LA-UR- 08-5371

Approved for public release; distribution is unlimited.

Title:	A Gene Regulatory Network Armature for T-Lymphocyte Specification.
Author(s):	E. Fung, Z # 222887, T-10/T Division
Intended for:	Journal: Proceeding of the National Academy of Sc.

Los Alamos
NATIONAL LABORATORY
EST. 1943

Form 836 (7/06)

A Gene Regulatory Network Armature for T-Lymphocyte Specification

Constantin Georgescu*, William J. R. Longabaugh†, Deirdre D. Scripture-Adams*, Elizabeth-Sharon David-Fung*, Mary A. Yui*, Mark A. Zarnegar*, Hamid Bolouri*†‡, Ellen V. Rothenberg*‡ will be revised to include: Theoretical Biology and Biophysics, Los Alamos, NM 87544.

*Division of Biology 156-29, California Institute of Technology, Pasadena, CA 91125

†Institute for Systems Biology, 1441 N. 34th Street, Seattle, WA 98103

‡Corresponding authors: Ellen Rothenberg, Tel 626-395-4992, Fax 626-449-0756, E-mail evroth@its.caltech.edu; Hamid Bolouri, Tel 206-339-4852, Fax 206-339-4852, E-mail hbolouri@caltech.edu

Abstract

Choice of a T-lymphoid fate by hematopoietic progenitor cells depends on sustained Notch-Delta signaling combined with tightly-regulated activities of multiple transcription factors. To dissect the regulatory network connections that mediate this

process, we have used high-resolution analysis of regulatory gene expression trajectories from the beginning to the end of specification; tests of the short-term Notch-dependence of these gene expression changes; and perturbation analyses of the effects of overexpression of two essential transcription factors, namely PU.1 and GATA-3. Quantitative expression measurements of >50 transcription factor and marker genes have been used to derive the principal components of regulatory change through which T-cell precursors progress from primitive multipotency to T-lineage commitment. Distinct parts of the path reveal separate contributions of Notch signaling, GATA-3 activity, and downregulation of PU.1. Using BioTapestry, the results have been assembled into a draft gene regulatory network for the specification of T-cell precursors and the choice of T as opposed to myeloid/dendritic or mast-cell fates. This network also accommodates effects of E proteins and mutual repression circuits of Gfi1 against Egr-2 and of TCF-1 against PU.1 as proposed elsewhere, but requires additional functions that remain unidentified. Distinctive features of this network structure include the intense dose-dependence of GATA-3 effects; the gene-specific modulation of PU.1 activity based on Notch activity; the lack of direct opposition between PU.1 and GATA-3; and the need for a distinct, late-acting repressive function or functions to extinguish stem and progenitor-derived regulatory gene expression.

\body

Introduction

Exclusion of alternative fates is integral to cell type specification and one of the key features explained by the gene regulatory networks for development in well-studied embryological systems. Cell-type specific gene activation is tightly coupled with prohibition of alternative gene programs, through three basic elements of gene network architecture: positive autoregulation of major cell type-specific transcription factors, feed-forward relationships between these factors and their collaborators, and mutual antagonisms between the drivers of alternative cell fates. The collective impact of these mechanisms is usually to create within tight spatial and temporal boundaries a swift cascade of regulatory changes that become effectively irreversible. Yet this is not the only way that cell type specification can occur. In stem-cell based systems such as exist in adult mammals, the multipotent state is actively maintained over an extended number of cell cycles. Even as differentiation of these precursors begins, there can be considerable delay before the cell fate decision is determined. For example, many of the cell fate decisions of mouse hematopoietic stem cell progeny may be controlled by dynamic balance of regulatory factors such as PU.1, C/EBPα, and GATA-2 through the intermediate stages of the process. Even in collaboration, these factors appear to drive up to four different cell fates depending on the ratios and fluxes of their activities (1, 2). This behavior is a clue that a distinctive gene network architecture may lie at the core of stem-cell based cell-type specification.

A good place to dissect this mode of specification is in mammalian T-lymphocyte development. T-cell precursors acquire their specific immunological function through a mechanism that preserves stem-cell-like features, such as an ongoing variety of developmental options and a capacity for extensive, though tightly regulated, proliferation even after the cells become committed to a T-cell fate. T-cell development begins with the migration of hematopoietic precursors into the specialized microenvironment of the thymus, where these multipotent cells adopt T-lineage characteristics and gradually give up the ability to give rise to other kinds of blood cells. The early stages of this process are well marked and experimentally accessible. Much work shows that lineage exclusion is not only slow but discontinuous for T-cell precursors in the thymus: there is a delay of multiple cell cycles between the time cells lose certain non-T options (red blood cell, B cell) and the time they finally become committed to a T-cell fate [rev. in (3)]. T-cell specification emerges through the response to a combination of at least eight transcription factors acting under the influence of Notch pathway signals from the thymic microenvironment. The challenge has been to understand the mechanisms operating in this multicomponent system.

Here, we have sought to make explicit the regulatory structures and some aspects of combinatorial control that underlie T-lineage specification in mice. This synthesis combines evidence from the following: (*i*) purification of staged T-cell precursors, taking advantage of the discontinuities from the stem cell through the commitment stage; (*ii*) definition of multiple transcription factor gene expression changes that distinguish these stages *in vivo*; (*iii*) characterization of the impacts of

Notch signaling on gene expression at individual stages, using *in vitro* culture systems to control delivery of Notch signals; and (*iv*) perturbation analysis based on manipulation of two key transcription factors that are thought to drive opposing network subcircuits in the T-cell development process. We compare the likely inputs of three regulators on the developmental trajectory of the cells and present a combinatorial map of regulatory connections, incorporating work from other groups, as a testable framework for reconstructing the full process.

