nature immunology

Intrathymic programming of effector fates in three molecularly distinct γδ T cell subtypes

Kavitha Narayan[1,5], Katelyn E Sylvia[1,5], Nidhi Malhotra[1], Catherine C Yin[1], Gregory Martens[2], Therese Vallerskog[2], Hardy Kornfeld[2], Na Xiong[3], Nadia R Cohen[4], Michael B Brenner[4], Leslie J Berg[1], Joonsoo Kang[1] & The Immunological Genome Project Consortium[6]

Innate γδ T cells function in the early phase of immune responses. Although innate γδ T cells have often been studied as one homogenous population, they can be functionally classified into effector subsets on the basis of the production of signature cytokines, analogous to adaptive helper T cell subsets. However, unlike the function of adaptive T cells, γδ effector T cell function correlates with genomically encoded T cell antigen receptor (TCR) chains, which suggests that clonal TCR selection is not the main determinant of the differentiation of γδ effector cells. A high-resolution transcriptome analysis of all emergent γδ thymocyte subsets segregated on the basis of use of the TCR γ-chain or δ-chain indicated the existence of three separate subtypes of γδ effector cells in the thymus. The immature γδ subsets were distinguished by unique transcription-factor modules that program effector function.

The Immunological Genome (ImmGen) Project is a consortium of immunologists and computational biologists assembled to comprehensively define gene regulatory networks in the immune system of the mouse[1]. In that context, we determined the global gene-expression profiles of subsets of intrathymic cells expressing the γδ T cell antigen receptor (γδTCR) to ascertain the heterogeneity of the T cell lineage and to identify gene networks that control the production of innate effector-cell subsets distinct from those that operate in conventional adaptive T cells. T cells in vertebrates are separated into two lineages on the basis of the expression of either γδTCRs or αβTCRs on their cell surface. The adaptive αβ T cell lineage is subdivided into T helper type 1 cells, T helper type 2 cells, interleukin 17 (IL-17)-producing helper T cells, follicular helper T cells, regulatory T cells and cytotoxic effector cells that are often considered distinct cell lineages. There are indications that the innate γδ T cell lineage is also composed of distinct subsets programmed to secrete a discrete cluster of effector cytokines[2,3], but the mechanism of innate effector specialization is unclear. As sentinels of the immune system, γδ T cells are located mainly in the mucosal epithelia where pathogens are first encountered[4], and thus the rapidity of their response to infection is paramount. Similar to other innate lymphocytes[5,6], γδ T cells are exported from the thymus as 'pre-made' memory-like cells, displaying cell-surface markers associated with cellular activation. After infection, γδ T cells rapidly produce effector cytokines and growth factors, similar to memory αβ T cells[7–9].

Although the genes that encode the γδTCR and αβTCR were identified contemporaneously, studies delving into the distinct development and function of γδ T cells have been greatly hampered by the scarcity of known molecules that distinguish them, other than the TCR and the transcription factor SOX13 (ref. 10), that can be classified as true γδ T cell lineage–specific markers. Identification of the genetic circuits that underpin two unique features of γδ T cell development is particularly critical. First, the segregation of effector functions of γδ T cells from adult mice is based on the use of genes encoding the γ-chain variable region (V_γ) and/or δ-chain variable region (V_δ)[11]; the production of IL-17, interferon-γ (IFN-γ) or IL-4 is associated with the $V_\gamma2^+$ (V2), $V_\gamma1.1^+V_\delta6.3^-$ (V1) or $V_\gamma1.1^+V_\delta6.3^+$ (V6) γδ T cell subset, respectively. In comparison, conventional αβ T cells are classified into functional subsets on the basis of the repertoire of effector cytokines produced, not by TCR repertoire, which is vastly diverse for each subset. Second, the function of γδ T cell subsets seems to be programmed in the thymus[12–14]. In contrast, conventional αβ T cells differentiate into effector subsets after encountering pathogens in peripheral tissues. How and when γδ subsets are programmed and/or selected toward distinct effector-cell fates in the thymus is not well understood. So far, γδ T cell development has been studied mainly by analysis of total γδTCR$^+$ cells as one uniform population, not as distinct subsets defined by use of V_γ or V_δ chains. We compared the gene-expression profiles of emergent γδTCR$^+$ thymocytes from adult mice segregated on the basis of the use of genes encoding V_γ and/or V_δ

[1]Department of Pathology, University of Massachusetts Medical School, Worcester, Massachusetts, USA. [2]Department of Medicine, University of Massachusetts Medical School, Worcester, Massachusetts, USA. [3]Department of Veterinary and Biomedical Sciences, The Pennsylvania State University, University Park, Pennsylvania, USA. [4]Brigham and Women's Hospital and Department of Medicine, Division of Rheumatology, Immunology and Allergy, Harvard Medical School, Boston, Massachusetts, USA. [5]These authors contributed equally to this work. [6]Complete list of authors and affiliations appears at the end of this paper. Correspondence should be addressed to J.K. (joonsoo.kang@umassmed.edu).

