e-mail: evroth@its.caltech.edu*

https://doi.org/10.1038/s41577-020-00426-6

The T cell developmental programme is initiated by a prolonged iteration of environmental signals within the thymus and by the changes in transcription factor expression that are induced in a stepwise manner by those signals. Genetic and molecular biological evidence has identified many of these regulatory factors, and recent advances have shed light on the mechanisms that underlie their actions across the genome to result in the generation of T cell precursors. This Review introduces the main transcriptional regulators of T cell development and summarizes recent evidence for the ways in which they work. The results indicate general lessons that can be applied more broadly to immune cell development. As described herein, developing T cells need to balance precursor expansion with quality-controlled differentiation. This requirement is met by a system in which chromatin-based epigenetic constraints and transcription factor interactions modulate the activities of the factors involved, which not only regulate the speed of differentiation but also give the factors highly stage-specific roles. Similar mechanisms might operate in other systems of immune cell development that are not yet as well understood.

This Review introduces the main known transcriptional regulators that initiate T cell development. It describes work mainly in mice, where extensive test perturbations have been performed, which has shed light on the mechanisms of action of these factors. We connect the actions of the ensemble of these factors with the two major processes that determine the pace and progression of T cell lineage differentiation. Importantly, evidence shows that the transcription factors work in highly context-dependent manners, as their activities are modulated, potentiated and redeployed by interactions with chromatin states and with each other. These interactions contribute strongly to the irreversibility of T cell lineage commitment.

Overview of early T cell development

Multipotent blood progenitor cells launch the T cell programme when they migrate into the thymus and receive Notch pathway signals from the thymic microenvironment. The thymus is specialized to make T cells, and its stroma provides extracellular signals to progenitor T cells, including stem cell factor (SCF; also known as KIT ligand), FLT3 ligand, IL-7 and Notch ligands[1,2]. T cell development begins in cells that lack expression of both CD4 and CD8, known as double-negative (DN) cells, which subsequently acquire CD4 and CD8 expression to become double-positive (DP) cells and then differentiate into mature CD4 or CD8 single-positive (SP) cells (FIG. 1).

Each day, very few progenitor cells migrate into the thymus, but they proliferate extensively while initiating the T cell differentiation programme. Next, they undergo T cell lineage commitment, followed by T cell receptor (TCR) gene rearrangements at the DN and DP stages[3]. DP cells that successfully express a functional αβ TCR are subjected to positive selection and negative selection, then differentiate into $CD4^+$ T helper cells (T_H cells) or $CD8^+$ cytotoxic T cells[4]. The early differentiation transitions are accompanied by extensive proliferation, which

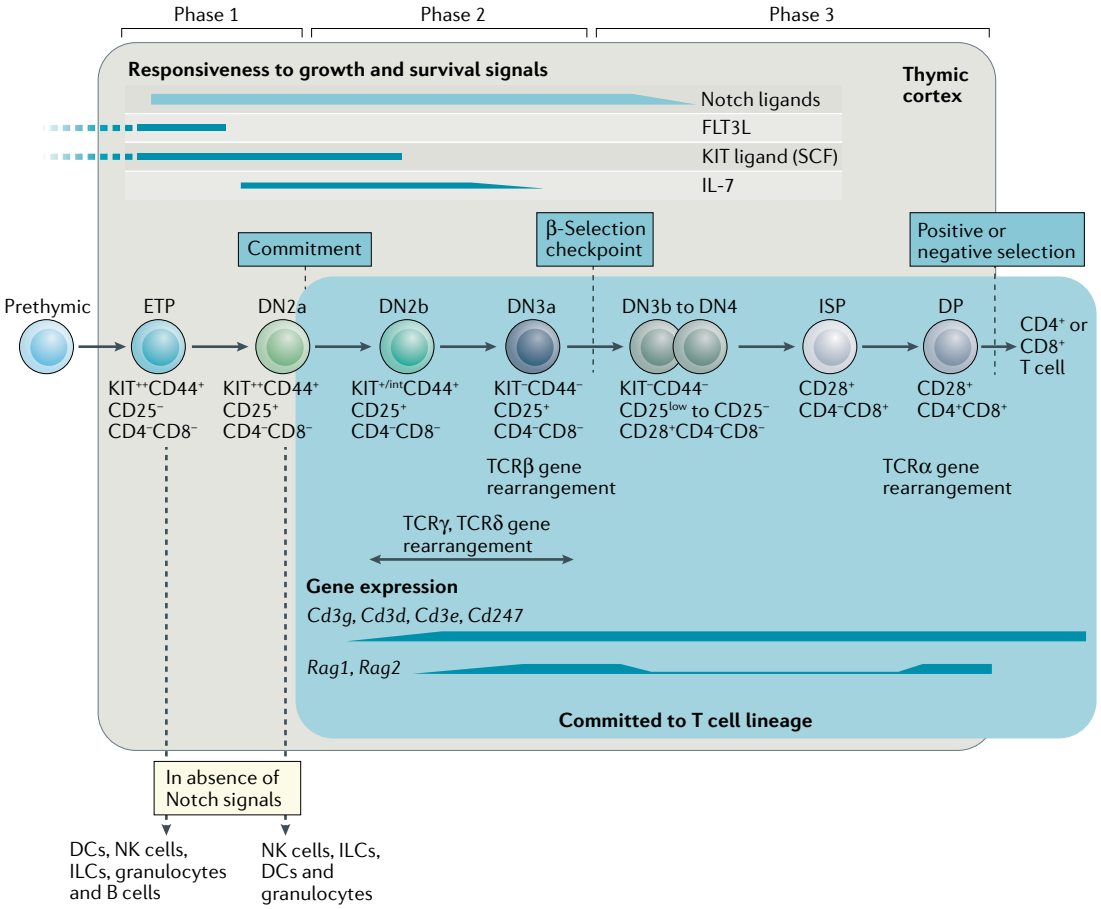

Fig. 1 | **Early T cell developmental stages in mice.** Stages of development within the mouse thymus are shown up to T cell receptor-αβ (TCRαβ)-dependent positive and negative selection, including the developmental checkpoints and major cell-surface phenotype markers. Phases of responsiveness to growth and survival signals from the environment are indicated. The approximate timing of TCR gene rearrangements is shown below the cells; 'TCRβ gene rearrangement' indicates the stage at which V–DJβ rearrangement can produce a complete TCRβ chain. Expression of the genes encoding RAG1 and RAG2 recombinases and components of the CD3 complex is also depicted. Blue shading shows the stages after commitment to the T cell lineage, as defined by the loss of ability to generate non-T cells when placed in an alternative lineage-promoting environment. This coincides with expression of BCL-11B at the single-cell level[159]. The main alternative fates that are accessible to the indicated precommitment cell types if they are withdrawn from Notch signalling are shown. The ability to generate B cells seems to be confined to the most immature early T cell progenitors (ETPs); cells up to the double-negative 2 (DN2) stage are also reported to generate mast cells and macrophages under Notch withdrawal conditions (not depicted). Commitment to the T cell lineage occurs at a considerably earlier stage for fetal thymocytes than for postnatal thymocytes (reviewed in REFS[2,164]). $^{+/int}$, intermediate positive level; DC, dendritic cell; DP, double positive; ILC, innate lymphoid cell; ISP, immature single-positive cell; NK cell, natural killer cell; SCF, stem cell factor.

Positive selection
Once immature thymocytes at the double-positive stage have expressed a complete T cell receptor-αβ (TCRαβ) complex, the cells are doomed to die unless that TCR can interact with cell-surface molecules on thymic epithelial cells. Positive selection is the TCR-dependent rescue of the cells from death and the choice of helper or killer fate that results from that rescue.

Negative selection
If the newly expressed T cell receptor (TCR) on double-positive and immature single-positive thymocytes interacts too strongly with surface molecules on thymic antigen-presenting cells, the thymocytes are induced to commit suicide rather than enabled to survive and mature. Negative selection designates this TCR stimulation-dependent suicide.

generates a large number of distinct TCR$^+$ cells so that later steps in the thymus can positively select those with useful TCR recognition specificities and consign the rest to death through negative selection. However, the expression of TCR itself depends on recombination processes that are confined to non-cycling cells[5]. Therefore, the T cell programme must not only equip cells with the expression of genes required for TCR activity but also allow time for extensive proliferation to occur before cells express the TCR.

DN thymocytes in mice are divided into multiple phenotypically distinct stages defined by the expression of CD44, growth factor receptor KIT (also known as CD117) and CD25 (REFS[2,6,7]) (FIG. 1; details given in BOX 1). All of the stages before the expression of TCR proteins are referred to as pro-T cells. To undergo conversion into definitive T lineage cells, the DN pro-T cells must activate growth factor receptor and signalling genes associated with T cell function — encoding kinases such as LCK, ITK and ZAP70 and adaptor molecules such as LAT and GADS (also known as GRAP2) — as well as the genes involved in TCR gene rearrangement, cell-surface TCR complex assembly and TCR-dependent selection. Upregulation of IL-7 receptor (IL-7R) and signalling genes, including those encoding the CD3δ, CD3γ, CD3ε and CD247 (also known as CD3ζ) chains of the TCR complex, starts in the DN2a or DN2b stage, whereas the genes associated with TCR rearrangement and expression are upregulated soon afterwards, in the DN2b to DN3a stages (FIG. 1). DN stages can be separated into three phases on the basis of known precursor–product relationships, the requirements of the cells

for extracellular signals and their commitment status[2]. Phase 1, including early T cell progenitor (ETP; also known as DN1) and DN2a stages, encompasses the proliferation of uncommitted T cell precursors and is Notch dependent. Phase 2 includes T cell lineage-committed DN2b and DN3a stages, in which cells are more Notch dependent but proliferate less and undergo TCR gene rearrangement. Phase 3 includes the DN3b to DP stages, during which the T cell lineage-committed cells proliferate in response to pre-TCR or γδ TCR signals and ultimately lose Notch dependency.

Transcription factors and chromatin

Developmental changes are generally driven by sequence-specific transcription factors, which alter gene expression patterns through their binding to specific genomic sites. They can also alter chromatin states that otherwise resist cell-type identity changes. This section introduces the ensembles of transcription factors that are important to control early T cell development and the chromatin state contexts for their work.

Two groups of transcription factors. The study of haematopoiesis has provided classic paradigms of lineage-determining transcription factors. Some can promote a cell identity ectopically, such as CCAAT/enhancer-binding protein-α (C/EBPα) and PU.1 for myeloid fates[8–10]. Some are distinctive positive regulators of cell type-specific genes that are uniquely expressed in the given lineage, such as PAX5 and EBF1 (also known as COE1) for B cells[11,12]. However, for early T cell development, no cell type-specific transcription factor set is known to have 'master-like' activity in these senses. This raises the question of how T cell genes are activated specifically. Nearly all factors that promote the T cell programme have relatives that bind to the same target motifs, working in other lineage programmes. Nevertheless, through their combined actions, overlapping with each other and with Notch signalling, a well-defined ensemble of factors establishes T cell identity. Such factors include the basic helix–loop–helix (bHLH) factors E2A and HEB, GATA3, TCF1 (encoded by *Tcf7*), BCL-11B, RUNX family factors and Ikaros (*Ikzf*) family factors, and PU.1 (encoded by *Spi1*) in a transient early role, all of which are indispensable for appropriate T cell development[2,6,7,13,14] (FIG. 2). Many of these factors are also expressed throughout later T cell stages and have roles in later developmental choices.

This group of T cell transcription factors is not activated on an empty regulatory background. Instead, the early pro-T cells continue to express a distinctive group of phase 1 transcription factors that are normally associated with haematopoietic stem and progenitor cells (HSPCs) in the bone marrow (FIG. 2): PU.1, LMO2 (also known as RBTN2), MEF2C, BCL-11A, HHEX, LYL1, GFI1B, ERG, MYCN and initially also HOXA9 and MEIS1 (REFS[2,15,16]). Forced experimental expression of LMO2 can upregulate the expression of other phase 1 factors[17]. Furthermore, the preferred binding motif for PU.1 is the most highly enriched motif among all of the chromatin sites that are selectively open before T cell lineage commitment[18–20]. These results suggest that PU.1 and LMO2 actively help to maintain the precursor state of early pro-T cells. The genes encoding all of these factors are turned off around the time of T cell lineage commitment, albeit with individual kinetics (for example, *LMO2* earlier, *ERG* later)[2,15]. However, as confirmed by single-cell transcriptome analyses in both mouse and human systems, the phase 1 transcription factors provide the context within which the initial regulatory events specific for T cell development begin[16,21,22].

Recent advances in genome-wide and multi-omics approaches and single-cell analyses have led to the discovery of key features in the molecular mechanisms by which both the T cell-associated and the progenitor cell-associated transcription factors control early T cell development. Unexpectedly, stage-dependent biochemical roles of these transcription factors enable the developing cells to satisfy both the need for extensive precursor proliferation and the need for coordinated differentiation once they reach the right stage.