RESULTS AND DISCUSSION

Early T-cell developmental progression through regulatory gene expression space

Mouse T cell precursors traverse a canonical sequence of stages between entry into the thymus and full commitment to the T-cell lineage. They are distinguishable by changes in surface markers and also robustly associated with quantitatively distinct patterns of gene expression. The same stages are used in fetal T-cell development, adult T-cell development, and when T-cell precursors are induced to differentiate in stromal culture *in vitro*. Cells proliferate in each of the first two or three stages for several days before moving on to the next one. The gene expression signatures of each stage remain consistent but the kinetics of the progression differ among fetal, adult, and *in vitro* development, (4-6)(D.D.S.-A. unpublished data). Thus, in contrast to embryonic systems where autonomous progression from state to state is "hard wired" in the

regulatory circuitry, T-cell specification is inherently discontinuous: progression between states may depend on repeated microenvironmental stimulation (7).

The characteristic gene expression patterns of these stages are benchmarks for comparison of normal and perturbed versions of T-cell development (Table 1). To represent these complex changes in a simpler form, we have used expression profiles of over 140 genes (4-6)(E.-S.D.-F., R. Butler, M. Morales, and M.A.Y., unpublished) to calculate principal components of regulatory change across the T-lineage specification process, from the earliest, still multipotent intrathymic precursors (DN1 cells) through commitment (to DN3a cells), then through the first T-cell receptor signaling responses, "β-selection" (to DN4). Gene expression levels used for these calculations were based on quantitative real-time PCR measurements from four representative sample series (from 6-8 series analyzed overall): first, the c-Kit$^+$ "DN1" or "Early T-cell Precursor" (ETP) stage; second, the c-Kit$^+$ CD25$^+$ CD44$^+$ "DN2" stage; third, the CD25$^+$ CD44$^-$ CD27-low "DN3a" stage; fourth, the CD25$^+$ CD44$^-$ CD27$^+$ "DN3b" stage; and fifth, the CD25$^-$ CD44$^-$ "DN4" stage (8)(and E.-S.D.-F. and M.A.Y., unpublished). By partial least squares analysis, we calculated coordinates for each gene and for each stage along axes representing the first four principal components of variance (Supplementary Table 1).

As shown in Fig. 1A and SFig. 1 online, the distinctive gene expression profile for each stage defines a vector from the origin in this principal component space. Most of the first principal component accounts for differences between the DN1 to the DN3a/3b stages, and is dominated by the downregulation of prethymically expressed regulatory genes such as *Sfpi1* (encoding PU.1), *Tal1* (encoding SCL), *Gata2*, and *Gfi1b*. Most of the

second principal component accounts for differences between DN3a cells and β-selected cells, with subtler distinctions (e.g. DN1 vs. DN2) emphasized in the third and fourth principal components.

The developmental process can be seen as a path connecting the tips of these stage vectors. The first part of this path, till commitment, is most nearly parallel to the first principal component axis; then, successful T-cell receptor expression triggers a shift along the second principal component axis. The intermediate DN2 stage involves a distinct excursion along the third axis (SFig. 1). This trajectory is what we seek to explain in terms of transcriptional linkages in a gene regulatory network for T-cell specification. Three vital nodes of the network are examined in this paper.

PU.1 and GATA-3 as regulatory inputs

The transcription factors PU.1 and GATA-3 are both required for early T-cell development. Loss of either transcription factor from a prethymic stage virtually eliminates T-cell development, but overexpression of either one in early intrathymic stages also inhibits, blocking the generation of cells capable of undergoing β-selection. When overexpressed, both can push DN thymocytes toward alternative hematopoietic developmental fates: PU.1 toward the dendritic cell or monocytic lineages, GATA-3 toward the mast-cell lineage. High doses of either PU.1 or GATA-3 inhibit particular sets of T-cell genes while activating distinct sets of non-T genes (9-11).

These factors have been expected to oppose one another in T-cell development, based on a key precedent in other blood cell fate decisions. The related GATA factor

GATA-1 and PU.1 apparently act as mutually inhibitory competitive antagonists in a bistable switch to control the choice between erythroid or megakaryocytic fates, on the one hand, and all myeloid or lymphoid cell fates, on the other hand [rev. in (12-14)]. Notably, GATA-1 and PU.1 titrate each other's activities antagonistically at the protein level, while promoting their own respective expression via positive transcriptional autoregulation. To adapt such a model to the genetic requirement for both PU.1 and GATA-3 in T-cell development, PU.1 might support gene expression associated with "immaturity" whereas GATA-3 could promote gene expression associated with T lineage commitment. The expression patterns and functions of PU.1 and GATA-3 are indeed divergent (5, 6, 15, 16): from a high initial level of expression, PU.1 is sharply and permanently downregulated during T-lineage commitment (DN1-DN3a) *in vivo*, while GATA-3 rises gradually and functions repeatedly throughout T-cell development (Table 1). These expression patterns are situated near opposite ends of the first principal component axis in Fig. 1A; analyzing detailed impacts of exogenous PU.1 or GATA-3 should detect whether opposition between these two factors actually controls the position of cells along this axis.

As T-cell development is blocked when either PU.1 or GATA-3 is overexpressed, and since either factor at high level can sequester the other, double overexpression experiments may not prove that antagonism is relevant *in vivo*. However, if PU.1 and GATA-3 did titrate each other in normal thymocytes, each should normally be limiting the other's activity during the DN1 and DN2 stages when both are present. Then, experimentally increased expression of either factor should oppose the distinctive

pattern of gene expression effects attributable to the other factor (Fig. 1B,C). The magnitudes of the PU.1 and GATA-3 effects described and the statistical threshold for significance are shown in SFig. 2 and Supp. Table 2 online.