Received 10 November 2011; accepted 17 January 2012; published online 1 April 2012; doi:10.1038/ni.2247

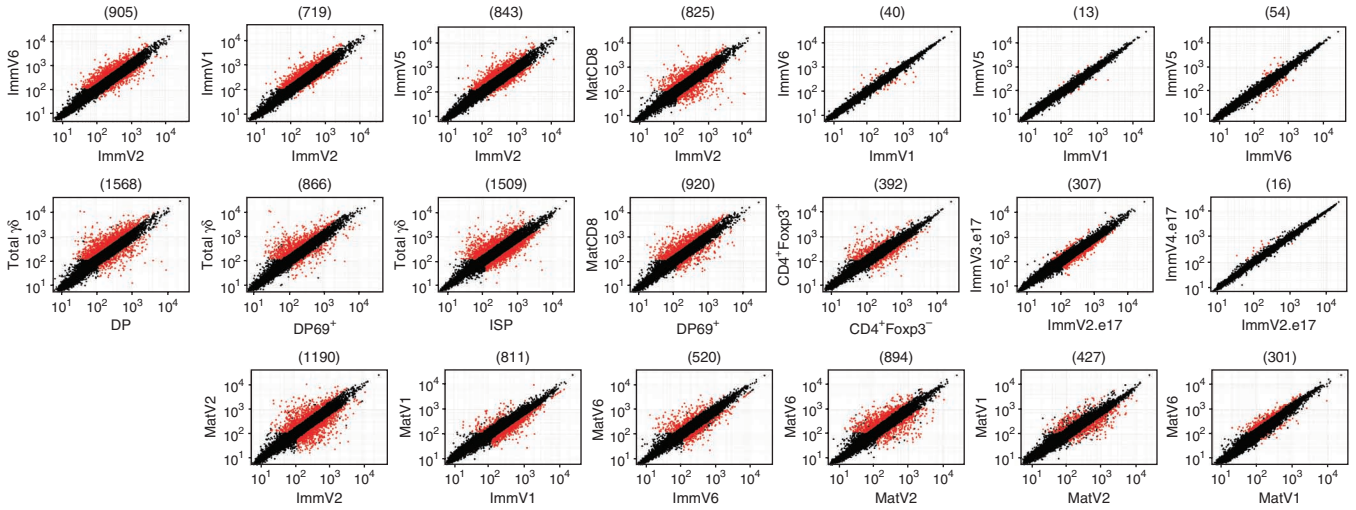

Figure 1 Distinct global gene-expression profiles of the γδ cell subsets defined by TCR repertoire. Expression of consolidated probe sets by populations of immature thymocytes from adult or fetal mice and mature thymocytes from C57BL/6 mice. Each dot represents one gene (mean of all probe sets); red indicates genes with expression changed by more than twofold ($P < 0.05$ (Student's *t*-test); coefficient of variation < 0.5; mean expression value (MEV) > 120 in one subset); numbers in parentheses above plots indicate total number of these genes. $CD4^+Foxp3^-$ and $CD4^+Foxp3^+$ samples were isolated from the spleen. Similar results were obtained for the comparison of $CD8^+CD24^{int}$ thymocytes and immV2 cells (615 genes). Designations along axes correlate with ImmGen populations as follows: ImmV2, immTgd.vg2+.Th; ImmV1, immTgd.vg1+vd6-.Th; ImmV6, immTgd.vg1+vd6+.Th; ImmV5, immTgd.vg5+.Th (sorted in duplicate); MatV2, matTgd.vg2+.Th (sorted in duplicate); MatV1, matTgd.vg1+vd6-.Th (sorted in duplicate); MatV6, matTgd.vg1+vd6+.Th; Semi-matCD8, T.8SP24int.Th; DP, T.DP.Th (CD69-); DP CD69+, T.DP69+.Th; Total γδ, Tgd.Th; ISP (immature single-positive), T.ISP.Th; MatCD8, T.8SP24-.Th; $CD4^+FoxP3^-$, T.4FP3-.Sp $CD4^+FoxP3^+$, T.4FP3+25+.Sp; ImmV2.e17, immTgd.vg2.e17.Th; ImmV3.e17, immTgd.vg3.e17.Th; ImmV4.e17, immTgd. vg4.e17.Th; ETP, preT.ETP.Th; DN2, preT.DN2.Th; DN3A, preT.DN3A.Th; MatCD4, T.4SP24-.Th; DN4, T.DN4.Th; *i*NKT, NKT.44+NK1.1+.Th (unless otherwise specified). Data are from two to three independent experiments with 4–30 mice each.

and found that the main subsets were as distinct from each other as they were from αβTCR$^+$ thymocyte subsets. Most notably, the profiles of the emergent immature γδ T cell subsets were already embedded with unique gene programs that direct subset-specific effector function, which indicates that γδ function is molecularly programmed in the thymus before or concurrent with TCR expression.