A substantial shift in epigenome organization. The cell biological changes that are observed during T cell development, including changes in developmental potential and cell-surface markers, result from sequential transformations of chromatin organization and genome activity in the cells as they pass through the pro-T cell stages. From the time that ETPs enter the thymus to the time that DP cells finish TCR gene rearrangement, the organization of the genome into topologically interacting DNA loops, the particular regions of chromatin that are accessible and the patterns of DNA methylation and histone modification all change markedly[18,23–25]. The epigenetic state in any given stage makes it easy for cells to maintain current patterns of gene expression as a default, whereas the changes in epigenetic state between stages are caused by the actions of transcription factors, probably requiring specific combinations of transcription factors

Box 1 | **Early stages of T cell development in mice**

The earliest T cell precursors in the thymus, a tiny subset within the minority of thymocytes that lack expression of CD4 and CD8 (double-negative (DN) cells), are known as early T cell progenitors (ETPs) or KIT^{hi} DN1 cells ($KIT^{++}CD44^{+}CD25^{-}$ in mice). At the following DN2a ($KIT^{++}CD44^{+}CD25^{+}$) stage, expression of the T cell lineage marker CD25 is induced, but DN2a cells can still gain access to alternative, non-T cell, fates if they are removed from the presence of Notch signals. At the transition from the DN2a stage to the DN2b ($KIT^{+/int}CD44^{+}CD25^{+}$) stage, which is marked by a decrease in KIT expression, pro-T cells become intrinsically committed to the T cell lineage. They then start preparing for T cell receptor (TCR) β-chain gene (*Tcrb*) rearrangement, which mostly occurs at the DN3a ($KIT^{-}CD44^{-}CD25^{+}$) stage. In parallel, some cells rearrange the *Tcrg* and *Tcrd* genes instead, which results in differentiation to the γδ T cell lineage. Successful rearrangement of the *Tcrb* locus encodes a TCRβ chain that can complex with an invariant surrogate α-chain and signalling partners to form a pre-TCR. Pre-TCR assembly triggers progression to the DN3b stage (known as β-selection), strong proliferation and progression through the DN4 ($KIT^{-}CD44^{-}CD25^{-}$) and immature single-positive ($CD4^{-}CD8^{+}$ in mice, lacking αβ TCR) stages to the subsequent double-positive (DP; $CD4^{+}CD8^{+}$) stage. Only then does the *Tcra* locus rearrange to generate the TCRα chain of the mature αβ TCR. $CD4^{+}CD8^{+}$ DP cells then undergo positive and negative selection on the basis of their αβ TCR specificity, giving rise to all later types of αβ T cell. In human pro-T cells, the cell-surface markers that define these stages of T cell development differ from those of mouse pro-T cells, but commitment to the T cell lineage is similarly marked by a specific phenotypic change, from $CD34^{+}CD7^{+}CD1a^{-}$ cells to $CD34^{+}CD7^{+}CD1a^{+}$ cells[165–167].

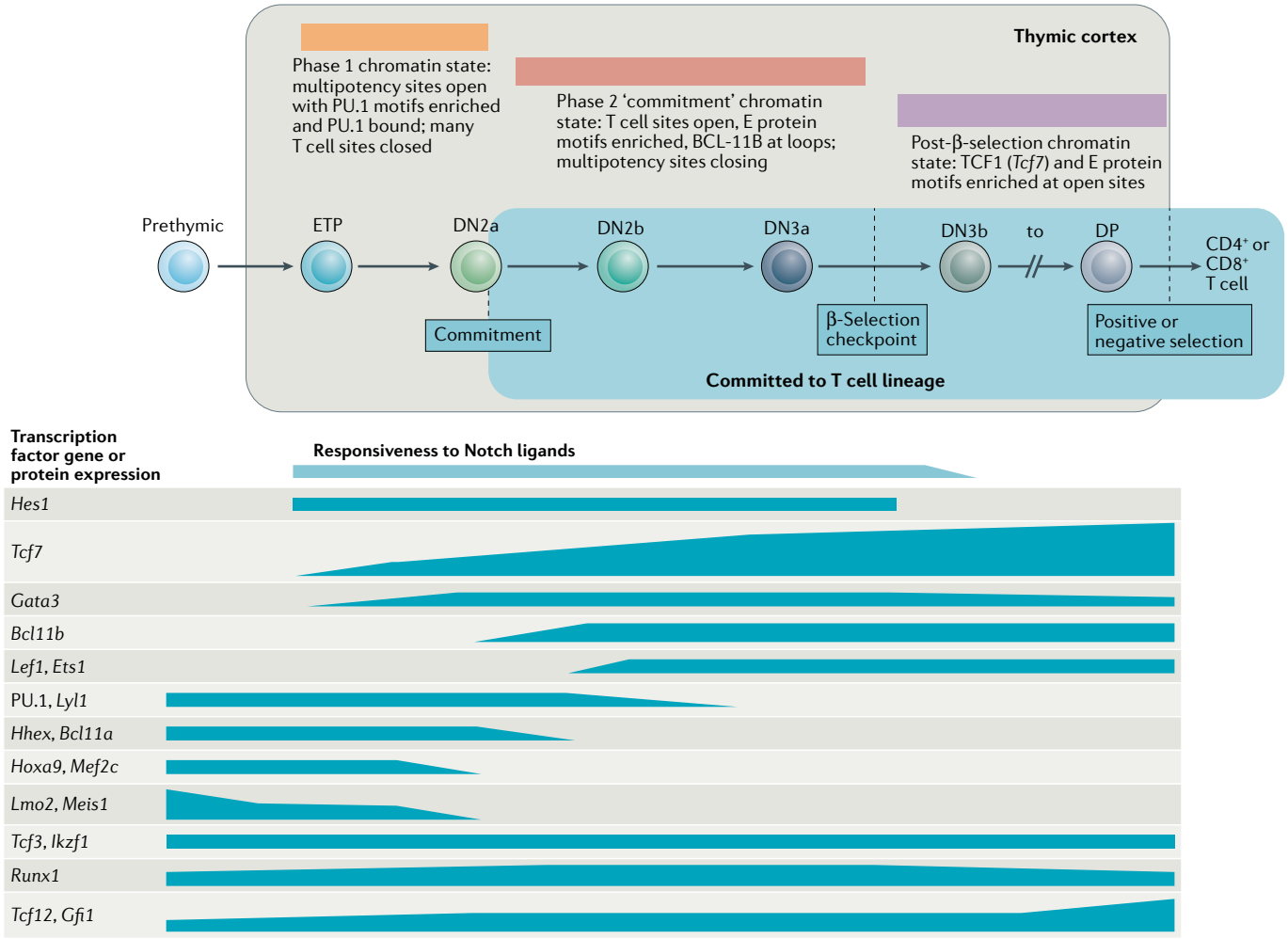

Fig. 2 | **Major changes in epigenetic state and transcription factor expression in mouse pro-T cells.** Changes in genomic accessibility and patterns of transcription factor expression during T cell development based on data in REFS[18,25] and REFS[23,34], respectively, are summarized. Indicated levels of transcription factor expression approximate a logarithmic scale. Transcription factor expression indicates RNA expression levels, except for PU.1, for which protein expression is given. Whereas *Spi1* RNA (encoding PU.1) is expressed in a similar manner to *Hhex* and *Bcl11a*, PU.1 persists for longer owing to its high stability[16]. Note the overlap in expression of the genes encoding progenitor transcription factors (such as *Lyl1* and PU.1) and T cell lineage specification transcription factors (such as *Tcf7* and *Bcl11b*) in late-stage early T cell progenitors (ETPs), double negative 2a (DN2a) cells and DN2b cells. This overlap extends through multiple cell cycles and has been validated at the single-cell level[16]. DP, double positive.

functioning together to trigger molecular switches. The new chromatin states are important for biasing T cell development towards irreversible commitment because they in turn constrain the future actions of transcription factors, as discussed next.

Detailed analysis of genome-wide chromatin accessibility and 3D organization from HSPCs to the DP thymocyte stage has shown that the greatest number of changes in chromatin site accessibility occur between the DN2 pro-T cell stage and the DN3 pro-T cell stage, which is the transition associated with T cell commitment[18] (FIG. 2). This transition is also associated with widespread changes in histone modification and genome-wide transcriptional patterns[23]. In terms of chromatin accessibility (as measured by DNase I hypersensitive site sequencing (DNase-seq) or assay for transposase-accessible chromatin using sequencing (ATAC-seq)) and in terms of the transcription factor motifs that are enriched at genome-accessible sites, cells from the DN2b stage onwards cluster together with more mature thymocytes, whereas ETPs (KIT^{++} DN1 cells) and DN2a cells consistently cluster with prethymic HSPCs[18,25]. Thus, although early pro-T cells in phase 1 stages proliferate in the thymus and receive intrathymic Notch signals, they do not convert into a definitive epigenetic state of the T cell lineage until the DN2b stage. This is because several factors in addition to Notch determine the timing of epigenetic change.

In the next section, the regulatory drivers that promote T cell identity choice are profiled in detail. The first group are those that are activated early in thymus-settling progenitors, namely Notch signals and the two T cell specification transcription factors, TCF1 and GATA3. Then, stably expressed factors that must collaborate

with TCF1 and GATA3 to execute the T cell programme are described: E proteins, Ikaros proteins and RUNX proteins. Finally, we introduce BCL-11B, the regulatory factor that is most closely linked to lineage commitment in mouse pro-T cells. The activities of these factors that support the T cell lineage programme are then described in an ensemble, as they are initially opposed by the progenitor-associated factors and then as they overcome that resistance to promote T cell commitment.

Notch signals in T cell development

The most important intercellular signalling pathway required for T cell development is the Notch pathway, and one of the most important roles of the thymus is to provide Notch ligands to progenitor T cells. Conditional deletion of *Notch1* in HSPCs, or of the gene encoding the Notch ligand Delta-like 4 (*Dll4*) in thymic epithelium, leads to a complete block of T cell development accompanied by the appearance of B cells in the thymus[26,27]. In addition, providing Notch signals to fetal liver or bone marrow-derived non-T progenitor cells induces T cell development in vitro, or ectopically in vivo[28,29]. Notch family molecules function both as transmembrane cell-surface receptors and as transcription factors[30]. Engagement of cell-surface Notch by Notch ligands on neighbouring cells triggers the proteolytic release of the intracellular domain of Notch (ICN), which undergoes nuclear translocation to become a co-activator for the DNA-binding protein RBPJ. Notch signalling promotes the developmental progression and survival of pro-T cells through phases 1 and 2, and also supports viability and proliferation competence into β-selection[31,32].

In early T cell development, Notch signalling helps to activate the expression of several genes encoding functionally important transcription factors, including *Gata3*, *Tcf7* and, later, *Bcl11b*. Well-studied mechanisms for NOTCH1 activity emphasize its function as a positive transcriptional regulator, through the recruitment of Mastermind-like factors to the RBPJ–ICN complex[33]. However, Notch pathway signalling can also trigger transcriptional repression, at least in part because Notch directly activates genes encoding transcriptional repressors from the HES family of class V helix–loop–helix (HLH) transcription factors. *Hes1* is strongly activated throughout the ETP to DN3 stages of early T cell development[23,34], and it has a role in T cell commitment as well as in pro-T cell survival[35]. Repression of *Pten* by HES1 enhances the latent proliferation competence of DN3 cells, which is unleashed when they leave the pro-T cell stages via β-selection[36].

The dominant roles of Notch signalling in early T cell development have been extensively studied[14,33]; however, there are still several important questions to be answered. First, how does Notch signalling specifically induce the T cell programme in the thymus, as distinct from its instructive roles in many other cell fate decisions in a multiplicity of embryological contexts[30,37]? Second, recent evidence shows that future T cell precursors need to be primed by Notch signalling even before they reach the thymus[38,39]; what then makes intrathymic presentation of Notch ligands necessary for T cell specification? Also, as NOTCH1 is expressed in prethymic and intrathymic T cell precursors alike, why does Notch signalling have stage-specific activities in early T cell development that differ between phase 1 and phase 2 (REFS[40–42])? Notch signalling controls the expression of *Il2ra* (encoding CD25), *Dtx1* (encoding deltex 1), *Notch3*, *Hes1* and *Ptcra* (encoding pre-TCRα) in DN cells, but many of these target genes are not activated immediately in ETPs, but only in DN2b and DN3 cells. Conversely, *Nrarp*, a negative-feedback regulator of Notch signalling, is one of the earliest Notch-activated genes to be expressed in ETPs, but its expression is shut off at the T cell lineage commitment checkpoint (between the DN2 stage and the DN3 stage)[2,43]. Therefore, specific mechanisms must exist to select distinct T cell lineage-specific and stage-specific Notch target genes even within the same cell lineage[42].

Some of the answers clearly involve the fact that Notch signals are received in the context of different combinations of other transcription factors, which are upregulated and downregulated in overlapping waves as T cell development proceeds[2,16,23,25,34] (FIG. 2). At each stage of T cell development, some of the key factors that collaborate with Notch signalling to control expression of the target genes of that stage are encoded by genes that were themselves induced in response to Notch signals in previous stages.

T cell specification factors

In ETPs, Notch signalling activates *Tcf7* (encoding TCF1) and *Gata3*, which are crucial regulatory genes for T cell specification[2]. *Tcf7* and *Gata3* seem to be upregulated within early ETPs, almost as early as the canonical Notch target gene *Hes1* (REF.[16]). These factors are indispensable for initiating the T cell programme. Although their expression is not maximally activated until later in T cell development, knocking out either *Tcf7* or *Gata3* markedly decreases the survival as well as differentiation of ETPs[44,45]. Expression of GATA3 reaches a plateau at the end of the DN2a stage, whereas for TCF1, which is already strongly expressed in DN2a cells, expression increases further at later stages. Both GATA3 and TCF1 positively regulate T cell identity genes and contribute to the later activation of a third T cell lineage-associated regulatory gene, *Bcl11b*, concomitant with lineage commitment at the transition from phase 1 to phase 2.

TCF1. The high mobility group box transcription factor TCF1 is a direct target gene of Notch signalling in ETPs. When thymic progenitors recognize Notch ligands from the thymic microenvironment, ICN is translocated to the nucleus and directly binds with RBPJ to the upstream region (–31.5 kb) of the *Tcf7* locus, as part of the mechanism that activates *Tcf7* expression[45,46]. The involvement of *Tcf7* in Notch signalling is stage specific, probably owing to additional regulatory inputs and TCF1 feedback itself, which may replace Notch in phase 2 and later stages[46,47].