The overexpression effects of PU.1 generally go against the trend of the DN1 to DN3a progression (Fig. 1C), as PU.1 inhibits a wide range of genes utilized in the DN3 stages. Although these may not all be direct targets, they support the idea that high-level endogenous PU.1 in the DN1 stage could play a major role delaying gene expression changes associated with T-lineage progression, perhaps to allow continued self-renewal. Conversely, high-level GATA-3 not only slightly increases expression of a key T-lineage regulatory gene, Ikaros (*Ikzf1*), but also can repress PU.1 (*Sfpi1*) RNA expression in thymocytes (9, 11, 17)(Fig. 1B). This effect is confirmed by GATA-3 loss-of-function phenotype (D.D.S.-A., unpublished) and by effects of GATA-3 coexpression with a PU.1 cis-regulatory sequence reporter in myeloid cells (M.A.Z., unpublished). While additional regulatory inputs are likely needed to explain the steepness of PU.1 repression in thymocytes (18, 19)(M.A.Z., unpublished), GATA-3 may be a substantial contributor to the repression mechanism.

However, high-level GATA-3 does not show other effects expected for a T-lineage promotion factor. Even setting aside its activation of "non-T" mast cell genes (11) and focusing on T-lineage genes only, many effects of high-level GATA-3 also tend to oppose DN3a-specific gene expression (Fig. 1B). Consider the induction of key T-cell genes *Myb*, *Gfi1*, *Cd3e*, *Lck*, *LAT*, *Rag1*, *Tcf7* (encoding TCF-1), *Ets1*, or *Ets2* from DN1 to DN3a (Table 1). All of these genes are downregulated by exogenous high-level PU.1,

but they are not accelerated in their induction by elevating GATA-3; instead, these genes are all either unaffected or downregulated by high-level GATA-3 as well. Furthermore, *Lef1*, another DN3-stage upregulated gene, is actually inhibited by GATA-3 but unaffected by PU.1. Two additional features of the PU.1: GATA-3 subnetwork contrast with those of PU.1: GATA-1 interaction. First, PU.1 shows little or no ability to inhibit GATA-3 expression in thymocytes (10). Also, GATA-3 is quite unable to enhance its own expression in these cells (11).

These data imply that the relationship between PU.1 and GATA-3 in T-cell development is distinct from the bistable switch model imputed to PU.1 and GATA-1. If GATA-3 promotes progression to the DN3a state, it probably does so not by titrating PU.1 directly to relieve PU.1 repression of T-cell genes, but rather by helping to repress transcription of PU.1 itself. GATA-3 clearly must provide other unique regulatory functions for early T-cell development, but these have not emerged from GATA-3 overexpression experiments; they are likely governed by distinctive GATA-3 dose-response requirements for enhancesome assembly rather than by the opposition between GATA-3 and PU.1.

Notch-Delta signaling as a direct regulator of T-lineage specification genes

T-cell specification is highly dependent on at least one regulatory input from the environment, i.e. Notch pathway triggering by interaction with Delta-class ligands. Notch-Delta signaling is needed from the earliest DN1 stage through commitment and into β-selection. Its effects on gene expression can be quantitated by incubating defined

populations of immature thymocytes with stromal cells that either do (OP9-DL1) or do not (OP9-control) express Delta (20), then reisolating the thymocytes and measuring gene expression. The magnitudes of these effects, drawn from experiments which compared them with PU.1 and GATA-3 perturbations, are shown in SFig. 5 online. Fig. 1D shows that many of the regulatory effects, direct and indirect, of Notch-Delta signaling in thymocytes are to activate genes associated with the DN3-4 stages. The aggregate effects of Notch signaling on gene expression broadly appear "opposite" to those of PU.1 overexpression (cf. Fig. 1C). Thus Notch-Delta signaling appears to be more closely associated with DN3-specific regulatory events than does GATA-3 (see above).

This raises the question whether the DN3a stage phenotype is simply a reflection of increasing activation of direct Notch target genes, or to what extent additional factors also play a rate-limiting role. Fig. 2 shows a direct comparison of the developmental regulation and Notch-Delta dependence of multiple thymocyte genes; more extensive data are shown in SFig. 3 online. In Fig. 2, genes are shown in order of highest to lowest ratio of natural expression in DN3a to DN1 stage (measurements in two independent experiments shown as line graphs)(4, 5), while the effects of short-term Notch signaling on their expression (10, 11) are shown by bars. Known direct Notch target genes *Deltex1*, *Hes1*, and *Ptcra* (encoding the T-cell receptor surrogate α chain, preTα) that peak in expression at the DN3a stage all are strongly affected by the presence or absence of Notch-Delta interaction, as expected (Fig. 2). However, this is not the case for all T-lineage specification genes, such as the transcription factor genes *Gata3* and *Bcl11b*, and

the genes encoding T-cell receptor complex proteins *Cd3g* and *Cd3e* (5, 6). These genes are only induced in hematopoietic precursors through a Notch-dependent regulatory cascade (6, 21), so they show a weak Notch effect in cells just beginning specification (Fig. 2A). However, once activated in Thy-1$^+$ thymocytes (DN2 stage or beyond), these landmark genes and others with maximal expression at the DN3a stage become much less dependent on Notch-Delta interaction (Fig. 2B). Thus, regulatory inputs besides Notch (and GATA-3) contribute to the gene expression trajectory through the DN2 and DN3 stages, and these must be taken into account in the T-cell gene regulatory network.

Modification of regulatory factor effects by interactions with Notch

Notch signaling does make another kind of contribution to the regulatory state of the developing cells. This is to modulate the effects of other regulators in a gene-specific and factor-specific way. In prethymic hematopoietic precursors, Notch signaling synergizes with the effects of the basic helix-loop-helix factor E2A to enhance activation of T-lineage-associated genes such as *Ptcra* (22). In fetal thymocytes, Notch signaling modulates the effects of PU.1 selectively to relieve its repression of T-lineage genes (10). Note that cells can receive both Notch and PU.1 regulatory inputs independently without direct antagonism. Notch-Delta signaling does not repress *Sfpi1* (PU.1 expression), and even at high levels PU.1 does not inhibit the expression of Notch1 or Notch3 themselves in thymocytes, nor interfere with Notch-Delta dependent induction of *Ptcra* or *Deltex1*. However, when effects of PU.1 on its target genes are assessed in the

presence or absence of Notch-Delta signaling, a statistically strong interaction is seen, such that Notch/Delta signals block PU.1 effects on many genes.