RESULTS

Distinct subtypes among emerging γδTCR$^+$ thymocytes

A correlation between effector function and use of the TCR V_γ and/or V_δ chain[11] raised the possibility that the earliest identifiable γδ T cell populations in the thymus are composed of molecularly heterogeneous cell subtypes distinguishable by the expression of unique germline-encoded γδTCRs. To address this possibility, we determined the gene-expression signatures of four subsets of γδ T cells from adult mice on the basis of expression of TCRγ and/or TCRδ and maturational stages on the basis of expression of the cell surface marker CD24 (heat-stable antigen). We compared expression of the V2 ($V_\gamma2^+$; partner δ-chain repertoire and frequency among total γδTCR$^+$ thymocytes, Online Methods), V1 ($V_\gamma1.1^+$), V6 ($V_\gamma1.1^+V_\delta6.3^+$) and V5 ($V_\gamma5^+$; intraepithelial lymphocyte) thymocyte subsets with that of other thymic subsets of the T cell lineage profiled by ImmGen. The paucity of mature ($CD24^{lo}$) V5 thymocytes prohibited their analysis. In addition, we established gene-expression profiles of the following three subsets of immature ($CD24^{hi}$) γδ T cells from fetal mice: V3 ($V_\gamma3^+$; dendritic epidermal T cells), V4 ($V_\gamma4^+$) and V2 ($V_\gamma2^+$).

Pairwise comparisons showed that the earliest identifiable immature V2 (immV2) subset of cells from adult mice, the largest population in the γδ T cell lineage, was distinct in its global gene-expression profile compared with that of all other γδ T cell subsets (**Fig. 1**). At an arbitrary threshold of a twofold change in expression, the number of genes that distinguished the immV2 subset from other immature γδ subsets was roughly equivalent to the difference between the immV2 subset and CD8$^+$ thymocytes of the αβ lineage (**Fig. 1**). Putting this finding in the proper context, genes with different expression in immature $CD4^+CD8^+$ double-positive (DP) $CD69^-$ cells of the αβ lineage and DP $CD69^+$ cells that had received TCR signals versus total γδTCR$^+$ thymocytes numbered ~1,600 and ~900, respectively. In the αβ T cell lineage, $Foxp3^+$ regulatory T cells versus conventional $CD4^+$ T cells in the spleen differed in the expression of ~400 genes (**Fig. 1**). The distinct gene-expression signature of immV2 cells was in contrast to the similarity among immature subsets of V1, V5 and V6 cells, which differed in the expression of at most 54 genes. However, the higher expression of the natural killer T cell (NKT cell) lineage marker PLZF (encoded by *Zbtb16*) and granzyme A, which has shown enrichment in intraepithelial lymphocytes[15], in immature V6 (immV6) and immature V5 (immV5) cells, respectively, suggested that despite their similarity, they were already marked with the molecular features of their fully differentiated peripheral counterparts (**Fig. 2a** and data not shown). We confirmed the amount of selected transcription factors whose genes had different expression in subsets of γδ T cells from adult mice (**Fig. 2a**) and obtained confirmatory data for additional genes with different expression (**Supplementary Fig. 1**), which captured the considerable heterogeneity in the phenotype of γδ subtypes. The immV2 and immature V4 (immV4) cells from fetal mice (the subset of cells from fetal mice programmed for IL-17 production[11,16]) had very similar global gene-expression profiles, which correlated with their common IL-17-producing effector phenotype. In contrast, immature V3 (immV3) cells from fetal mice, the precursors of dendritic epidermal T cells, were distinct from other γδ T cell subsets from fetal mice (**Fig. 1**), consistent with their known distinct developmental origin and requirements[17,18]. Thus, the γδ T cell lineage can be divided into the following three distinct subtypes on the basis of unique gene-expression profiles at the immature developmental stage: V3, V4-V2 and V1-V6-V5.