Both loss-of-function and gain-of-function experiments identify TCF1 as a key T cell developmental regulator[20,45,46]. As TCF1 is also an important mediator of WNT signalling in other systems, many years of debate have revolved around whether signalling through WNT

β-Selection
The first step of T cell development that depends on a form of the T cell receptor (TCR), in this case a special immature form of the TCR consisting of only a TCRβ chain plus an invariant pre-TCRα surrogate chain. This complex is generated when double-negative 3a (DN3a) thymocytes successfully rearrange the genes encoding the TCRβ chain, and its assembly is required to enable the cells to proliferate and differentiate further to become double-positive thymocytes.

WNT signalling
The WNT pathway is a multistep developmental signalling pathway, often involved in self-renewal of tissue stem cells and in embryonic pattern formation in many organisms. In mammals, a soluble ligand from the large WNT family binds to a cell-surface receptor (of the FZD family), which enables β-catenin to avoid degradation in the cytoplasm and translocate to the nucleus, where it becomes a co-activator for transcription factors of the TCF/LEF family.

pathway components is required to instruct initiation of T cell differentiation[48]; the balance of evidence now suggests that it does not[46,49,50]. *Tcf7*-mutant (TCF1-deficient) progenitor cells can migrate into the thymus[45,46,49], but they are profoundly affected in terms of T cell development in adult mice, with survival and differentiation defects from the earliest stage of intrathymic T cell development[45,46]. Molecular characterization of TCF1 so far has focused mainly on the effects of gene disruption after commitment to the T cell lineage, abnormalities of rare, late survivors of germline mutation, or biochemical analyses of the roles of TCF1 in DP cells (which constitute 85% of thymocytes). Thus, the functions of TCF1 in earlier T cell development remain to be fully understood. Other complexities involve the roles of TCF1 isotypes with distinct functions and the partial redundancy of TCF1 with its paralogue LEF1, which is expressed after T cell lineage commitment[51,52]. However, forced expression of full-length *Tcf7* in prethymic progenitors is sufficient to trigger aspects of T cell development in the absence of Notch signalling, which confirms that TCF1 can activate a subset of T cell signature genes, including *Gata3*, *Bcl11b*, *Il2ra*, *Lck*, *Cd3g* and *Rag2* (REF.[46]).

As the epigenetic state changes during early T cell development, recent evidence indicates that TCF1 functions as a pioneer-like factor[53] to establish aspects of the T cell lineage-specific chromatin landscape, starting at the earliest stages of T cell development[20]. In a comparison of chromatin accessibility across T cell development from HSPCs to CD4$^+$ SP T cells and CD8$^+$ SP T cells, TCF1-binding motifs were found to be enriched at the genomic regions that are more 'open' in ETPs than in pre-thymic progenitors, and especially at sites that remain open in later T development. TCF1-binding motifs were also enriched at many groups of sites that are more accessible at later T cell stages. The few DP cells that emerge in *Tcf7*-knockout mice have defective T cell lineage chromatin landscapes and gene expression profiles[20], and DP cells in which *Tcf7* was recently deleted (with CD4–Cre) have reduced genomic accessibility at TCF1-binding sites[54]. In addition, forced ectopic expression of *Tcf7* in fibroblasts[20] also showed a potential pioneer-like activity for TCF1. TCF1 seemed to be able to bind to closed chromatin, erase locally pre-existing repressive histone modifications and generate de novo accessible chromatin at its target sites to induce parts of the T cell lineage-specific epigenetic and transcriptomic profiles in these fibroblasts[20]. Hence, *Tcf7*, which is an early direct Notch target gene in the T cell lineage programme, encodes a factor that is functionally important to initiate the T cell-specific chromatin landscape and to establish T cell lineage identity.

GATA3. GATA3 belongs to the GATA family of zinc-finger transcription factors and has two highly conserved C4-type zinc-fingers, each of which is followed by a conserved basic region. GATA3 expression is detected in a subset of haematopoietic stem cells and has important roles throughout T cell development in the thymus and periphery[55,56]. However, GATA3 is also important in multiple other tissues, ranging from the developing jaw to mammary glands, kidneys and sympathetic neurons[57–59]. Loss-of-function mutations of *Gata3* are associated with an autosomal dominant disease known as HDR syndrome that is characterized by hypoparathyroidism, deafness and renal disease[60]. Therefore, GATA3 has crucial dose-dependent and cell context-dependent roles in different developmental contexts.

Under physiological conditions, the very low levels of *Gata3* expression in prethymic progenitor cells are upregulated by Notch signalling on migration of the cells into the thymus. Notch signalling and GATA3 then seem to collaborate, most likely in a feedforward relay, to exclude the B cell fate in the ETP and DN2 stages[61–63]. The molecular mechanism through which Notch signalling induces upregulation of *Gata3* transcription is not fully understood, owing to the complexity of *Gata3* regulatory elements[64,65]; in part, *Gata3* transcription could be indirectly mediated by TCF1 (REF.[46]). The levels of GATA3 are tightly regulated during early T cell development, and both overexpression and knockout of *Gata3* can be toxic in ETPs[66,67]. However, *Gata3* expression gradually increases during phase 1 and then regulates several checkpoints of early T cell development, including T cell lineage commitment and β-selection, at the transition from phase 1 to phase 2 and from phase 2 to phase 3, respectively[2].

GATA3 has distinct roles in a succession of T cell developmental choices long after its role in specifying the precursors of all T cells[56,68], and it also supports other lymphoid fates, particularly that of group 2 innate lymphoid cells (ILC2s)[69]. In αβ TCR$^+$ thymocytes, GATA3 promotes the CD4$^+$ fate over the CD8$^+$ fate and subsequently promotes a T helper 2 (T_H2) cell fate over T_H1 cell or T_H17 cell fates. GATA3 binds to markedly different genomic sites in phase 1 pro-T cells to in later T cell developmental stages, which may help to explain its various roles[23,70]. This context dependence of GATA3 contrasts with the apparent ability of TCF1 to open many of the same 'T cell lineage' genomic sites in fibroblasts as those it binds to in thymocytes. Interactions with other transcription factors could affect the genomic target preferences of GATA3 at different developmental stages (see later). Another contributing mechanism could be that GATA3 is functionally affected by post-translational modification, controlled by environmental signalling. In mature T_H2 cells and ILC2s, GATA3 undergoes distinct post-translational modifications, including acetylation of lysine, phosphorylation of serine and threonine, and methylation of arginine, that not only enhance its expression but also control its nuclear translocation and the organization of GATA3 complexes[71–76]. Thus, extracellular signals from the different types of vascular endothelial cells and thymic epithelial cells in the thymus[77] might be important to control post-translational modifications of GATA3 and thus affect its functions in a stage-specific manner.

E proteins in T cell versus ILC fate

The Notch-induced factors HES1, TCF1 and GATA3 operate in a rich context of other transcription factors that are already present in prethymic precursors. Among the most important of these are E proteins. E proteins, in particular E2A, are already highly expressed in lymphoid-competent precursors before the cells enter

the thymus[78–80]. There is strong molecular evidence that they contribute to both phase 1 and later events in T cell development[81]. However, their most prominent role is to activate definitive T cell lineage genes in phase 2.

E proteins belong to the class I bHLH family of transcription factors[82] and have essential roles in the generation of haematopoietic stem cells, lymphoid-primed multipotent progenitors, B cells and T cells[82,83]. Several E proteins arise through alternative transcriptional start site usage or differential splicing from E protein loci, including *Tcf3* (E2A, encoding E12 and E47), *Tcf4* (commonly known as E2-2, encoding E2-2can and E2-2alt) and *Tcf12* (HEB, encoding HEBcan and HEBalt). E proteins associate into homodimers or heterodimers with other bHLH or HLH factors and bind to a consensus E-box motif. E protein function is antagonized by class IV HLH family inhibitor of DNA binding (ID) factors, which lack the basic domain required for DNA binding and form stable E protein–ID heterodimers that cannot bind to DNA. Among E protein dimers, the E2A homodimer and E2A–HEB heterodimer have crucial roles in early T cell development[84,85]. E2A is essential not only for proper NOTCH1 expression in lymphoid-primed multipotent progenitors and pro-T cells but also for promoting T cell lineage commitment cooperatively with Notch signalling[86–88]. The arrest of T cell development at the lineage commitment checkpoint in *Tcf3*-deficient thymocytes can be partially rescued by the introduction of ICN[86].

E proteins and their ID family antagonists have a central role in defining the fate choice between T cells and ILCs. ID2 is a crucial regulator of ILC and natural killer (NK) cell development, and it is highly expressed in common helper ILC precursors[89–92], as described later[93]. *Id2* expression, either intrinsic or exogenously introduced, enables fetal thymic pro-T cells to generate NK cells[94]. The role of ID2 in ILCs and NK cells seems to be fully explained by its ability to suppress the DNA-binding activity of E2A and HEB[81,95], as either inappropriate expression of *Id2* or deletion of both *Tcf3* and *Tcf12* similarly induces abnormally increased development of at least two classes of ILCs — ILC2s and lymphoid tissue inducer-like cells — in the thymus[81].

In developing T cell precursors, the roles of E proteins remain highly dynamic through T cell development and in mature T cell function. E2A itself is expressed constantly throughout, but HEB, ID3 and other heterodimerization partners have fluctuating levels of expression, with ID3 being transiently upregulated during each response to TCR signalling[83,96,97]. E2A and HEB activities are crucial for TCR gene rearrangement[98]. In fetal thymocytes, which develop on an accelerated schedule compared with adult thymocytes, E protein complexes assemble even at the ETP stage to establish T cell lineage chromatin landscapes around T cell signature loci, including *Notch1*, *Rag1*, *Rag2*, *Ptcra*, *Cd3g*, *Cd3d*, *Cd3e* and *Tcrb*[81], probably through direct binding on the basis of data from postnatal DN3 cells[96]. E2A is also required to control expression levels of the dose-dependent factor GATA3 around the T cell lineage commitment checkpoint, specifically to limit expression levels of *Gata3* at the DN2 stage to enable optimal T cell lineage specification[66]. High E protein levels in DN3 and DP thymocytes arrest development at those stages until cells achieve successful pre-TCR or TCR signalling[99]. However, after successful TCRβ rearrangement, pre-TCR signalling-mediated induction of *Id3* expression attenuates E protein activity, resulting in a decrease of *Notch1* expression[88]. E protein activity increases again after β-selection and has crucial roles at the DP stage[99], in close partnership with TCF1 (REF.[54]), but is again transiently neutralized during positive selection and during the maturation of $CD8^+$ SP lineage cells[100,101].

In summary, in pro-T cells, E proteins control stage-specific expression of *Notch1* and of indispensable T cell lineage genes, modulate Notch-induced expression of GATA3 and ultimately repress intrathymic development of ILCs to establish T cell identity.

Ikaros and RUNX factors in pro-T cells

Like Notch and E proteins, Ikaros family and RUNX family transcription factors are already strongly expressed by precursors before they enter the thymus and from the ETP stage throughout T cell lineage commitment. Together with the zinc-finger repressor GFI1 (REF.[102]) and MYB[103,104], Ikaros family and RUNX family factors are extremely important for the early T cell programme[105–107] despite their relatively small changes in expression level from phase 1 to phase 2. They are bifunctional, both repressing and activating target genes. Ikaros family factors function as tumour suppressors and to enforce developmental checkpoints[108,109], specifically by making the activation and repression of different waves of regulatory genes more switch-like, rather than gradual, from one developmental stage to the next[110]. Among RUNX family factors, RUNX1 is expressed very highly in DN pro-T cells and has been found to make a key contribution to the activities of both phase 1-specific and phase 2-specific factors, as described later[93,111]. Individual members of the Ikaros and RUNX families have different patterns of expression[105,112]. However, on the whole, both the Ikaros family and the RUNX family factors have their effects reinforced by the highly overlapping expression of different family members in the same early T cells. Knockout of a single family member gene results in a limited phenotype, whereas the importance of these factor families emerges clearly from studies perturbing the whole family at once, such as multiple family member gene knockouts, loss of function of the RUNX family shared cofactor core-binding factor subunit-β (CBFβ), or overexpression of dominant-negative mutants[107,113,114].

BCL-11B in T cell lineage commitment

The highly T cell lineage-specific factor BCL-11B is not activated immediately in ETPs at the same time as TCF1 and GATA3 but seems to be expressed for the first time in late DN2a stage cells, during the phase 1 to phase 2 transition (FIG. 2). It then becomes essential for specific aspects of T cell lineage commitment, for all αβ T cell development and for successful passage through β-selection, all of which are nearly eliminated in *Bcl11b*-knockout mice (reviewed in REFS[115,116]), although certain fetal γδ T cell subsets in mice are less dependent on BCL-11B[117–119]. BCL-11B has roles in the brain and several other tissues, as well as T cells, but

its role in haematopoiesis is restricted to T cell lineage cells and ILC2s[116,120]. Once it is induced at the late DN2a stage, BCL-11B expression is then sustained at some level in essentially all αβ T cell effector lineages, including NKT cells, regulatory T cells, cytotoxic T cells and peripheral effector T_H cells[121], where it often sets thresholds for effector responses[122]. In a human patient, a heterozygous missense mutation altering the second zinc-finger of BCL-11B (N441K; involved in DNA binding) caused severe T cell immunodeficiency as well as neurological defects. The resulting mutant BCL-11B had dominant-negative activity and somehow blocked effective DNA binding by BCL-11B, despite the expression of wild-type BCL-11B from the other allele[123]. Targeted mutations have shown that the DNA-binding and amino-terminal repression domains of BCL-11B are both crucial for T cell commitment, and the carboxy-terminal zinc-finger also becomes important at later stages of T cell development during the $CD4^+$ versus $CD8^+$ lineage choice[124,125].