Of 23 early T-cell genes affected by PU.1 in our analysis, as many as 11 of them showed supra-additive protection from PU.1 by the presence of Notch-Delta signals (*Myb*, *Hes1*, *Ikzf1*, *Gfi1*, *Cd3e*, *Rag1*, *Lat*, *Bcl11b*, *Zap70*, *Ets1*, and *Tcf12*), and one more was blocked from induction by PU.1 by Notch-Delta signals (*Id2*) (Fig. 1D; Supplementary Table 2A). The Notch-protected genes were particularly associated with the DN3 states, and the interaction was specific, because it did not apply to other genes regulated by PU.1 and Notch, such as genes associated with myeloid lineage redirection (10). In contrast, effects of high-level GATA-3 were much less influenced by Notch signaling. Of 22 early T-lineage genes affected by GATA-3, Notch-Delta signaling only modified three in a supra-additive way, diminishing the positive effects of GATA-3 on *Hes1* and *Mitf* and protecting *Tcf7* from repression (Supplementary Table 2B).

These results show that relative inputs from PU.1 activity and from Notch-Delta signaling can be key regulators of the progression of thymocytes from DN1 to DN3a.

Assembly of a framework for the T-cell specification gene regulatory network

We have used BioTapestry software (23) to make explicit the network of regulatory relationships that appear to operate through the emergence of committed early T-cell precursors from hematopoietic stem cells and other pluripotent progenitors. Such a network integrates all available data on regulatory inputs into each of the important genes in a process. It provides a validation map for assessing the extent to

which available information can account for the pattern of expression of individual genes and for the coordination of expression of groups of genes through the course of the process. Although yet incomplete, this network provides a useful armature for the regulatory relationships involved in T-cell specification.

To construct this model, we combined gene expression data for normal thymocyte subsets and perturbation data for PU.1, GATA-3, and Notch regulatory effects. Because PU.1 and GATA-3 can promote lineage redirection to myeloid/dendritic and mast-cell fates, respectively, regulatory pathways involved in these fate alternatives (24, 25) are also indicated. Also incorporated is evidence from the literature on the following: connections involving the basic helix-loop-helix E proteins and their antagonists (22, 26-29); some additional Notch inputs (30); possible links both upstream [TCF-1 (*Tcf7*), Runx (18, 19)] and downstream [Gfi1, Egr2; (1)] of PU.1; and additional data on the sharp changes in gene expression that follow commitment, during TCR-dependent selection in the DN3b and DN4 stages (5, 8, 31). A static view of this BioTapestry network is shown in Fig. 3. Full-sized screen shots and full annotation of the individual links are provided in SFig. 4 and Supplementary Table 3. In the interactive BioTapestry viewer posted on line (http://www.its.caltech.edu/~tcellgrn), we also provide dynamic views of changing network states through the DN1 to DN4 transitions, all the raw data from our group's publications on which the network links are based, and continuous updating of the annotations of data for individual links.

The results discussed here have required three modifications of the usual gene regulatory network depiction (23). First, the effects of Notch signaling on the activities

of other regulators have had to be taken into account. These go beyond independent cis-regulatory inputs of Notch and its transcription factor CSL and require "processing" of the effects of transcription factors such as PU.1 by effectors of Notch-Delta signaling. Second, GATA-3 demonstrates in particularly stark form the need to provide different network links for factors when expressed at different concentrations. For example, among the genes that GATA-3 shuts off when it is overexpressed is *Il7ra*, normally coexpressed with *Gata3*, while it induces some (*Tal1*, *Gata2*, *Gfi1b*) that would normally be turned off at stages when *Gata3* expression peaks.

Third, the data analysis reviewed above identifies three obvious gaps in using Notch, GATA-3, PU.1, or other known factors, to account for T-cell development. These have been filled by "placeholders". A specification inducer is needed to account for the initial upregulation of *Bcl11b*, *HEBalt*, and the *Cd3* genes at the DN2 stage. A DN3-specific gene activator seems necessary to supplement known effects of Notch signaling (cf. Fig. 2) and E protein activity, to account for the full pattern of DN3a-stage gene activation. Finally, a repressive T-lineage commitment function is needed to account for the timing of repression of *Sfpi1*, *Tal1*, *Gfi1b*, and other non-T-lineage promoting regulatory genes during the DN2 to DN3 transition. By placing these functions in a network context, equivalences should eventually emerge among parallel regulatory pathways involving these placeholders and known factors, and thus help to establish the actual molecular identities of these agents.

CONCLUDING REMARKS

T-cell specification is not governed by a few dedicated transcription factors operating in a simple regulatory pathway. It can only be understood in gene network terms. It depends on multiple transcription factors almost every one of which is also used, in other combinations, for other hematopoietic programs. The close linkage of T-cell development to other hematopoietic fates is hammered home by the ease of diverting T-cell precursors to other lineages, when the same transcription factors that are normally part of the T-cell program are overexpressed. This sharing is typical for hematopoietic lineage decisions, in which the same transcription factors expressed in different ratios or in different temporal orders yield different cell types.