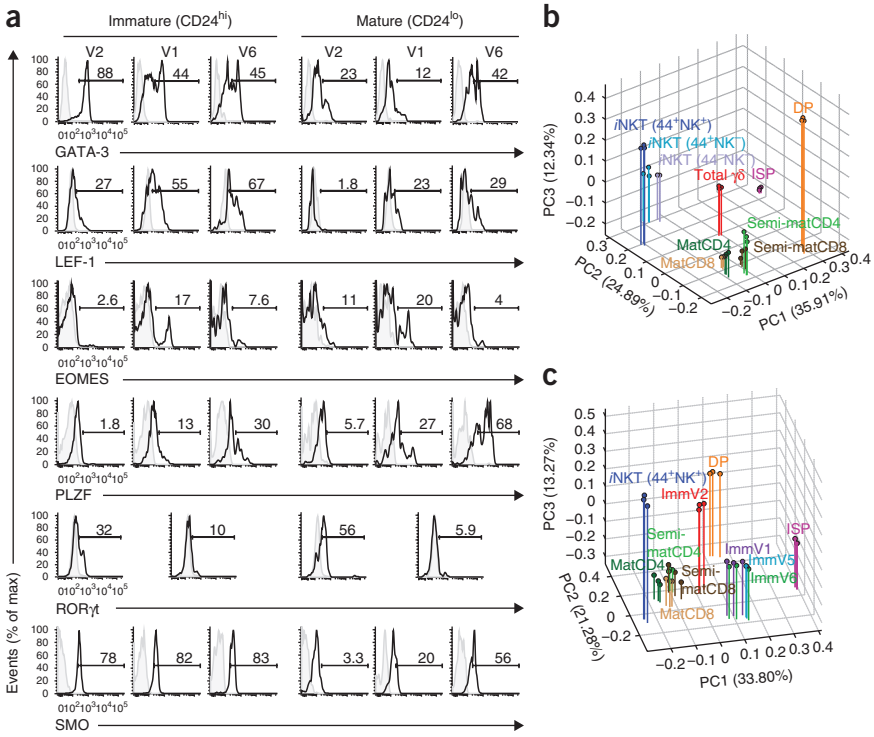

Figure 2 Distinct expression of transcription factors and divergence of γδ T cell subsets. (**a**) Expression of transcription factors in immature ($CD24^{hi}$) and mature ($CD24^{lo}$) V2, V1 and V6 γδ thymocytes, generated by gating on total $TCR\delta^+$ cells, with cell subsets defined on the basis of expression of $V_\gamma2$, $V_\gamma1.1$, and $V_\delta6.3$, and gating on $CD24^{hi}$ or $CD24^{lo}$ cells in each subset. Plots 'straddling' the V1 and V6 columns are gated on total $V_\gamma1.1^+$ cells. Gray shaded histograms, staining with isotype-matched control antibody. For GATA-3, a flow-cytometry-minus-one control was used; for RORγt, a negative control was used ($TCR\beta^{hi}$ cells negative for RORγt expression). SMO, transcription factor with similar expression among immature subsets that is downregulated after maturation. In some cases, gates were drawn on cells with high expression to best present the relative difference in expression among the γδ subsets. Numbers above bracketed lines indicate percent cells expressing the marker. (**b**) PCA of the 15% most variable genes in various populations of cells (colors of bars and labels indicate population; MEV > 120 in at least one population; 1,594 genes). The first three principal components (PC1–PC3) are presented here, along with the proportion of the total variability represented by each component (in parentheses along axes). Expression of CD44 (44) and NK1.1 (NK) by the *i*NKT cell subsets is included in parentheses. Semi-matCD4, semi-mature $CD4^+$ cells; Semi-matCD8, semi-mature $CD8^+$ cells. (**c**) PCA of the 15% most variable genes in various populations of cells as in **b** (MEV > 120 in at least one population; 1,597 genes). Data are representative of at least three independent experiments with at least three mice per experiment (**a**) or are from two to three independent experiments with 4–30 mice each (**b**,**c**).

Lineage relatedness among diverse subsets can be visualized by principal-component analysis (PCA), a mathematical transformation that decreases the dimensionality of gene-expression data to demonstrate the main components of variability among the samples. We focused on subsets of γδ cells from adult mice, as this data set is complete. We first used PCA to compare the gene-expression profiles of nonsegregated total γδTCR$^+$ thymocytes with those of three αβ thymocyte subsets ranging in maturity and lineages using the 15% of genes with the greatest difference in expression among the samples, on the basis of expression and variance across the populations (**Fig. 2b**). Mature αβ thymocyte subsets were positioned together, increasingly displaced from precursor subsets (immature single-positive $CD8^+$ and DP) as a function of their relative maturity. Innate-like αβ invariant NKT cells (*i*NKT cells) at three stages of maturation clustered separately from other αβ cell subsets, and total γδ T cells were positioned away from the four distinct clusters of αβ thymocyte subsets (**Fig. 2b**). We obtained similar results when we measured subset relatedness by Euclidean distance and Pearson's correlation coefficients for the 15% of genes with the greatest difference in regulation used for PCA (**Supplementary Fig. 2a**,**b**). When we used the segregated γδ T cell subsets for this analysis, however, it was clear that the immV2 subset was distinct from all other γδ thymocyte subsets, positioned uniquely in relation