The functional importance of BCL-11B at the T cell lineage commitment checkpoint has been demonstrated by analysis of *Bcl11b*-deficient mice and cells. Conditional deletion of *Bcl11b* in haematopoietic cells induces developmental arrest of T cells in the thymus at a distorted DN2–DN3 stage with some DN2a-like features[93,118,126,127], and it also induces or allows the abnormal activation of gene expression associated with ILC1s, NK cells and/or myeloid cells[118,128]. *Bcl11b*-deficient DN2a cells generated cells with NK cell and/or dendritic cell markers even in the presence of Notch signalling[93,118,127,128] and generated myeloid cells in the absence of Notch signals[118,127]. Thus, BCL-11B is indispensable for appropriate lymphopoiesis in humans and mice.

BCL-11B is bifunctional (activating and repressive) like Ikaros and RUNX factors, and it also seems to have an active role in chromatin organization, as deletion of *Bcl11b* starting in DP thymocytes disrupted the promoter–enhancer interactions that control dichotomous *Cd4* or *Cd8a* expression in positive selection[125]. A genome-wide survey of developmental changes in chromatin looping topology from HSPCs to DP cells showed that the newly interacting DNA sites after T cell lineage commitment were substantially enriched in BCL-11B binding[18]. When *Bcl11b* expression was disrupted (albeit later, in naive mature $CD4^+$ T cells), chromatin interactions and loop formations were globally reduced near BCL-11B-deprived loop anchor sites[18].

The mechanisms through which BCL-11B regulates specific target genes in early pro-T cells include both direct, binding site-specific recruitment of chromatin-modifying proteins and indirect effects via gene network circuitry. Many binding sites occupied by BCL-11B may themselves be determined by its collaboration with other factors, as they are most enriched in motifs that bind to RUNX and ETS family factors, and many of the genomic sites bound by BCL-11B are also bound by RUNX1 (REF.[93]). During T cell lineage commitment, BCL-11B is important for both the activation and the repression of target genes[93,126], as it forms complexes with both the SWI/SNF complex and the nucleosome remodelling deacetylase complex (NuRD complex)[129–131]. In pro-T cells and a pro-T cell-like cell line, BCL-11B was shown by proteomic analysis to interact with the NuRD complex, Polycomb repressor complex 1 (PRC1), REST transcriptional repressor complex and KDM1A histone demethylase complexes, as well as the transcription factor RUNX1 (REF.[93]). BCL-11B seems to act directly to recruit these factors to many genomic sites, as shown by chromatin immunoprecipitation followed by sequencing (ChIP-seq) analysis of control and *Bcl11b*-deficient cells to compare binding patterns of BCL-11B partners in the presence and absence of BCL-11B itself[93]. Overall, as is seen for many other transcription factors, BCL-11B and its associated complexes occupied genomic sites near both non-regulated genes and BCL-11B-regulated genes alike. However, the regions near functionally BCL-11B-regulated target loci were distinguished by enrichment of specific sites where the recruitment of cofactors such as RUNX1 and NuRD complex components was dependent on BCL-11B or redistributed by BCL-11B[93]. Thus, BCL-11B works at many of its functional target sites by correctly directing the nucleation of complexes with RUNX1 and chromatin-modulating complexes to establish the activation and repression of genes in a T cell lineage-specific pattern. In summary, BCL-11B has an essential role in excluding T cell progenitors at the transition from phase 1 to phase 2 from potential access to alternative lineages, and it also functions to maintain T cell lineage-specific higher-order chromatin structure after phase 2.

PU.1 and resistance to commitment

At least two general mechanisms slow the initial response of cells to the T cell specification factors TCF1 and GATA3 and to Notch signalling. One mechanism involves the starting epigenetic state of the cells as they enter the thymus, in which substantial genomic regions around key T cell-specific genes are inaccessible owing to repressive marks on histones, DNA methylation, compaction of chromatin and/or intranuclear localization[18,23,24]. One of the genes that is under a strong *cis*-acting constraint by an initially closed epigenetic state is *Bcl11b*[132,133], as discussed later. The other restrictive mechanism is the activity of a robust, alternative gene regulatory network operating in the T cell precursors initially, involving phase 1 transcription factors such as PU.1. Recent single-cell transcriptome analysis using a sensitive method confirmed that the individual ETPs that enter the T cell programme do so initially while expressing an extensive set of phase 1 transcription factors[16]. Regulatory linkages between these transcription factors show that they actively sustain a distinct subcircuit within the larger gene regulatory network of T cell specification[134]. The phase 1 network opposes progression to T cell commitment and operates throughout the phase 1 pro-T cell stages, sustaining multipotency while allowing multiple cell cycles before full TCR gene rearrangement. The pro-T cells need to deactivate this initial network to progress through T cell lineage commitment.

The best-studied member of the phase 1 transcription factors in pro-T cells is PU.1. PU.1 (encoded by *Spi1*) is an ETS family transcription factor with a broad

SWI/SNF complex
A nucleosome remodelling protein complex in eukaryotic cells that generally opens chromatin to allow greater transcription factor access. This is thought to be an important step involved in transcriptional activation of many genes.

Nucleosome remodelling deacetylase complex
(NuRD complex). A nucleosome remodelling protein complex that is recruited by many transcription factors and is often involved in target gene repression. Although the histone deacetylase activity of the complex is often used for repression, the complex as a whole can be involved in various transcriptional regulatory activities.

range of roles in haematopoiesis. It regulates lineage specification of macrophages, granulocytes, dendritic cells and B cells. However, PU.1 is also modestly but consistently expressed in phase 1 stages of T cell development, where it supports the proliferation of phase 1 cells to maintain the size of the T cell progenitor pool before β-selection[3,135–137]. It is required to generate early T cell precursors, as *Spi1*-deficient HSPCs fail to contribute to the T cell lineage[138]. Despite modest levels of *Spi1* expression in phase 1 in terms of RNA copies per cell[16], the stability of PU.1 at the protein level[139] enables it to have a substantial impact on the genomic activity of early pro-T cells within the thymus.

PU.1 directly regulates many target genes in the phase 1 state, activating several that encode other phase 1 transcription factors, such as *Bcl11a*, *Lmo2* and *Mef2c*, which may themselves have roles in controlling T cell differentiation. The preferences of PU.1 binding to different genomic sites in pro-T cells are largely based on site affinities, PU.1 concentrations and the initial chromatin accessibilities of the sites[19] (FIG. 3a,b). PU.1 can function as a pioneer-like factor, recognizing target sites in closed chromatin and recruiting other factors for lineage determination[140–143]. It induces chromatin accessibility at a subset of initially closed sites when it binds, which is associated with activating its target genes[19,111,144]. However, PU.1 binding site choices across the genome are also affected by cell type-specific binding partners[145] with which PU.1 forms cell type-specific protein complexes. Such partners include C/EBPα and nuclear factor-κB in myeloid cells, and IRF4, IRF8 or E proteins in B lineage cells. In pro-T cells, ChIP-seq and proteomic analysis of PU.1-interacting molecules have shown that PU.1 forms functional protein complexes with RUNX1 and SATB1 (REF.[93]). Chromatin sites that are open specifically during the phase 1 pro-T cell stages are highly enriched in the PU.1 recognition motif[18–20], and ChIP-seq shows that PU.1 binds to tens of thousands of sites in the genomes of ETPs and DN2a cells, with the number of sites decreasing as PU.1 levels decrease in DN2b cells. When experimentally introduced, PU.1 can rapidly open chromatin at its binding sites; conversely, downregulation of PU.1 during T cell lineage commitment is associated with the closure of a large fraction of its open binding sites, contributing to the loss of expression of linked phase 1-specific target genes[19].

PU.1 not only is a strong 'placeholder' for chromatin sites that are kept open during precommitment stages but also potently influences the site choices of other transcription factors during these stages. As described in BOX 2, PU.1 can competitively recruit factors including RUNX1, SATB1 and, to some extent, GATA3, in the process depleting them from the genomic sites that they would otherwise occupy after PU.1 is downregulated[111] (FIG. 3c–f). This expands the effects of PU.1 beyond the genomic sites that it binds directly.

In phase 1 pro-T cells within the thymus, PU.1 supports proliferation while inhibiting differentiation. Using a Cre-encoding retroviral vector or CRISPR–Cas9 systems for phase 1-specific disruption of the *Spi1* gene[111,135], loss of PU.1 seemed to enable phase 1 pro-T cells to progress to phase 2 more rapidly than control cells but also reduced their proliferation and survival, greatly decreasing the number of cells that were eventually recovered. Nevertheless, PU.1 expression needs to be downregulated during T cell lineage commitment. Inefficient silencing of *Spi1*, or abnormal expression of PU.1 or its positively regulated target genes, can cause T cell leukaemia[2,131,146,147]. The ability of pro-T cells to avoid myeloid differentiation during PU.1-expressing stages depends on their sustained response to Notch signals[42,47,148,149]; thus, it may be important for *Spi1* to be silenced before the cells lose Notch responsiveness just after β-selection.

In addition to PU.1, other members of the phase 1 transcription factor network, including BCL-11A, LYL1 and LMO2, have important roles in T cell progenitor survival and/or proliferation before TCR gene rearrangement[17,150,151]. Thus, the phase 1 state may delay differentiation and sustain multipotency partly as a side effect of these population-sustaining activities, which are important for the ultimate yield of T cells.

How T cell commitment is established

During commitment to the T cell lineage, pro-T cells relinquish access to at least three types of alternative regulatory state: myeloid cells and dendritic cells, ILCs and NK cells, or continuation as a multipotent progenitor. Evidence currently suggests that loss of access to the PU.1-dependent myeloid and dendritic cell fates is attributable to silencing of *Spi1*, whereas loss of access to ILC and NK cell fates is mediated by a stage-specific function of E proteins, reinforced by a gene network effect of newly expressed BCL-11B.

There is good correlation between the naturally occurring downregulation of PU.1 expression and the loss of access to myeloid cell fates during T cell lineage commitment[152–156]. If Notch signalling is removed or attenuated from PU.1-expressing ETPs and DN2a cells, they can generate myeloid cells and dendritic cells; phase 2 cells, which have naturally downregulated PU.1 expression, cannot, but the reintroduction of PU.1 restores the myeloid potential of these cells[47,148,149]. Therefore, downregulation of PU.1 expression in phase 1 cells is one of the crucial events that excludes myeloid cell and dendritic cell potential at the T cell lineage commitment checkpoint. Notch signalling itself does not repress *Spi1*, but GATA3, RUNX1, TCF1 and certain *cis*-regulatory elements around the *Spi1* locus have been reported to participate in *Spi1* silencing[63,147,157,158]. However, the molecular genetic mechanisms explaining how PU.1 expression is repressed in a stage-specific manner are still not fully clear, and they may also involve the downregulation of other *Spi1* activators in the phase 1 transcription factor set.

A relatively small number of T cell lineage-associated regulatory genes are upregulated during the transition to T cell lineage commitment. Gene expression evidence indicates that there is a change in E protein activity at this point[96], despite only a small increase in the level of the E protein HEB and no increase in the level of E2A. Instead, the major regulatory genes that undergo a much sharper increase in expression from phase 1 to phase 2 are *Bcl11b*, *Ets1* and *Lef1*. Commitment of mouse pro-T