However, this kind of system also makes gene network construction very difficult. The regulatory meaning of every transcription factor is completely dependent on level as well as on context. To explain how factors like GATA-3 can be used for T-cell development at all, it has been crucial to build into the network diagram dose-sensing nodes, which channel transcription factor input to different downstream genes at high factor levels than at low factor levels. Such nodes are not commonly needed in the embryonic specification gene networks developed to date, where transcription factors can act in a quasi-Boolean way due to strongly forward-driving network architecture. For T-cell development, a most important part of the context is provided by the sustained role of Notch pathway signaling throughout T-cell specification. Notch not only provides its own transcriptional input but also modifies the effects of both high-level GATA-3 and PU.1. Thus, to explain the use of PU.1, it has been crucial to include a

node for filtering this transcription factor's regulatory effects through transformations, still poorly understood, that depend on Notch-Delta signaling.

The network framework presented here necessarily remains incomplete. Only a limited number of perturbations have been tested so far in the temporally defined and stage-specific way that is needed to discern proximal downstream regulatory targets. Internal network structure depends on identifying the targets of other regulators. Also, as we have emphasized here, there are a number of stage-specific functions required for T-cell development that clearly remain to be identified, as they cannot be accounted for by effects of Notch signaling or GATA-3, and do not appear from other evidence to be explained by the expression patterns of other known factors. Two net positive T-lineage promoting functions are needed at different stages, and at least one lineage exclusion function is needed to explain commitment. Whether these are mediated by single factors or network subcircuits, cooperative or double negative, remains to be defined. The linkages that have been defined in the work described here reveal unexpected aspects of the roles of key factors, like GATA-3, and bring into focus the places additional connections need to be made.

MATERIALS AND METHODS

See full Materials and Methods in Supplementary Material Online. Regulatory and marker gene expression was measured by quantitative real-time PCR as described before (4, 6, 31). Expression levels were normalized to β-actin expression in the same

samples, $\log_{10}$ transformed, and then submitted to partial least squares and Analysis of Variance (ANOVA) analysis as described in Supplementary Material.

Acknowledgments

We thank Dr. Howard Petrie (Scripps Florida) for valuable discussions and sharing unpublished data for comparison with ours. Robert Butler, Ni Feng, and Koorosh J. Elihu provided excellent technical help. This work was supported by NIH grants R33 HL089102 (to H.B.) and R33 HL089123 (to E.V.R.), using data from support by CA90233, CA98925, DK73658, AI064590 (to M.A.Y.), and the DNA Sequencer Royalty Fund.

REFERENCES

1. Laslo, P., Spooner, C. J., Warmflash, A., Lancki, D. W., Lee, H. J., Sciammas, R., Gantner, B. N., Dinner, A. R. & Singh, H. (2006) Multilineage transcriptional priming and determination of alternate hematopoietic cell fates. *Cell* **126**, 755-766.

2. Iwasaki, H., Mizuno, S., Arinobu, Y., Ozawa, H., Mori, Y., Shigematsu, H., Takatsu, K., Tenen, D. G. & Akashi, K. (2006) The order of expression of transcription factors directs hierarchical specification of hematopoietic lineages. *Genes Dev* **20**, 3010-3021.

3. Rothenberg, E. V. (2007) Negotiation of the T lineage fate decision by transcription-factor interplay and microenvironmental signals. *Immunity* **26**, 690-702.

4. David-Fung, E. S., Yui, M. A., Morales, M., Wang, H., Taghon, T., Diamond, R. A. & Rothenberg, E. V. (2006) Progression of regulatory gene expression states in fetal and adult pro-T cell development. *Immunol Rev.* **209**, 212-236.

5. Rothenberg, E. V., Moore, J. E. & Yui, M. A. (2008) Launching the T-cell-lineage developmental programme. *Nat. Rev. Immunol* **8**, 9-21.

6. Tydell, C. C., David-Fung, E.-S., Moore, J. E., Rowen, L., Taghon, T. & Rothenberg, E. V. (2007) Molecular dissection of prethymic progenitor entry into the T lymphocyte developmental pathway. *J Immunol* **179**, 421-438.

7. Petrie, H. T. & Zuniga-Pflucker, J. C. (2007) Zoned out: functional mapping of stromal signaling microenvironments in the thymus. *Annu. Rev. Immunol* **25**, 649-679.

8. Taghon, T., Yui, M. A., Pant, R., Diamond, R. A. & Rothenberg, E. V. (2006) Developmental and molecular characterization of emerging β- and γδ-selected pre-T cells in the adult mouse thymus. *Immunity* **24**, 53-64.

9. Anderson, M. K., Hernandez-Hoyos, G., Dionne, C. J., Arias, A., Chen, D. & Rothenberg, E. V. (2002) Definition of regulatory network elements for T-cell development by perturbation analysis with PU.1 and GATA-3. *Devel. Biol.* **246**, 103-121.

10. Franco, C. B., Scripture-Adams, D. D., Proekt, I., Taghon, T., Weiss, A. H., Yui, M. A., Adams, S. L., Diamond, R. A. & Rothenberg, E. V. (2006) Notch/Delta signaling constrains re-engineering of pro-T cells by PU.1. *Proc Natl Acad Sci U S A* **103**, 11993-11998.

11. Taghon, T., Yui, M. A. & Rothenberg, E. V. (2007) Mast cell lineage diversion of T lineage precursors by the essential T-cell transcription factor GATA-3. *Nat. Immunol.* **8**, 845-855.

12. Huang, S., Guo, Y. P., May, G. & Enver, T. (2007) Bifurcation dynamics in lineage-commitment in bipotent progenitor cells. *Dev Biol* **305**, 695-713.

13. Cantor, A. B. & Orkin, S. H. (2002) Transcriptional regulation of erythropoiesis: an affair involving multiple partners. *Oncogene* **21**, 3368-3376.

14. Laiosa, C. V., Stadtfeld, M. & Graf, T. (2006) Determinants of lymphoid-myeloid lineage diversification. *Annu. Rev. Immunol* **24**, 705-738.