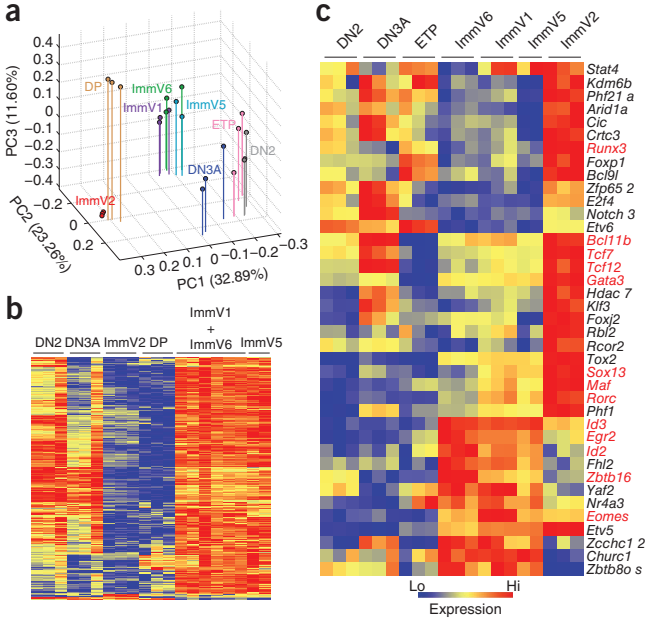

Figure 3 Expression of transcription factors and genes encoding molecules involved in metabolic processes distinguishes γδ thymocyte subsets. (**a**) PCA (as in **Fig. 2b**) of genes with different regulation among immV2 and immV1 or immV6 cells (1,006 genes; **Supplementary Fig. 3a** and **Supplementary Tables 1** and **2**). (**b**,**c**) Heat maps of the expression of genes encoding molecules involved in metabolic processes (**b**) or transcription factors (**c**) in thymocyte subsets, showing genes with differences in regulation among immV2 and immV1 or immV6 cells; data were log transformed, centered by gene row and hierarchically clustered by gene and subset. Red font indicates transcription factors discussed in text. In **b**, interspersed immV1 and immV6 replicates are grouped together. DN2, double-negative ($CD4^-CD8^-$) stage 2; DN3A, double-negative stage 3A. Data are from two to three independent experiments with 4–30 mice each.

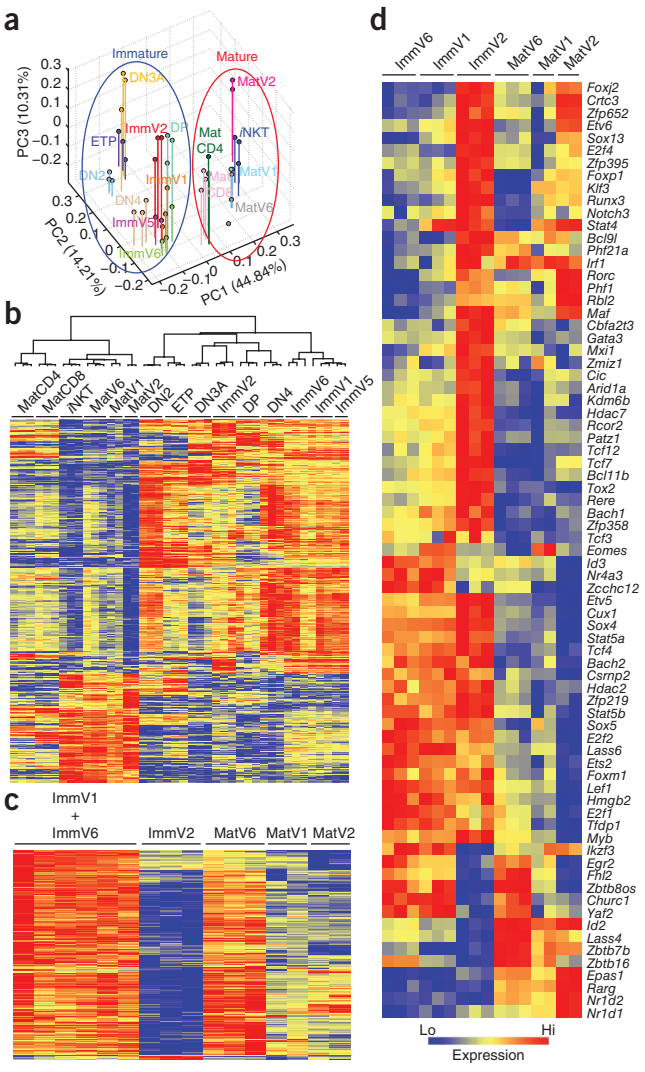

Figure 4 Convergence of the gene-expression profiles of γδ subsets after maturation. (**a**) PCA (as in **Fig. 2b**) of the 495 genes with different regulation after the maturation of γδ T cells from adult mice (**Supplementary Fig. 5** and **Supplementary Table 7**). Blue oval, immature cluster; red oval, mature cluster. (**b**) Heat map of the expression of the 495 genes of the γδ maturation gene signature in precursor, αβ, and γδ T cell subsets. Dendrogram (above) shows the relatedness of the populations based on hierarchical clustering. (**c**,**d**) Heat maps of the expression of genes encoding molecules involved in metabolic processes (**b**) or transcription factors (**c**) in subsets of immature and mature γδ cells (presented as in **Fig. 3b,c**). In **c**, interspersed immV1 and immV6 replicates are grouped together. Data are from two to three independent experiments with 4–30 mice each.

to other γδ and αβ subsets (**Fig. 2c** and **Supplementary Fig. 2c**,**d**). Second, for genes with expression in immV2 subsets different from that in other immature γδ cell subsets (1,006 genes; **Supplementary Fig. 3a** and **Supplementary Tables 1** and **2**), the results obtained by PCA, Euclidian distances and correlation scores supported the idea of close similarity between immV2 and αβ DP thymocytes, which shared >50% of the variance and segregated distantly from T precursor subsets (**Fig. 3a** and **Supplementary Fig. 3b**,**c**). When DP populations were segregated, PCA indicated that immV2 cells were most similar to DP CD69$^+$ cells (**Supplementary Fig. 3d**). These results showed that immV2 cells distinguished by germline-encoded TCR domains were different from other immature γδ T cells and, conversely, that γδ T cells expressing the $V_\gamma2$ chain from the locus encoding the γ-chain constant region 1 ($C_\gamma1$), despite the extensive TCR clonal diversity generated from their TCRδ partners and variable-joining (VJ)–variable-diversity-joining (VDJ) junctional diversity, nevertheless represented a unique cell type.