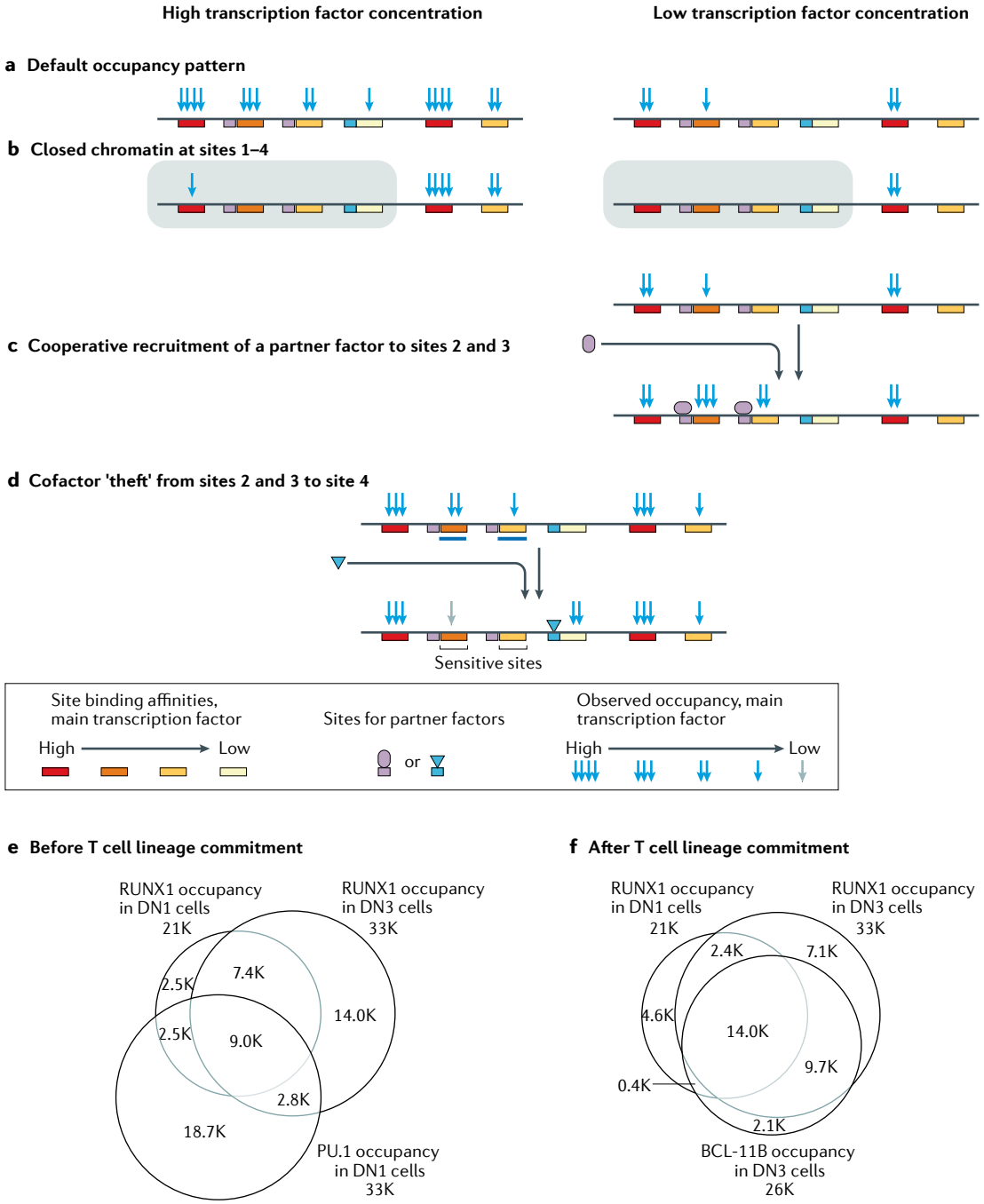

Fig. 3 | **Conditionality of transcription factor binding at genomic sites.** Illustrations of how the same transcription factor can differentially occupy genomic sites on the basis of their intrinsic affinities for binding by the factor, their chromatin accessibility status and their comparative advantage when a second transcription factor with its own binding specificities can interact with the first factor. The schematics in parts **a**–**d** are drawn from examples in REFS[19,111,144]. Default occupancy patterns for an idealized transcription factor are shown on six genomic sites that it recognizes with different intrinsic affinities, for high and low concentrations of the transcription factor (part **a**). Alterations of this default occupancy pattern are shown in a cell type for which some genomic sites are occluded by closed chromatin (part **b**). This part of the figure schematizes results seen for PU.1 in pro-T cells[19]. Sites in closed chromatin can still be bound at high transcription factor concentrations if they have high-affinity motifs[19]. Cooperative recruitment (part **c**) is illustrated by the ability of a potential interaction partner (light purple oval) to enhance the occupancy of marginal sites by the main transcription factor by coordinated binding. Cofactor 'theft' (part **d**) involves loss of binding by the main transcription factor from a subset of occupancy sites ('sensitive sites'), as can be observed when certain partner factors (blue triangle) recruit it to some alternative sites. The same transcription factor can have either role in different contexts. The effects of the mechanisms described in parts **a**–**d** on the actual patterns of occupancy by RUNX1 before (part **e**) and after (part **f**) T cell lineage commitment are shown. There is biased overlap of the precommitment pattern of RUNX1 binding in double-negative 1 (DN1) cells with sites occupied by PU.1 in DN1 cells (data from REF.[111]) (part **e**). Extensive interaction of sites occupied by RUNX1 with BCL-11B binding occurs after T cell lineage commitment in DN3 cells (data from REF.[93]) (part **f**). In parts **e** and **f**, peak numbers in each category are given in thousands (K).

cells coincides at the single-cell level with the abrupt onset of *Bcl11b* expression[159].

Bcl11b activation is a result of the combined actions of Notch signalling, TCF1, GATA3 and RUNX1 (REF.[159]), but its activation is slowed by being trapped initially in a repressive chromatin configuration[132,133] (reviewed in REF.[160]). Current evidence suggests that TCF1 and GATA3 are needed specifically during the ETP stage[159] to sensitize the *Bcl11b* locus for a slow-acting chromatin opening mechanism, which is accelerated by Notch signalling and enhanced by a far-distal enhancer[132] that initiates a chromatin compartment flip[18]. A specifically activated long non-coding RNA from the enhancer complex region, ThymoD, is also required[161]. These changes open the chromatin at the *Bcl11b* locus, multiple loops are established from enhancer regions to the *Bcl11b* promoter[18] and *Bcl11b* is finally activated. It may be downmodulated transiently after TCR signalling but remains expressed in T cells thereafter.

Once expressed, BCL-11B directly represses certain genes that have constitutive functions in ILCs and NK cells but are needed only during antigen-activated effector function in T cells[93]. Among these BCL-11B-repressed genes, *Id2* and *Zbtb16* (encoding PLZF) are functionally important direct targets. As discussed earlier, suppression of *Id2* expression is crucial for T cell development, in particular to maintain E protein activity, and to avoid inappropriate development of ILCs in the thymus[81]. Downregulation of *Id2* or *Zbtb16* expression reverses part of the abnormal phenotype of *Bcl11b*-deficient pro-T cells[93]. Thus, BCL-11B functions in T cell lineage commitment through three mechanisms: direct activation and repression of target genes involving several different protein complexes, large-scale chromatin reorganization and repression of *Id2* and *Zbtb16* expression.

The global impact of T cell lineage commitment on gene expression is more than the expression or loss of a single key transcription factor, such as BCL-11B or PU.1, respectively. Importantly, other transcription factors shift their deployment across the genome, even those such as GATA3 and RUNX1 that have been present throughout the commitment transition. RUNX1 binds to markedly different genomic sites before and after commitment (FIG. 3e,f). Multiple changes of binding by these transcription factors occur around developmentally regulated genes, including coordinated gains of occupancy around some activated genes and losses of occupancy around some downregulated genes (FIG. 4). This could reflect changes in regional chromatin accessibility owing to the global switch from a progenitor-like chromatin state to a definitive T cell lineage chromatin state. Thus, not only the direct transcriptional effects of factors such as PU.1 and BCL-11B but also indirect effects mediated by both cofactor redeployment and chromatin accessibility changes (FIG. 3) likely contribute to the regulation of specific loci (FIG. 4).

Conclusions

Notch signalling powerfully directs multipotent progenitors into the T cell pathway by activating genes encoding transcription factors that propagate a cascade of regulatory changes, both activating and repressive, to transform cell identity. The overall shifts in cellular developmental potential result from multistep expression changes of groups of transcription factors, both activation of T cell factors and repression of progenitor-specific factors. The coordinated timing of multiple changes in transcriptional regulators is responsible for T cell lineage commitment rather than the expression of a single 'master regulator' during commitment. Coordinated action of these transcription factors is also likely to catalyse the epigenetic changes that transform default genomic accessibility profiles. Furthermore, gene regulatory networks that mobilize intermediate transcription factors also contribute to the global state switches.

The sequence of regulatory changes, their relation to changes in target gene expression and their relation to stepwise changes in developmental potential, as reviewed here, are characterized for early T cell development in a level of detail that is uncommon in mammalian systems. These changes can thus be instructive for other stem cell-based systems. One notable feature of T cell development is the long overlap, over multiple cell cycles, between expression of various transcription factors associated with multipotency or immaturity and expression of various transcription factors associated with T cell development (FIG. 2). The results reviewed here emphasize that the context of other transcription factors can substantially modify the manner in which any given transcription factor is deployed genomically. Such an effect could underlie the way that E proteins, despite near-constant expression throughout pro-T cell development, become dominant regulators of signature genes that are turned on only after T cell lineage commitment. It also explains how the same factor, such as

Box 2 | **Cofactor redeployment by lineage-determining transcription factors**

Transcription factors work together at active enhancers to increase each other's likelihood of occupancy, by opening chromatin and/or through direct protein–protein interactions[54,168–173]. However, the recruitment of transcription factors to new sites may result in the loss of those transcription factors from other genomic sites (FIG. 3c,d), which means that stably expressed transcription factors can have stage-specific effects. For example, the bifunctional transcription factor RUNX1 associates with PU.1 and contributes to PU.1-mediated gene regulation in phase 1, positively regulates *Bcl11b* expression at the late double-negative 2a (DN2a) stage and, after commitment to the T cell lineage, collaborates functionally with BCL-11B in both the activation and the repression of target genes[93,111,159]. Whereas *Runx1* expression increases only moderately from phase 1 to phase 2, RUNX1-binding sites change markedly[93,111] (FIGS 3e,f,4). These binding site choices not only overlap with but are also strongly affected by PU.1 in phase 1 and BCL-11B and other factors in phase 2 (REFS[93,111]).

When PU.1 is forcibly expressed in DN3 cells (after commitment to the T cell lineage), it primarily activates its own local binding targets but also represses other genes even without obvious local DNA binding. This repression is often associated with PU.1-induced loss of RUNX1 from sites that RUNX1 was otherwise occupying[111,135,174]. In normal postcommitment pro-T cells, in which RUNX1 often binds with BCL-11B at regulated and unregulated loci alike, the specific subset of RUNX1-binding sites that disappear or redistribute if BCL-11B is removed are enriched at loci that change expression on *Bcl11b* disruption[93]. Similar scenarios have been reported for the deployment of heart specification factors during cardiogenesis[174], for the action of Ikaros in tethering nucleosome remodelling deacetylase complexes to restrain leukaemic transformation[175], and for the effect of T-bet on GATA3 in developing human T helper 1 cells[176]. The system-level redeployment of cofactors by lineage-determining transcription factors could contribute to many unknown phenomena ascribed to secondary or tertiary effects of transcription factors.

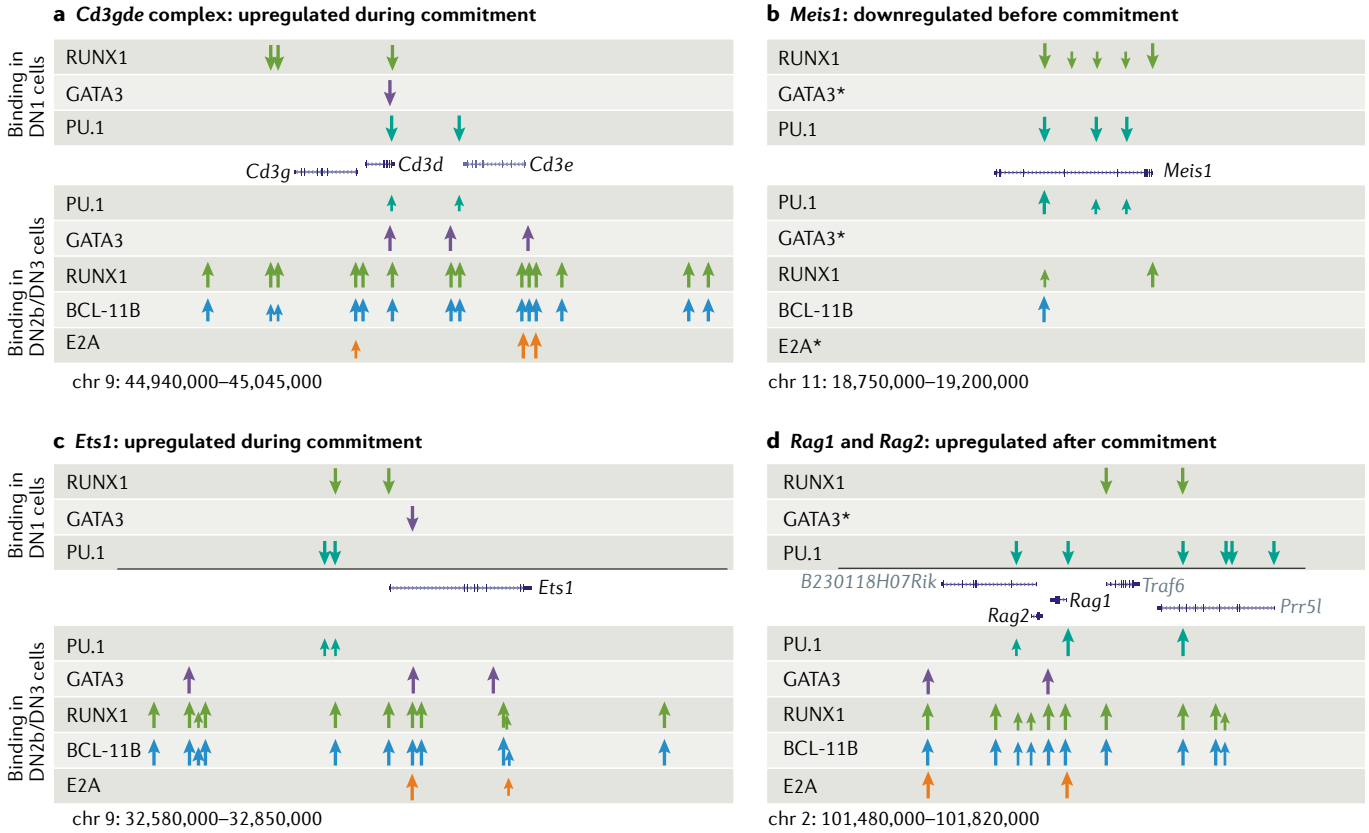

Fig. 4 | **Transcription factor binding changes at key developmentally regulated loci.** Summary schematics are shown for transcription factor occupancies observed by chromatin immunoprecipitation followed by sequencing at indicated loci before T cell lineage commitment (in early T cell progenitors (ETPs) (double-negative 1 (DN1) cells)) and after T cell lineage commitment (in DN2b and DN3 cells), comparing PU.1 in DN1 samples, BCL-11B in DN2b samples, E2A in DN3 samples (*Rag2*-knockout thymocytes) and GATA3 and RUNX1 in both DN1 and DN2b/DN3 samples (parts **a**–**d**). Original data are from REFS[23,93,96,111], aligned after remapping to the mm10 build of the mouse genome. *Cd3gde* cluster genes (part **a**), *Ets1* (part **c**) and *Rag1* and *Rag2* (part **d**) are upregulated sharply from the DN2a stage to the DN2b stage during T cell lineage commitment, whereas the progenitor cell regulatory gene *Meis1* (part **b**) is downregulated before T cell lineage commitment. Genomic regions depicted are shown at the bottom of each panel in mm10 coordinates, and transcription factor binding positions are shown to scale. Smaller arrows indicate lower detected levels of occupancy of the indicated transcription factor. Note the changes in the binding patterns of GATA3 and RUNX1 from the precommitment stage to the postcommitment stage, despite only modest changes in protein expression level. The bifunctional transcription factor RUNX1 undergoes single-site changes in occupancy at some loci but multisite increases in occupancy at others, which indicates that its binding might be regulated by broader genomic domain opening. chr, chromosome. *, no binding seen in region.