15. Anderson, M. K., Weiss, A., Hernandez-Hoyos, G., Dionne, C. J. & Rothenberg, E. V. (2002) Constitutive expression of PU.1 in fetal hematopoietic progenitors blocks T-cell development at the pro-T stage. *Immunity* **16**, 285-296.

16. Rothenberg, E. V. & Taghon, T. (2005) Molecular Genetics of T Cell Development. *Annu. Rev. Immunol* **23**, 601-649.

17. Hernández-Hoyos, G., Anderson, M. K., Wang, C., Rothenberg, E. V. & Alberola-Ila, J. (2003) GATA-3 expression is controlled by TCR signals and regulates CD4/CD8 differentiation. *Immunity.* **19**, 83-94.

18. Huang, G., Zhang, P., Hirai, H., Elf, S., Yan, X., Chen, Z., Koschmieder, S., Okuno, Y., Dayaram, T., Growney, J. D. *et al.* (2008) PU.1 is a major downstream target of AML1 (RUNX1) in adult mouse hematopoiesis. *Nat Genet.* **40**, 51-60.

19. Rosenbauer, F., Owens, B. M., Yu, L., Tumang, J. R., Steidl, U., Kutok, J. L., Clayton, L. K., Wagner, K., Scheller, M., Iwasaki, H. *et al.* (2006) Lymphoid cell

growth and transformation are suppressed by a key regulatory element of the gene encoding PU.1. *Nat Genet.* **38**, 27-37.

20. Schmitt, T. M. & Zúñiga-Pflücker, J. C. (2002) Induction of T cell development from hematopoietic progenitor cells by Delta-like-1 in vitro. *Immunity.* **17**, 749-756.

21. Taghon, T. N., David, E.-S., Zúñiga-Pflücker, J. C. & Rothenberg, E. V. (2005) Delayed, asynchronous, and reversible T-lineage specification induced by Notch/Delta signaling. *Genes Dev.* **19**, 965-978.

22. Ikawa, T., Kawamoto, H., Goldrath, A. W. & Murre, C. (2006) E proteins and Notch signaling cooperate to promote T cell lineage specification and commitment. *J Exp. Med.* **203**, 1329-1342.

23. Longabaugh, W. J. R., Davidson, E. H. & Bolouri, H. (2005) Computational representation of developmental genetic regulatory networks. *Dev. Biol.* **283**, 1-16.

24. Swiers, G., Patient, R. & Loose, M. (2006) Genetic regulatory networks programming hematopoietic stem cells and erythroid lineage specification. *Dev Biol* **294**, 525-540.

25. Friedman, A. D. (2002) Transcriptional regulation of granulocyte and monocyte development. *Oncogene* **21**, 3377-3390.

26. Schwartz, R., Engel, I., Fallahi-Sichani, M., Petrie, H. T. & Murre, C. (2006) Gene expression patterns define novel roles for E47 in cell cycle progression, cytokine-mediated signaling, and T lineage development. *Proc Natl Acad Sci U S A* **103**, 9976-9981.

27. Wang, D., Claus, C. L., Vaccarelli, G., Braunstein, M., Schmitt, T. M., Zuniga-Pflucker, J. C., Rothenberg, E. V. & Anderson, M. K. (2006) The basic helix-loop-helix transcription factor HEBAlt is expressed in pro-T cells and enhances the generation of T cell precursors. *J Immunol* **177**, 109-119.

28. Xu, W. & Kee, B. L. (2007) Growth factor independent 1B (Gfi1b) is an E2A target gene that modulates Gata3 in T-cell lymphomas. *Blood* **109**, 4406-4414.

29. Glimcher, L. H., Townsend, M. J., Sullivan, B. M. & Lord, G. M. (2004) Recent developments in the transcriptional regulation of cytolytic effector cells. *Nat. Rev. Immunol* **4**, 900-911.

30. Maillard, I., Tu, L., Sambandam, A., Yashiro-Ohtani, Y., Millholland, J., Keeshan, K., Shestova, O., Xu, L., Bhandoola, A. & Pear, W. S. (2006) The requirement for Notch signaling at the β-selection checkpoint in vivo is absolute and independent of the pre-T cell receptor. *J Exp. Med.* **203**, 2239-2245.

31. Yui, M. A. & Rothenberg, E. V. (2004) Deranged early T cell development in immunodeficient strains of nonobese diabetic mice. *J Immunol* **173**, 5381-5391.