Distinct regulatory circuits in immature subtypes

About one third of the genes with lower expression immV2 cells than in V1 or V6 cells encoded proteins involved in DNA synthesis, RNA biogenesis, oxidative phosphorylation and protein translation (~270 genes), which suggested lower metabolism and energy production in immV2 cells (**Supplementary Tables 1**, **3** and **4**) than in other γδ subsets. The expression of genes encoding metabolic molecules in immV2 cells corresponded to the pattern in immature DP cells of the αβ lineage, which provided the main element of similarity between these two populations (**Fig. 3b**, **Supplementary Table 4** and data not shown). The distinct gene signatures of γδ subsets were not a consequence of different cell-cycle properties or susceptibility to death, as the immature subsets incorporated the thymidine analog BrdU similarly and had similar frequencies of cells expressing Ki67, a marker of the non-G0 state of the cell cycle (**Supplementary Fig. 4a**,**b**), and annexin V, a marker of apoptotic cells (data not shown). These results suggested that immV2 cells develop distinctly with unique maturational transitions.

Lineage-specific transcription factors propel the lineage-commitment process by turning on lineage-associated genes and turning off lineage-mismatched genes. Distinct programming of the dominant transcription factors in subset-specific effector function was already evident in emerging γδ T cell subsets. Hierarchical clustering analysis of transcription factors with different expression in immV2 cells versus immV1, immV5 and immV6 cells indicated that the immV2 subset had much higher expression of most transcription factors, the most prominent of those belonging to the SOX-TCF1-TOX high-mobility-group box family of transcription factors (**Fig. 3c** and **Supplementary Tables 2**, **5** and **6**). Among cells of the immune system, *Sox13* expression is restricted to developing γδ T cells. Its product, SOX13, interacts with TCF1 (encoded by *Tcf7*) and LEF1 (encoded by *Lef1*), the nuclear effectors of Wnt signaling, to establish, in part, the pan–γδ lineage gene expression signature[10]. Conversely, expression of *Id2* and *Id3*, which encode the negative regulators of the E helix-loop-helix transcription factors E47 (encoded by *Tcfe2a*) and HEB (encoded by *Tcf12*), along with the gene encoding their upstream regulator (*Egr2*), was higher in immV1, immV5 and immV6 cells than in immV2 cells[19,20]. Furthermore, expression of *Bcl11b*, which encodes a negative regulator of NK cell development[21], was lower in immV1, immV5 and immV6 cells than in immV2 cells (**Fig. 3c** and **Supplementary Table 6**).

In the periphery, V2 CCR6$^+$CD27$^-$ γδ T cells are biased to produce IL-17, whereas V1 CD27$^+$ T cells secrete IFN-γ[8,14]. We further verified this effector pattern during the γδ T cell response to infection of mice with *Mycobacterium tuberculosis*[7]. At 14 d after infection, all γδ subsets in the lung synthesized IFN-γ, but V2 and V4 cells were the predominant producers of IL-17 (**Supplementary Fig. 4c–f**). The transcription factors RORγt (encoded by *Rorc*), Eomes (encoded by *Eomes*) and PLZF (encoded by *Zbtb16*) regulate the expression of IL-17, of IFN-γ and of both IL-4 and IFN-γ, respectively, in αβ T cells[22–25]. Emergent immature V2, V1 and/or V5, and V6 thymic subsets were distinguished by selective expression of these three transcription factor genes, which correlated with their peripheral γδ T cell subset–specific functions (**Figs. 2a** and **3c**). As further support