GATA3, can regulate different genes to control different developmental branch points, even within the same lineage. An intriguing possibility is that competitive transcription factor interactions could also contribute to the high dosage dependency that is seen for many of the transcription factors in this system.

The evidence available now provides a strong vantage point for tackling the developmental decisions that need future elucidation. Key unanswered questions include what the specific repressive mechanisms are that determine the shutoff of progenitor genes. Also unclear is how distinct branches of T cell-related development diverge, specifically how the cells choose to adopt molecularly distinct fates. The αβ and γδ T cell lineages may begin to separate at the DN2 stage, whereas T cell and ILC precursors, which can be distinguished by levels of ID2 expression, presumably define an even earlier branch point. Another question is how developing T cells gain access to alternative, stereotypic, effector gene programmes, even though these programmes remain latent until different developmental milestones in αβ T cells, γδ T cells, invariant NK T cells and ILCs (see REF.[160]). Finally, for humans, the success of attempts to rejuvenate T cell development in disease or ageing may depend on understanding how progenitor proliferation is regulated before TCR expression in the human thymus. Mouse evidence suggests that these stages could account for at least half of all the cell cycles that precursors undergo before becoming mature T cells[3,162,163]. Thus, transiently promoting the precommitment state and guiding cells through the commitment transition accurately, to stimulate precursor proliferation while avoiding leukaemogenesis and lineage infidelity, could become a clinical tool of the future.

Note in proof

After this manuscript was in production, a highly relevant study was published that dissects the respective roles of TCF1, GATA3 and BCL-11B in the T cell commitment process[177].

Published online 11 September 2020

1. Hosokawa, H. & Rothenberg, E. V. Cytokines, transcription factors, and the initiation of T-cell development. *Cold Spring Harb. Perspect. Biol.* **10**, a028621 (2018).
2. Yui, M. A. & Rothenberg, E. V. Developmental gene networks: a triathlon on the course to T cell identity. *Nat. Rev. Immunol.* **14**, 529–545 (2014).
3. Lu, M. et al. The earliest thymic progenitors in adults are restricted to T, NK, and dendritic cell lineage and have a potential to form more diverse TCRβ chains than fetal progenitors. *J. Immunol.* **175**, 5848–5856 (2005).
4. Singer, A., Adoro, S. & Park, J. H. Lineage fate and intense debate: myths, models and mechanisms of $CD4^-$ versus CD8-lineage choice. *Nat. Rev. Immunol.* **8**, 788–801 (2008).
5. Desiderio, S. Temporal and spatial regulatory functions of the V(D)J recombinase. *Semin. Immunol.* **22**, 362–369 (2010).
6. Rothenberg, E. V., Moore, J. E. & Yui, M. A. Launching the T-cell-lineage developmental programme. *Nat. Rev. Immunol.* **8**, 9–21 (2008).
7. Yang, Q., Jeremiah Bell, J. & Bhandoola, A. T-cell lineage determination. *Immunol. Rev.* **238**, 12–22 (2010).
8. Romanoski, C. E., Link, V. M., Heinz, S. & Glass, C. K. Exploiting genomics and natural genetic variation to decode macrophage enhancers. *Trends Immunol.* **36**, 507–518 (2015).
9. Laiosa, C. V., Stadtfeld, M. & Graf, T. Determinants of lymphoid-myeloid lineage diversification. *Annu. Rev. Immunol.* **24**, 705–738 (2006).
10. Takahashi, K. & Yamanaka, S. A decade of transcription factor-mediated reprogramming to pluripotency. *Nat. Rev. Mol. Cell Biol.* **17**, 183–193 (2016).
11. Schebesta, A. et al. Transcription factor Pax5 activates the chromatin of key genes involved in B cell signaling, adhesion, migration, and immune function. *Immunity* **27**, 49–63 (2007).
12. Lin, Y. C. et al. A global network of transcription factors, involving E2A, EBF1 and Foxo1, that orchestrates B cell fate. *Nat. Immunol.* **11**, 635–643 (2010).
13. Seo, W. & Taniuchi, I. Transcriptional regulation of early T-cell development in the thymus. *Eur. J. Immunol.* **46**, 531–538 (2016).
14. Thompson, P. K. & Zúñiga-Pflücker, J. C. On becoming a T cell, a convergence of factors kick it up a Notch along the way. *Semin. Immunol.* **23**, 350–359 (2011).
15. Rothenberg, E. V., Ungerback, J. & Champhekar, A. Forging T-lymphocyte identity: intersecting networks of transcriptional control. *Adv. Immunol.* **129**, 109–174 (2016).
16. Zhou, W. et al. Single-cell analysis reveals regulatory gene expression dynamics leading to lineage commitment in early T cell development. *Cell Syst.* **9**, 321–337 e9 (2019).
This study focuses on patterns of transcription factor expression in the earliest stages of T cell differentiation and their relationship to developmental potentials at the single-cell level using single-molecule fluorescence in situ hybridization and RNA sequencing for high-sensitivity quantitation of the expression of multiple transcription factor genes in the same single cells.
17. Cleveland, S. M. et al. Lmo2 induces hematopoietic stem cell-like features in T-cell progenitor cells prior to leukemia. *Stem Cells* **31**, 882–894 (2013).
18. Hu, G. et al. Transformation of accessible chromatin and 3D nucleome underlies lineage commitment of early T cells. *Immunity* **48**, 227–242 e8 (2018).
Revealing major changes in chromatin organization during commitment, this study shows that PU.1 motifs predominate in chromatin regions that are open early in pro-T cells. Then PU.1 occupancy gives way to BCL-11B occupancy of open sites across the genome, with evidence that BCL-11B is important as a factor associated with genomic loops in T lineage cells.
19. Ungerbäck, J. et al. Pioneering, chromatin remodeling, and epigenetic constraint in early T-cell gene regulation by SPI1 (PU.1). *Genome Res.* **28**, 1508–1519 (2018).
This study uses a motif quality metric for a global dissection of how PU.1 functions in open chromatin maintenance, pioneer-like activity and gene regulation in early pro-T cells. It shows that PU.1 respects chromatin constraints but can bind to closed sites if they are of sufficiently high affinity.
20. Johnson, J. L. et al. Lineage-determining transcription factor TCF-1 initiates the epigenetic identity of T cells. *Immunity* **48**, 243–257 e10 (2018).
This study establishes TCF1 as a controller of open chromatin states in T lineage cells. The authors show enrichment of TCF1 motifs in T cell lineage-associated chromatin sites of two types: one set becoming accessible initially in ETPs, and a larger set becoming accessible around the time of or after commitment.
21. Lavaert, M. et al. Integrated scRNA-Seq identifies human postnatal thymus seeding progenitors and regulatory dynamics of differentiating immature thymocytes. *Immunity* **52**, 1088–1104 e6 (2020).
22. Le, J. et al. Single-cell RNA-Seq mapping of human thymopoiesis reveals lineage specification trajectories and a commitment spectrum in T cell development. *Immunity* **52**, 1105–1118 e9 (2020).
Lavaert et al. (2020) and Le et al. provide complementary single-cell transcriptome analyses of pro-T cells in postnatal human thymus. Lavaert et al. focus on potential regulatory network connections, and Le et al. focus on lineage commitment. These two reports elegantly document broad similarities but some differences between early human and mouse stages.
23. Zhang, J. A., Mortazavi, A., Williams, B. A., Wold, B. J. & Rothenberg, E. V. Dynamic transformations of genome-wide epigenetic marking and transcriptional control establish T cell identity. *Cell* **149**, 467–482 (2012).
24. Ji, H. et al. Comprehensive methylome map of lineage commitment from haematopoietic progenitors. *Nature* **467**, 338–342 (2010).
25. Yoshida, H. et al. The cis-regulatory atlas of the mouse immune system. *Cell* **176**, 897–912 e20 (2019).
In this comprehensive survey, the authors show that enhancer activity correlates better with gene expression than promoter accessibility and that E protein and TCF family motifs predominate in enhancer sites that become accessible after T cell lineage commitment. A global analysis further relates specific factors to chromatin accessibility changes.
26. Radtke, F. et al. Deficient T cell fate specification in mice with an induced inactivation of Notch1. *Immunity* **10**, 547–558 (1999).
27. Hozumi, K. et al. Delta-like 4 is indispensable in thymic environment specific for T cell development. *J. Exp. Med.* **205**, 2507–2513 (2008).
28. Schmitt, T. M. & Zúñiga-Pflücker, J. C. Induction of T cell development from hematopoietic progenitor cells by Delta-like-1 in vitro. *Immunity* **17**, 749–756 (2002).
29. Hozumi, K., Abe, N., Chiba, S., Hirai, H. & Habu, S. Active form of Notch members can enforce T lymphopoiesis on lymphoid progenitors in the monolayer culture specific for B cell development. *J. Immunol.* **170**, 4973–4979 (2003).
30. Bray, S. J. Notch signalling: a simple pathway becomes complex. *Nat. Rev. Mol. Cell. Biol.* **7**, 678–689 (2006).
31. Yashiro-Ohtani, Y., Ohtani, T. & Pear, W. S. Notch regulation of early thymocyte development. *Semin. Immunol.* **22**, 261–269 (2010).
32. Ciofani, M. & Zúñiga-Pflücker, J. C. A survival guide to early T cell development. *Immunol. Res.* **34**, 117–132 (2006).
33. Radtke, F., Macdonald, H. R. & Tacchini-Cottier, F. Regulation of innate and adaptive immunity by Notch. *Nat. Rev. Immunol.* **13**, 427–437 (2013).
34. Mingueneau, M. et al. The transcriptional landscape of αβ T cell differentiation. *Nat. Immunol.* **14**, 619–632 (2013).
35. De Obaldia, M. E. et al. T cell development requires constraint of the myeloid regulator C/EBP-α by the Notch target and transcriptional repressor Hes1. *Nat. Immunol.* **14**, 1277–1284 (2013).
36. Wong, G. W., Knowles, G. C., Mak, T. W., Ferrando, A. A. & Zúñiga-Pflücker, J. C. HES1 opposes a PTEN-dependent check on survival, differentiation, and proliferation of TCRβ-selected mouse thymocytes. *Blood.* **120**, 1439–1448 (2012).
37. Muñoz-Descalzo, S., de Navascues, J. & Martinez Arias, A. Wnt-Notch signalling: an integrated mechanism regulating transitions between cell states. *Bioessays* **34**, 110–118 (2012).
38. Chen, E. L. Y., Thompson, P. K. & Zúñiga-Pflücker, J. C. RBPJ-dependent Notch signaling initiates the T cell program in a subset of thymus-seeding progenitors. *Nat. Immunol.* **20**, 1456–1468 (2019).
The authors develop a mouse strain in which the Notch-interacting transcription factor RBPJ can be deleted and then inducibly restored. The results confirm the need for sustained intrathymic signalling and show, unexpectedly, that a Notch signal via RBPJ needs to be delivered to bone marrow multipotent progenitors even before they reach the thymus.
39. Yu, V. W. et al. Specific bone cells produce DLL4 to generate thymus-seeding progenitors from bone marrow. *J. Exp. Med.* **212**, 759–774 (2015).
40. Hirano, K. et al. Delta-like 4-mediated Notch signaling is required for early T-cell development in a three-dimensional thymic structure. *Eur. J. Immunol.* **45**, 2252–2262 (2015).
41. Wolfer, A., Wilson, A., Nemir, M., MacDonald, H. R. & Radtke, F. Inactivation of Notch1 impairs VDJβ rearrangement and allows pre-TCR-independent survival of early αβ lineage thymocytes. *Immunity* **16**, 869–879 (2002).
42. Romero-Wolf, M. et al. Notch2 complements Notch1 to mediate inductive signaling that initiates early T cell development. *J. Cell Biol.* **219**, e202005093 (2020).
43. Yun, T. J. & Bevan, M. J. Notch-regulated ankyrin-repeat protein inhibits Notch1 signaling: multiple Notch1 signaling pathways involved in T cell development. *J. Immunol.* **170**, 5834–5841 (2003).
44. Hosoya, T. et al. GATA-3 is required for early T lineage progenitor development. *J. Exp. Med.* **206**, 2987–3000 (2009).
45. Germar, K. et al. T-cell factor 1 is a gatekeeper for T-cell specification in response to Notch signaling. *Proc. Natl Acad. Sci. USA* **108**, 20060–20065 (2011).
46. Weber, B. N. et al. A critical role for TCF-1 in T-lineage specification and differentiation. *Nature* **476**, 63–68 (2011).
47. Del Real, M. M. & Rothenberg, E. V. Architecture of a lymphomyeloid developmental switch controlled by PU.1, Notch and Gata3. *Development* **140**, 1207–1219 (2013).
48. Staal, F. J. T. & Sen, J. M. The canonical Wnt signaling pathway plays an important role in lymphopoiesis and hematopoiesis. *Eur. J. Immunol.* **38**, 1788–1794 (2008).
49. Xu, Z. et al. Cutting edge: β-catenin-interacting Tcf1 isoforms are essential for thymocyte survival but dispensable for thymic maturation transitions. *J. Immunol.* **198**, 3404–3409 (2017).
50. Jeannet, G. et al. Long-term, multilineage hematopoiesis occurs in the combined absence of β-catenin and γ-catenin. *Blood.* **111**, 142–149 (2008).
51. Tiemessen, M. M. et al. The nuclear effector of Wnt-signaling, Tcf1, functions as a T-cell-specific tumor suppressor for development of lymphomas. *PLoS Biol.* **10**, e1001430 (2012).
52. Yu, S. et al. The TCF-1 and LEF-1 transcription factors have cooperative and opposing roles in T cell development and malignancy. *Immunity.* **37**, 813–826 (2012).
53. Zaret, K. S. & Carroll, J. S. Pioneer transcription factors: establishing competence for gene expression. *Genes Dev.* **25**, 2227–2241 (2011).
54. Emmanuel, A. O. et al. TCF-1 and HEB cooperate to establish the epigenetic and transcription profiles of $CD4^+CD8^+$ thymocytes. *Nat. Immunol.* **19**, 1366–1378 (2018).
The authors demonstrate the close cooperation between TCF1 and the bHLH factor HEB in establishing open chromatin and driving gene expression in DP thymocytes. TCF1 is shown to enhance the binding of HEB to many of its genomic sites and to promote HEB accumulation by blunting Notch signals that otherwise degrade HEB.
55. Frelin, C. et al. GATA-3 regulates the self-renewal of long-term hematopoietic stem cells. *Nat. Immunol.* **14**, 1037–1044 (2013).
56. Tindemans, I., Serafini, N., Di Santo, J. P. & Hendriks, R. W. GATA-3 function in innate and adaptive immunity. *Immunity.* **41**, 191–206 (2014).
57. Hasegawa, S. L. et al. Dosage-dependent rescue of definitive nephrogenesis by a distant *Gata3* enhancer. *Dev. Biol.* **301**, 568–577 (2007).
58. Lim, K. C. et al. *Gata3* loss leads to embryonic lethality due to noradrenaline deficiency of the sympathetic nervous system. *Nat. Genet.* **25**, 209–212 (2000).
59. Kouros-Mehr, H., Slorach, E. M., Sternlicht, M. D. & Werb, Z. GATA-3 maintains the differentiation of the luminal cell fate in the mammary gland. *Cell* **127**, 1041–1055 (2006).
60. Van Esch, H. et al. GATA3 haplo-insufficiency causes human HDR syndrome. *Nature* **406**, 419–422 (2000).
61. Hozumi, K. et al. Notch signaling is necessary for GATA3 function in the initiation of T cell development. *Eur. J. Immunol.* **38**, 977–985 (2008).