FIGURE LEGENDS

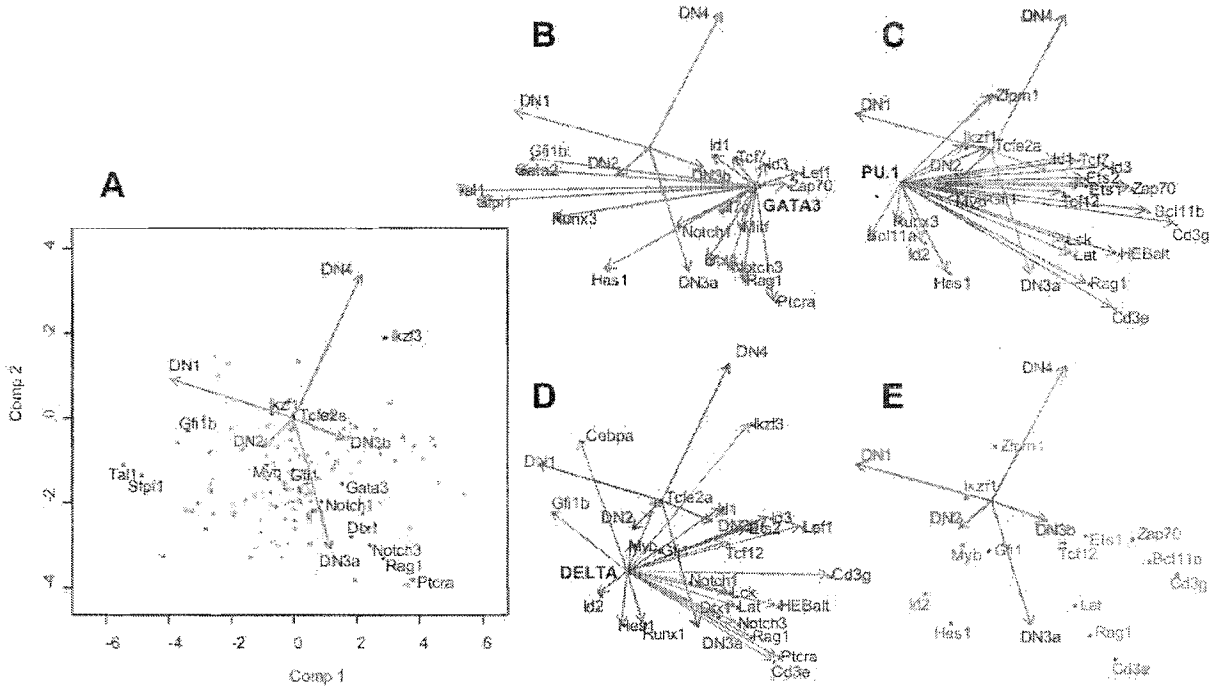
Figure 1. Gene expression changes during normal T-lineage specification and regulatory perturbation: depiction in principal component space. (A) Coordinates of key genes and gene expression signatures of DN1, 2, 3a, 3b, and 4 stages, projected on axes representing the first two principal components of gene expression change. Each stage can be depicted as a unit vector in 5 dimensions of principal component space (full data in SFig. 1, Supp. Table 1), with different principal components dominated by changes in expression of different genes. Genes (black or magenta stars) with stable expression are at the center of the graph. Genes with the highest change in expression are furthest from the center. The more similar the patterns of expression of two genes, the smaller the angle between them from the center. A DN stage vector appears longest against a principal component axis that captures most of its own difference from the average of all DN subsets. (B) Targets of GATA-3 overexpression in fetal thymocytes, positioned relative to first two principal component axes shown in (A) (11). The coordinates of the normal, adult DN1-DN4 stage phenotypes are shown for orientation by blue vectors. Genes positively affected by GATA-3 overexpression are targets of green arrows, with negatively affected genes indicated by red arrows. (C) Targets of PU.1 overexpression in fetal thymocytes (10), depicted as in (B). (D) Effects of short-term exposure to Notch-Delta signaling in fetal thymocytes. These effects, compiled from refs. (10, 11), are calculated independently of effects of PU.1 or GATA-3 in the experiments. (E) Supra-

additive modulation of PU.1 effects by Notch-Delta signaling (10). Genes that are protected from repression supra-additively by Notch-Delta signaling are shown in green. A gene that is prevented from upregulation by Notch-Delta signaling is shown in red. Threshold for interaction: $p<0.05$ (Supp. Table 2).

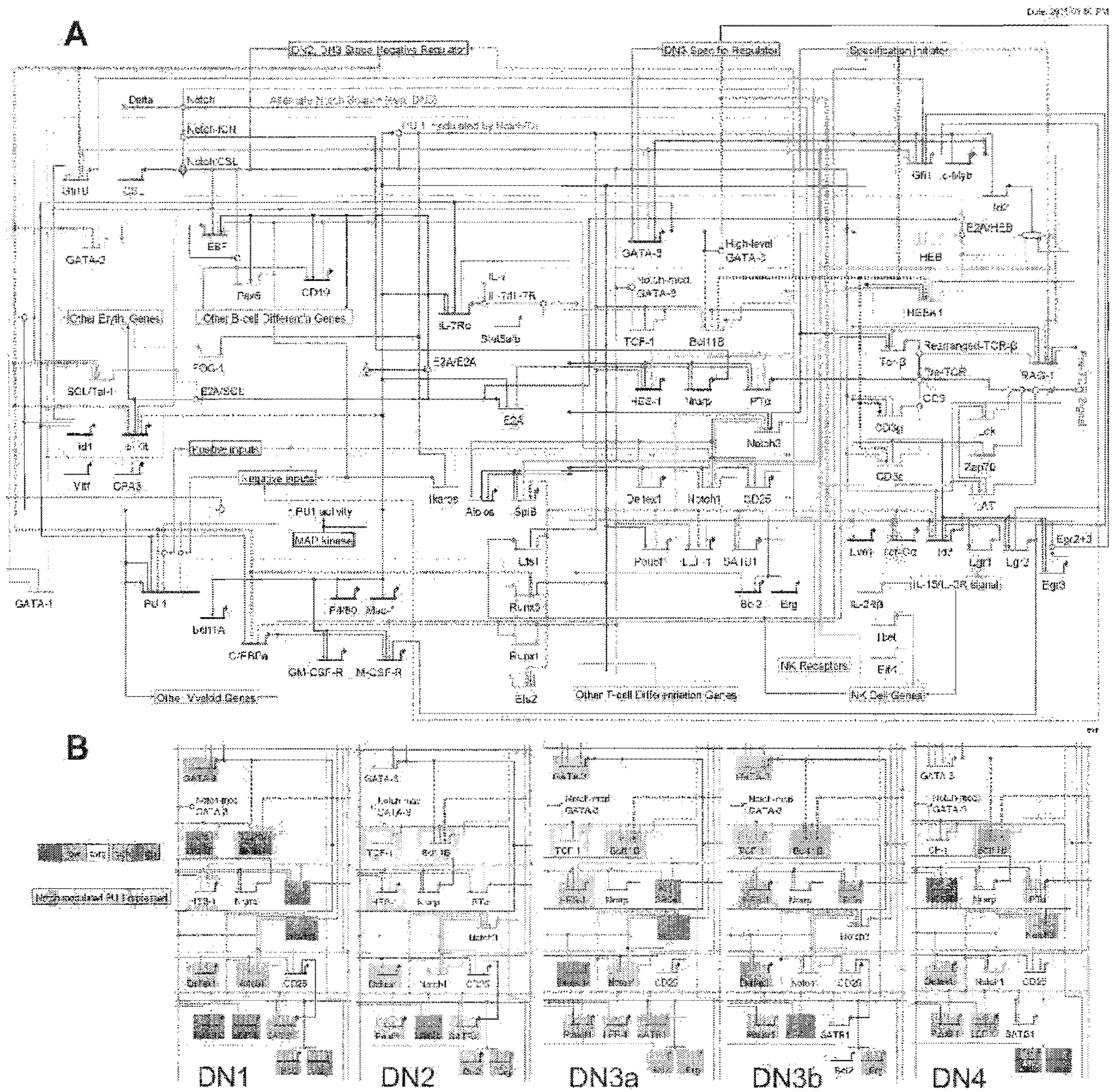
Figure 2. Direct Notch regulation of early T-cell genes compared with developmental regulation during the DN1 to DN3 transitions. Line graphs show the $\log_{10}$ of the ratio of expression of indicated genes in adult DN3a cells relative to DN1 cells in two independent studies (4, 5)(E.-S.D.-F. and M.A.Y., unpublished results); breaks indicate genes not included. Bars: effect of 24-hr Notch-Delta signaling on gene expression in Thy-1$^+$ E15.5 fetal thymocytes following a 16 hr preculture without Delta. Secondary y axis gives $\log_{10}$ of the expression ratio with and without restored Notch-Delta signaling. For additional results see SFig. 3. Data are from empty vector controls in refs.(10, 11)