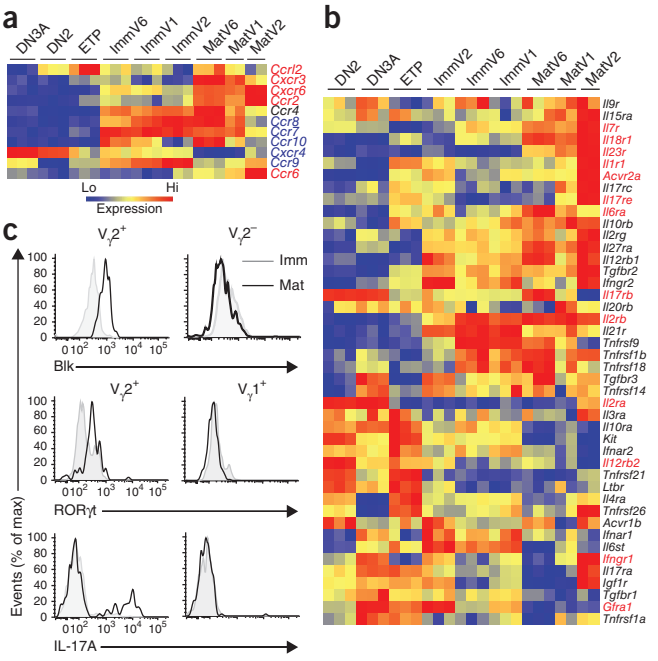

Figure 5 Generation of mature γδ cell subsets poised for the elaboration of effector function programs. (**a**) Heat map of the expression of genes encoding chemokine receptors, expressed in immature subsets and extinguished or downregulated as γδ T cells mature (blue) or induced precipitously during maturation and maintained in fully differentiated subsets (red). (**b**) Heat map of the expression of cytokine and growth factor receptors (presented as in **Fig. 3b,c**). (**c**) Expression of Blk, RORγt and IL-17A in immature (imm) and mature (mat) γδ thymocyte subsets, gated on total TCRδ$^+$ cells, separated into V$_\gamma$2$^+$ and V$_\gamma$2$^-$ populations, and gated on CD24^{hi} (light gray lines) or CD24^{lo} (black lines) to show marker expression in overlays. The expression of Blk and IL-17A could be discerned only in mature V$_\gamma$2$^+$ thymocytes, whereas immV2 cells had low expression of RORγt (**Fig. 2**). For analysis of IL-17A expression, cells were stimulated for 4 h with the phorbol ester PMA and ionomycin, followed by surface and intracellular staining of cytokines. Data are from two to three independent experiments with 4–30 mice each (**a**,**b**) or are from one experiment representative of six with at least three mice each (**c**).

for the proposal of thymic programming of immV2 cells, expression of the genes encoding the three markers of IL-17-producing γδ T cells—the signaing molecule Blk[26] (**Supplementary Fig. 4f**) and the scavenger receptors SCART1 and SCART2 (ref. 27)—was higher in immV2 cells than in other subsets (**Supplementary Tables 2**, **5** and **6**). The expression of genes encoding other factors required for the differentiation of αβ effector helper T cells, including c-Maf (encoded by *Maf*)[28], GATA-3 (encoded by *Gata3*)[29,30] and Runx3 (encoded by *Runx3*)[31], was also higher in immV2 cells than in other γδ subsets (**Figs. 2** and **3c**). These results demonstrated that before or concurrent with expression of various TCRs, γδ thymocyte subsets are regulated by divergent transcription-factor networks that are predicted to program distinct effector functions.

Maturation induces homing and functional markers

The biological processes associated with thymic γδ T cell maturation have not been systematically studied. We determined the gene-expression profiles of mature (CD24^{lo}) thymic γδ T cell subsets (V2, V1 and V6) to assess the duration and effect of the immature subset–specific gene networks (**Fig. 1**). Despite the substantial subset-specific differences among γδ T cells, there was a shared set of 495 genes that were coordinately regulated after maturation of all γδ T cells (**Supplementary Fig. 5a**,**b** and **Supplementary Table 7**). PCA, hierarchical clustering, Euclidian distances and correlation scores of these γδ maturation–dependent genes in other thymocyte subsets showed that the γδ maturation signature was conserved during αβ T cell maturation as well, as precursors and immature subsets formed a distinct cluster, separate from the cluster of mature subsets, regardless of T cell lineage (**Fig. 4a**,**b** and **Supplementary Fig. 5c**,**d**). Of the 495 γδ-maturation genes identified, ~78% were coordinately regulated after the maturation of αβ T cells subsets (comparison of DP cells with mature CD4$^+$ cells, mature CD8$^+$ cells and *i*NKT cells), whereas ~22% were uniquely modulated during γδ T cell maturation only (data not shown). These results established the loss of CD24 as a demarcation of the transition to maturity in the γδ T cell lineage.

Four overriding themes characterized γδ subset–specific maturation. First, the expression pattern of genes encoding molecules involved in metabolism or RNA-DNA biogenesis in mature V2 (matV2) cells and mature V1 (matV1) cells converged, whereas mature V6 (matV6) cells remained distinct (**Fig. 4c** and **Supplementary Table 8**). Second, after maturation, V2 cells downregulated subset-specific transcription factors to the amount of expression associated with V1 cells (**Fig. 4d** and **Supplementary Table 9**). Third, mature γδ T cell subsets modulated the expression of proteins directly involved in the execution of their peripheral subset–specific effector functions, including chemokine and cytokine receptors. Fourth, V1 cells and V6 cells diverged after transition to the mature state, with the latter acquiring the gene-expression profile of PLZF$^+$ αβ *i*NKT cells. The last two features of γδ T cell maturation are discussed in greater detail below.