62. Garcia-Ojeda, M. E. et al. GATA-3 promotes T-cell specification by repressing B-cell potential in pro-T cells in mice. *Blood* **121**, 1749–1759 (2013).
63. Scripture-Adams, D. D. et al. GATA-3 dose-dependent checkpoints in early T cell commitment. *J. Immunol.* **193**, 3470–3491 (2014).
64. Hosoya, T. et al. Global dynamics of stage-specific transcription factor binding during thymocyte development. *Sci. Rep.* **8**, 5605 (2018).
65. Ohmura, S. et al. Lineage-affiliated transcription factors bind the *Gata3* Tce1 enhancer to mediate lineage-specific programs. *J. Clin. Invest.* **126**, 865–878 (2016).
66. Xu, W. et al. E2A transcription factors limit expression of Gata3 to facilitate T lymphocyte lineage commitment. *Blood* **121**, 1534–1542 (2013).
67. Taghon, T., Yui, M. A. & Rothenberg, E. V. Mast cell lineage diversion of T lineage precursors by the essential T cell transcription factor GATA-3. *Nat. Immunol.* **8**, 845–855 (2007).
68. Hosoya, T., Maillard, I. & Engel, J. D. From the cradle to the grave: activities of GATA-3 throughout T-cell development and differentiation. *Immunol. Rev.* **238**, 110–125 (2010).
69. Zhu, J. GATA3 regulates the development and functions of innate lymphoid cell subsets at multiple stages. *Front. Immunol.* **8**, 1571 (2017).
70. Wei, G. et al. Genome-wide analyses of transcription factor GATA3-mediated gene regulation in distinct T cell types. *Immunity* **35**, 299–311 (2011).
71. Nakayama, T. et al. Th2 cells in health and disease. *Annu. Rev. Immunol.* **35**, 53–84 (2017).
72. Furusawa, J. et al. Critical role of p38 and GATA3 in natural helper cell function. *J. Immunol.* **191**, 1818–1826 (2013).
73. Hosokawa, H. et al. Methylation of Gata3 protein at Arg-261 regulates transactivation of the Il5 gene in T helper 2 cells. *J. Biol. Chem.* **290**, 13095–13103 (2015).
74. Hosokawa, H. et al. Akt1-mediated Gata3 phosphorylation controls the repression of IFNγ in memory-type Th2 cells. *Nat. Commun.* **7**, 11289 (2016).
75. Hosokawa, H. et al. Functionally distinct Gata3/Chd4 complexes coordinately establish T helper 2 (Th2) cell identity. *Proc. Natl Acad. Sci. USA* **110**, 4691–4696 (2013).
76. Yamagata, T. et al. Acetylation of GATA-3 affects T-cell survival and homing to secondary lymphoid organs. *EMBO J.* **19**, 4676–4687 (2000).
77. Buono, M. et al. A dynamic niche provides Kit ligand in a stage-specific manner to the earliest thymocyte progenitors. *Nat. Cell Biol.* **18**, 157–167 (2016).
78. Semerad, C. L., Mercer, E. M., Inlay, M. A., Weissman, I. L. & Murre, C. E2A proteins maintain the hematopoietic stem cell pool and promote the maturation of myelolymphoid and myeloerythroid progenitors. *Proc. Natl Acad. Sci. USA* **106**, 1930–1935 (2009).
79. Yang, Q. et al. E47 controls the developmental integrity and cell cycle quiescence of multipotential hematopoietic progenitors. *J. Immunol.* **181**, 5885–5894 (2008).
80. Dias, S., Mansson, R., Gurbuxani, S., Sigvardsson, M. & Kee, B. L. E2A proteins promote development of lymphoid-primed multipotent progenitors. *Immunity* **29**, 217–227 (2008).
81. Miyazaki, M. et al. The E-Id protein axis specifies adaptive lymphoid cell identity and suppresses thymic innate lymphoid cell development. *Immunity* **46**, 818–834 e4 (2017).
E protein activity in ILCs is tightly inhibited by ID2. This study shows that inappropriate expression of *Id2* or deletion of both *Tcf3* and *Tcf12* induces abnormal development of ILCs in the thymus. Thus, ID2 acts primarily through E protein neutralization to segregate T cells and ILCs.
82. Murre, C. Helix-loop-helix proteins and the advent of cellular diversity: 30 years of discovery. *Genes Dev.* **33**, 6–25 (2019).
83. Braunstein, M. & Anderson, M. K. HEB in the spotlight: transcriptional regulation of T-cell specification, commitment, and developmental plasticity. *Clin. Dev. Immunol.* **2012**, 678705 (2012).
84. Bain, G. et al. E2A deficiency leads to abnormalities in αβ T-cell development and to rapid development of T-cell lymphomas. *Mol. Cell Biol.* **17**, 4782–4791 (1997).
85. Wojciechowski, J., Lai, A., Kondo, M. & Zhuang, Y. E2A and HEB are required to block thymocyte proliferation prior to pre-TCR expression. *J. Immunol.* **178**, 5717–5726 (2007).
86. Ikawa, T., Kawamoto, H., Goldrath, A. W. & Murre, C. E proteins and notch signaling cooperate to promote T cell lineage specification and commitment. *J. Exp. Med.* **203**, 1329–1342 (2006).
87. Pereira de Sousa, A. et al. Inhibitors of DNA binding proteins restrict T cell potential by repressing Notch1 expression in Flt3-negative common lymphoid progenitors. *J. Immunol.* **189**, 3822–3830 (2012).
88. Yashiro-Ohtani, Y. et al. Pre-TCR signaling inactivates Notch1 transcription by antagonizing E2A. *Genes Dev.* **23**, 1665–1676 (2009).
89. Zook, E. C. & Kee, B. L. Development of innate lymphoid cells. *Nat. Immunol.* **17**, 775–782 (2016).
90. Moro, K. et al. Innate production of T_H2 cytokines by adipose tissue-associated c-Kit$^+$Sca-1$^+$ lymphoid cells. *Nature* **463**, 540–544 (2010).
91. Koues, O. I. et al. Distinct gene regulatory pathways for human innate versus adaptive lymphoid cells. *Cell* **165**, 1134–1146 (2016).
92. Boos, M. D., Yokota, Y., Eberl, G. & Kee, B. L. Mature natural killer cell and lymphoid tissue-inducing cell development requires Id2-mediated suppression of E protein activity. *J. Exp. Med.* **204**, 1119–1130 (2007).
93. Hosokawa, H. et al. Bcl11b sets pro-T cell fate by site-specific cofactor recruitment and by repressing *Id2* and *Zbtb16*. *Nat. Immunol.* **19**, 1427–1440 (2018).
This study reveals the molecular basis of both direct and indirect BCL-11B actions to promote T cell identity and to avoid alternative lineage phenotypes in the thymus. Among the BCL-11B-repressed targets, *Id2* and *Zbtb16* are shown to be functionally important to control distinct alternative programmes.
94. Ikawa, T., Fujimoto, S., Kawamoto, H., Katsura, Y. & Yokota, Y. Commitment to natural killer cells requires the helix-loop-helix inhibitor Id2. *Proc. Natl Acad. Sci. USA* **98**, 5164–5169 (2001).
95. Zook, E. C. et al. Transcription factor ID2 prevents E proteins from enforcing a naive T lymphocyte gene program during NK cell development. *Sci. Immunol.* **3**, eaao2139 (2018).
96. Miyazaki, M. et al. The opposing roles of the transcription factor E2A and its antagonist Id3 that orchestrate and enforce the naive fate of T cells. *Nat. Immunol.* **12**, 992–1001 (2011).
97. Jones, M. E. & Zhuang, Y. Stage-specific functions of E-proteins at the β-selection and T-cell receptor checkpoints during thymocyte development. *Immunol. Res.* **49**, 202–215 (2011).
98. Jones, M. E. & Zhuang, Y. Regulation of V(D)J recombination by E-protein transcription factors. *Adv. Exp. Med. Biol.* **650**, 148–156 (2009).
99. Jones, M. E. & Zhuang, Y. Acquisition of a functional T cell receptor during T lymphocyte development is enforced by HEB and E2A transcription factors. *Immunity* **27**, 860–870 (2007).
100. Bain, G. et al. Regulation of the helix-loop-helix proteins, E2A and Id3, by the Ras-ERK MAPK cascade. *Nat. Immunol.* **2**, 165–171 (2001).
101. Jones-Mason, M. E. et al. E protein transcription factors are required for the development of CD4$^+$ lineage T cells. *Immunity* **36**, 348–361 (2012).
102. Phelan, J. D. et al. Growth factor independent-1 maintains Notch1-dependent transcriptional programming of lymphoid precursors. *PLoS Genet.* **9**, e1003713 (2013).
103. Emambokus, N. et al. Progression through key stages of haemopoiesis is dependent on distinct threshold levels of c-Myb. *EMBO J.* **22**, 4478–4488 (2003).
104. Allen, R. D. III, Bender, T. P. & Siu, G. c-Myb is essential for early T cell development. *Genes Dev.* **13**, 1073–1078 (1999).
105. Georgopoulos, K. The making of a lymphocyte: the choice among disparate cell fates and the IKAROS enigma. *Genes Dev.* **31**, 439–450 (2017).
106. Ebihara, T., Seo, W. & Taniuchi, I. Roles of RUNX complexes in immune cell development. *Adv. Exp. Med. Biol.* **962**, 395–413 (2017).
107. Guo, Y., Maillard, I., Chakraborti, S., Rothenberg, E. V. & Speck, N. A. Core binding factors are necessary for natural killer cell development, and cooperate with Notch signaling during T cell specification. *Blood* **112**, 480–492 (2008).
108. Geimer Le Lay, A. S. et al. The tumor suppressor Ikaros shapes the repertoire of Notch target genes in T cells. *Sci. Signal.* **7**, ra28 (2014).
109. Winandy, S., Wu, L., Wang, J. H. & Georgopoulos, K. Pre-T cell receptor (TCR) and TCR-controlled checkpoints in T cell differentiation are set by Ikaros. *J. Exp. Med.* **190**, 1039–1048 (1999).
110. Arenzana, T. L., Schjerven, H. & Smale, S. T. Regulation of gene expression dynamics during developmental transitions by the Ikaros transcription factor. *Genes Dev.* **29**, 1801–1816 (2015).
111. Hosokawa, H. et al. Transcription factor PU.1 represses and activates gene expression in early T cells by redirecting partner transcription factor binding. *Immunity* **48**, 1119–1134 e7 (2018).
PU.1 directly activates its target genes in part by recruiting RUNX1 and SATB1. The result is to 'steal' RUNX1 and SATB1 from sites associated with T cell lineage-related genes, thus repressing some of these genes, indirectly, without direct DNA binding.
112. David-Fung, E. S. et al. Transcription factor expression dynamics of early T-lymphocyte specification and commitment. *Dev. Biol.* **325**, 444–467 (2009).
113. Talebian, L. et al. T-lymphoid, megakaryocyte, and granulocyte development are sensitive to decreases in CBFβ dosage. *Blood* **109**, 11–21 (2007).
114. Cortes, M., Wong, E., Koipally, J. & Georgopoulos, K. Control of lymphocyte development by the Ikaros gene family. *Curr. Opin. Biol.* **11**, 167–171 (1999).
115. Liu, P., Li, P. & Burke, S. Critical roles of Bcl11b in T-cell development and maintenance of T-cell identity. *Immunol. Rev.* **238**, 138–149 (2010).
116. Kominami, R. Role of the transcription factor Bcl11b in development and lymphomagenesis. *Proc. Jpn. Acad. Ser. B Phys. Biol. Sci.* **88**, 72–87 (2012).
117. Kastner, P. et al. Bcl11b represses a mature T-cell gene expression program in immature CD4$^+$CD8$^+$ thymocytes. *Eur. J. Immunol.* **40**, 2143–2154 (2010).
118. Li, L., Leid, M. & Rothenberg, E. V. An early T cell lineage commitment checkpoint dependent on the transcription factor Bcl11b. *Science* **329**, 89–93 (2010).
119. Shibata, K. et al. IFN-γ-producing and IL-17-producing γδ T cells differentiate at distinct developmental stages in murine fetal thymus. *J. Immunol.* **192**, 2210–2218 (2014).
120. De Obaldia, M. E. & Bhandoola, A. Transcriptional regulation of innate and adaptive lymphocyte lineages. *Annu. Rev. Immunol.* **33**, 607–642 (2015).
121. Avram, D. & Califano, D. The multifaceted roles of Bcl11b in thymic and peripheral T cells: impact on immune diseases. *J. Immunol.* **193**, 2059–2065 (2014).
122. Fang, D. et al. Bcl11b, a novel GATA3-interacting protein, suppresses Th1 while limiting Th2 cell differentiation. *J. Exp. Med.* **215**, 1449–1462 (2018).
123. Punwani, D. et al. Multisystem anomalies in severe combined immunodeficiency with mutant BCL11B. *N. Engl. J. Med.* **375**, 2165–2176 (2016).
124. Hirose, S. et al. Bcl11b prevents the intrathymic development of innate CD8 T cells in a cell intrinsic manner. *Int. Immunol.* **27**, 205–215 (2015).
125. Kojo, S. et al. Priming of lineage-specifying genes by Bcl11b is required for lineage choice in post-selection thymocytes. *Nat. Commun.* **8**, 702 (2017).
126. Longabaugh, W. J. R. et al. Bcl11b and combinatorial resolution of cell fate in the T-cell gene regulatory network. *Proc. Natl Acad. Sci. USA* **114**, 5800–5807 (2017).
127. Ikawa, T. et al. An essential developmental checkpoint for production of the T cell lineage. *Science* **329**, 93–96 (2010).
128. Li, P. et al. Reprogramming of T cells to natural killer-like cells upon Bcl11b deletion. *Science* **329**, 85–89 (2010).
129. Kadoch, C. & Crabtree, G. R. Mammalian SWI/SNF chromatin remodeling complexes and cancer: Mechanistic insights gained from human genomics. *Sci. Adv.* **1**, e1500447 (2015).
130. Cismasiu, V. B. et al. BCL11B functionally associates with the NuRD complex in T lymphocytes to repress targeted promoter. *Oncogene* **24**, 6753–6764 (2005).
131. Hosokawa, H. et al. Cell type-specific actions of Bcl11b in early T-lineage and group 2 innate lymphoid cells. *J. Exp. Med.* **217**, e20190972 (2020).
***Bcl11b* is expressed in pro-T cells and ILC2s. *Id2* is an important repressed target gene of BCL-11B in pro-T cells, whereas *Id2* must be co-expressed with *Bcl11b* in ILC2s. This study show that BCL-11B regulates distinct sets of genes in pro-T cells and ILC2s by organizing distinct protein complexes and binding to different genomic regions.**