Figure 3: Gene regulatory network model for T-cell specification. (A) "View from All Nuclei": comprehensive map of relationships included in the network, integrating over all stages. For sources of each link, see Supplementary Table 3. For expanded size figure, see SFig. 4A. For predicted differential activity of different network links at different stages, see SFig. 4G-K. (B) Close-ups of one region of the network with background highlighting indicating differential gene expression levels at five different developmental states. For full network versions, see SFig. 4B-F.

	Initiation Stem to DN1	**DN2 entry** DN1 to DN2	**Specification** DN1-DN3 continuous		**Commitment** DN3 stage		**Selection** β-selection	
	INCREASE	INCREASE	INCREASE	DECREASE	INCREASE	DECREASE	INCREASE	DECREASE
Transcription factors	GATA-3 TCF-1 HES1	Bcl11b HEB-alt	GATA-3 TCF-1 (Tcf7) HEB (Tcf12) Ikaros Gfi1 Runx1 Ets1	PU.1 SCL GATA-2 Id2 Runx2 Runx3 Gfi1B C/EBPα (& others)	Ets2 Pou6f1 SpiB HES1 LEF-1 Id3	PU.1 SCL	Aiolos SATB1 Tbet	SpiB HES1 Myb Runx1 HEB-alt Erg Id2 (& others)
Targets/other	Deltex1	CD3γ CD3ε	IL-7Rα CD25 Rag-1 CD3ε Lck LAT ZAP70	Mac-1	Deltex1 pTα (Ptcra) Rag-1 Notch3 Lck LAT ZAP70	c-Kit	TCR-Cα ZAP70	Deltex1 IL-7Rα pTα (Ptcra) Rag-1 CD25 Bcl2 Notch1 Notch3


A
B
C
D
E
DN4
DN1
DN2
DN3a
DN3b
Ikzf3
Ikzf1
Tcfe2a
Gfi1b
Tal1
Sfpi1
Myb
Gata3
Notch1
Dtx1
Notch3
Rag1
Ptcra
Comp 1
Comp 2
GATA3
Runx3
Hes1
Id1
Id2
Id3
Lef1
Zap70
Cebpa
DELTA
Runx1
Cd3g
Lck
Lat
HEBalt
Zbtb1
PU.1
Ets1
Tcf12
Bcl11b
Bcl11a
Cd3e

Log exp ratio, Dl.cell
Log expression ratio, DN2:DN1
Pbx3
CD36
CD3g
Bcl11B
HEBAlt
Rag1
Dsg1cy1
E1s2
E1s1
Zap70
Id3
LAT
Id2
Lck
Tcf7
Tcf12
Gata3
Runx1
Hoxa1
Cd1
Myb
Zfpm1
E2A
Il2rb
Runx3
Id2
Bcl11A
Cebpb
Cebpa
Sfpi1


A
B
DN2, DN3 Stage Negative Regulator
DN3 Spec fic Regulator
Specification Initiator
Delta
Notch
Notch-ICN
Notch/CSL
GATA-2
EBF
Pax5
CD19
Other Eryth. Genes
Other B-cell Differentn Genes
IL-7
IL-7Rα
Stat5a/b
E2A/E2A
FOG-1
E2A/SCL
SCL/Tal-1
E2A
Id1
Positive inputs
Negative inputs
PU1 activity
MAP kinase
Ikaros
Aiolos
SpiB
GATA-1
PU.1
Bcl11A
C/EBPα
GM-CSF-R
M-CSF-R
Other Myeloid Genes
Ets1
Runx3
Ets2
Other T-cell Differentiation Genes
GATA-3
High-level GATA-3
Notch-mod. GATA-3
TCF-1
Bcl11B
HES-1
Nrarp
PTα
Notch3
Deltex1
Notch1
CD25
Pou6f1
SATB1
Bcl2
Erg
Gfi1
c-Myb
Id2
E2A/HEB
HEB
HEBAlt
Tcr-β
Rearranged-TCR-β
Pre-TCR
RAG-1
CD3
CD3γ
CD3ε
Lck
Zap70
LAT
Lmo1
Lgr1
Egr2+3
Egr3
Tbet
NK Receptors
NK Cell Genes
DN1
DN2
DN3a
DN3b
DN4