132. Ng, K. K. et al. A stochastic epigenetic switch controls the dynamics of T-cell lineage commitment. *eLife* **7**, e37851 (2018).
This article shows that *cis*-acting epigenetic constraints cause delays of the order of days in the activation of *Bcl11b* for T cell lineage commitment. The authors establish two-colour *Bcl11b* reporter mice to monitor the activation status and enhancer dependence of *Bcl11b* alleles separately.
133. Li, L. et al. A far downstream enhancer for murine Bcl11b controls its T-cell specific expression. *Blood* **122**, 902–911 (2013).
134. Rothenberg, E. V. Dynamic control of the T-cell specification gene regulatory network. *Curr. Opin. Syst. Biol.* **18**, 62–76 (2019).
135. Champhekar, A. et al. Regulation of early T-lineage gene expression and developmental progression by the progenitor cell transcription factor PU.1. *Genes Dev.* **29**, 832–848 (2015).
136. Rothenberg, E. V., Hosokawa, H. & Ungerback, J. Mechanisms of action of hematopoietic transcription factor PU.1 in initiation of T-cell development. *Front. Immunol.* **10**, 228 (2019).
137. Kawamoto, H. et al. Extensive proliferation of T cell lineage-restricted progenitors in the thymus: an essential process for clonal expression of diverse T cell receptor β chains. *Eur. J. Immunol.* **33**, 606–615 (2003).
138. Scott, E. W., Simon, M. C., Anastasi, J. & Singh, H. Requirement of transcription factor PU.1 in the development of multiple hematopoietic lineages. *Science* **265**, 1573–1577 (1994).
139. Kueh, H. Y., Champhekar, A., Nutt, S. L., Elowitz, M. B. & Rothenberg, E. V. Positive feedback between PU.1 and the cell cycle controls myeloid differentiation. *Science* **341**, 670–673 (2013).
140. Heinz, S. et al. Effect of natural genetic variation on enhancer selection and function. *Nature* **503**, 487–492 (2013).
141. Heinz, S. et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589 (2010).
142. Ostuni, R. et al. Latent enhancers activated by stimulation in differentiated cells. *Cell* **152**, 157–171 (2013).
143. Escalante, C. R. et al. Crystal structure of PU.1/IRF-4/DNA ternary complex. *Mol. Cell* **10**, 1097–1105 (2002).
144. Minderjahn, J. et al. Mechanisms governing the pioneering and redistribution capabilities of the non-classical pioneer PU.1. *Nat. Commun.* **11**, 402 (2020).
145. Carotta, S. et al. The transcription factor PU.1 controls dendritic cell development and Flt3 cytokine receptor expression in a dose-dependent manner. *Immunity* **32**, 628–641 (2010).
146. Seki, M. et al. Recurrent SPI1 (PU.1) fusions in high-risk pediatric T cell acute lymphoblastic leukemia. *Nat. Genet.* **49**, 1274–1281 (2017).
147. Rosenbauer, F. et al. Lymphoid cell growth and transformation are suppressed by a key regulatory element of the gene encoding PU.1. *Nat. Genet.* **38**, 27–37 (2006).
148. Franco, C. B. et al. Notch/Delta signaling constrains reengineering of pro-T cells by PU.1. *Proc. Natl Acad. Sci. USA* **103**, 11993–11998 (2006).
149. Laiosa, C. V., Stadtfeld, M., Xie, H., de Andres-Aguayo, L. & Graf, T. Reprogramming of committed T cell progenitors to macrophages and dendritic cells by C/EBPα and PU.1 transcription factors. *Immunity* **25**, 731–744 (2006).
150. Yu, Y. et al. Bcl11a is essential for lymphoid development and negatively regulates p53. *J. Exp. Med.* **209**, 2467–2483 (2012).
151. Zohren, F. et al. The transcription factor Lyl-1 regulates lymphoid specification and the maintenance of early T lineage progenitors. *Nat. Immunol.* **13**, 761–769 (2012).
152. Yui, M. A., Feng, N. & Rothenberg, E. V. Fine-scale staging of T cell lineage commitment in adult mouse thymus. *J. Immunol.* **185**, 284–293 (2010).
153. Masuda, K. et al. T cell lineage determination precedes the initiation of TCR β gene rearrangement. *J. Immunol.* **179**, 3699–3706 (2007).
154. Bell, J. J. & Bhandoola, A. The earliest thymic progenitors for T cells possess myeloid lineage potential. *Nature* **452**, 764–767 (2008).
155. Wada, H. et al. Adult T-cell progenitors retain myeloid potential. *Nature* **452**, 768–772 (2008).
156. De Obaldia, M. E., Bell, J. J. & Bhandoola, A. Early T-cell progenitors are the major granulocyte precursors in the adult mouse thymus. *Blood* **121**, 64–71 (2013).
157. Zarnegar, M. A., Chen, J. & Rothenberg, E. V. Cell-type-specific activation and repression of PU.1 by a complex of discrete, functionally specialized cis-regulatory elements. *Mol. Cell. Biol.* **30**, 4922–4939 (2010).
158. Huang, G. et al. PU.1 is a major downstream target of AML1 (RUNX1) in adult mouse hematopoiesis. *Nat. Genet.* **40**, 51–60 (2008).
159. Kueh, H. Y. et al. Asynchronous combinatorial action of four regulatory factors activates Bcl11b for T cell commitment. *Nat. Immunol.* **17**, 956–965 (2016).
160. Rothenberg, E. V. Programming for T-lymphocyte fates: modularity and mechanisms. *Genes Dev.* **33**, 1117–1135 (2019).
161. Isoda, T. et al. Non-coding transcription instructs chromatin folding and compartmentalization to dictate enhancer-promoter communication and T cell fate. *Cell* **171**, 103–119 e18 (2017).
The authors focus on the distal superenhancer region of *Bcl11b* and show that it harbours a functionally important long non-coding RNA that acts positively on gene regulation. The authors find a severe, allele-specific loss of *Bcl11b* expression when long non-coding RNA transcription is truncated.
162. Manesso, E., Chickarmane, V., Kueh, H. Y., Rothenberg, E. V. & Peterson, C. Computational modelling of T-cell formation kinetics: output regulated by initial proliferation-linked deferral of developmental competence. *J. R. Soc. Interface* **10**, 20120774 (2013).
163. Krueger, A., Ziętara, N. & Łyszkiewicz, M. T cell development by the numbers. *Trends Immunol.* **38**, 128–139 (2017).
164. Cumano, A. et al. New molecular insights into immune cell development. *Annu. Rev. Immunol.* **37**, 497–519 (2019).
165. Casero, D. et al. Long non-coding RNA profiling of human lymphoid progenitor cells reveals transcriptional divergence of B cell and T cell lineages. *Nat. Immunol.* **16**, 1282–1291 (2015).
166. Ha, V. L. et al. The T-ALL related gene BCL11B regulates the initial stages of human T-cell differentiation. *Leukemia* **31**, 2503–2514 (2017).
167. Cantê-Barrett, K. et al. Loss of $CD44^{dim}$ expression from early progenitor cells marks T-cell lineage commitment in the human thymus. *Front. Immunol.* **8**, 32 (2017).
168. Garvie, C. W., Pufall, M. A., Graves, B. J. & Wolberger, C. Structural analysis of the autoinhibition of Ets-1 and its role in protein partnerships. *J. Biol. Chem.* **277**, 45529–45536 (2002).
169. Pufall, M. A. & Graves, B. J. Ets-1 flips for new partner Pax-5. *Structure* **10**, 11–14 (2002).
170. Soufi, A., Donahue, G. & Zaret, K. S. Facilitators and impediments of the pluripotency reprogramming factors' initial engagement with the genome. *Cell* **151**, 994–1004 (2012).
171. Iwafuchi-Doi, M. & Zaret, K. S. Cell fate control by pioneer transcription factors. *Development* **143**, 1833–1837 (2016).
172. Chronis, C. et al. Cooperative binding of transcription factors orchestrates reprogramming. *Cell* **168**, 442–459 (2017).
173. Weikum, E. R., Knuesel, M. T., Ortlund, E. A. & Yamamoto, K. R. Glucocorticoid receptor control of transcription: precision and plasticity via allostery. *Nat. Rev. Mol. Cell. Biol.* **18**, 159–174 (2017).
174. Luna-Zurita, L. et al. Complex interdependence regulates heterotypic transcription factor distribution and coordinates cardiogenesis. *Cell* **164**, 999–1014 (2016).
175. Zhang, J. et al. Harnessing of the nucleosome-remodeling-deacetylase complex controls lymphocyte development and prevents leukemogenesis. *Nat. Immunol.* **13**, 86–94 (2011).
176. Kanhere, A. et al. T-bet and GATA3 orchestrate Th1 and Th2 differentiation through lineage-specific targeting of distal regulatory elements. *Nat. Commun.* **3**, 1268 (2012).
177. Garcia-Perez, L. et al. Functional definition of a transcription factor hierarchy regulating T cell lineage commitment. *Sci. Adv.* **6**, eaaw7313 (2020).

Acknowledgements
The authors thank M. Romero-Wolf for helpful discussions and suggestions, and J. Ungerbäck, X. Wang, M. A. Yui and present and former members of the Rothenberg group, whose helpful discussion and published and unpublished data were important for the ideas in this Review. The authors apologize to colleagues whose relevant work could not be cited owing to space constraints. The authors gratefully acknowledge support from the Japan Society for the Promotion of Science KAKENHI (grant number JP19H03692), the Mochida Memorial Foundation for Medical and Pharmaceutical Research, the Naito Foundation and the Takeda Science Foundation (to H.H.), and from the US Public Health Service (R01AI135200, R01HL119102, R01HD100039 and R01HD076915) and the Albert Billings Ruddock Professorship (to E.V.R).

Author contributions
Both authors contributed to research and discussion of the content as well as the writing and reviewing of the manuscript.

Competing interests
The authors declare no competing interests.

Peer review information
Nature Reviews Immunology thanks J. C. Zúñiga-Pflücker and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

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

© Springer Nature Limited 2020