Transition to the mature stage endows subsets of γδ cells from adult mice with unique homing properties, as shown by the acquisition and loss of chemokine and integrin receptors. We noted the following two patterns of chemokine-receptor expression: expression in immature subsets that was extinguished or downregulated as γδ T cells matured; and expression induced precipitously during maturation and maintained in fully differentiated subsets (**Fig. 5a** and **Supplementary Table 10**). We confirmed the expression pattern of the chemokine receptor CCR10 by analysis of reporter mice that express green fluorescent protein driven by the promoter of the gene encoding CCR10; the greatest frequency of CCR10$^+$ immature thymocytes was in the V6 subset (**Supplementary Fig. 1**). Expression of the chemokine receptor CCR6 was considerably and uniquely upregulated in the V2 subset after transition to the mature stage, whereas induction of the chemokine receptor CXCR3 was associated with maturation of the V6 and V1 subsets (**Fig. 5a**, **Supplementary Fig. 1** and **Supplementary Table 10**). CCR6 expression is tightly correlated with IL-17-producing lymphocytes[32], whereas CXCR3 expression is controlled in part by Eomes[33] and permits trafficking to nonlymphoid tissues. Notably, matV2 cells are ITGB7^{hi}CXCR6^{hi}CCR6$^+$ RORγt$^+$IL-7R^{hi}, reminiscent of fetal liver–derived RORγt$^+$ CD3$^-$ innate lymphoid cells, some with the ability to secrete IL-17 or IL-22 (ref. 34). The finding of dynamic subset-specific alterations in tissue tropism after maturation further emphasizes the separation of V2 cells from other γδ T cells and suggests that the molecular program that specifies IL-17 production in γδ T cells may overlap that of TCR$^-$ innate lymphoid cells.

Maturation stage–dependent alterations in cytokine-receptor expression were also specific to γδ T cell subsets and endowed γδ subsets with the ability to sense environmental cues that fully engaged their programmed effector functions. After thymic maturation, V2 cells entered the effector-poised phase by abrupt super-induction of genes encoding cytokine receptors dedicated to IL-17 production and responsiveness: *Il23r* (the gene most induced; ~100-fold higher

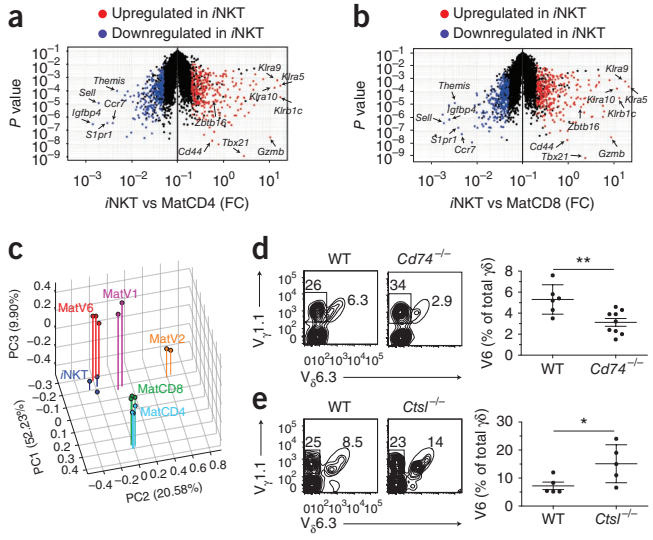

Figure 6 Common features of αβ *i*NKT cells and γδ matV6 cells. (**a**,**b**) Identification of *i*NKT cell signature genes on the basis of expression altered by twofold or more ('fold change' (FC)) in both αβ NKT cells versus (vs) mature CD4$^+$ cells (**a**) and αβ NKT cells versus mature CD8$^+$ cells (**b**), plus a coefficient of variation of <0.5 and an MEV of >120 in at least one subset. A total of 292 genes had higher expression in *i*NKT cells than mature CD4$^+$ or CD8$^+$ cells (red); 248 genes had lower expression in *i*NKT cells than in mature CD4$^+$ or CD8$^+$ cells (blue). (**c**) PCA (as in **Fig. 2b**) of the 540 genes with different regulation in *i*NKT cells versus mature CD4$^+$ or CD8$^+$ cells. (**d**,**e**) Flow cytometry analysis (left) of staining for V_γ1.1 and V_δ6.3 on C57BL/6 wild-type (WT) and *Cd74*$^{-/-}$ γδ thymocytes (**d**) or wild-type mice and *Ctsl*$^{-/-}$ γδ thymocytes (**e**), and frequency of V6 cells among γδ T cells (right; in **d**, $n = 6$ wild-type mice and $n = 9$ *Cd74*$^{-/-}$ mice; in **e**, $n = 5$ wild-type mice and $n = 5$ *Ctsl*$^{-/-}$ mice). Numbers adjacent to outlined areas (left) indicate percent cells in each. Each symbol (right) represents an individual mouse; small horizontal bars represent the mean (± s.d.). $^*P = 0.04$ and $^{**}P = 0.005$ (two-tailed Student's *t*-test). Data are from two to three independent experiments with 4–30 mice each (**a**–**c**) or are representative of (**d**,**e**, left) or pooled from (**d**,**e**, right) two independent experiments.

expression in CD24^{lo} cells than in CD24^{hi} cells), *Il17re, Il1r1, Il18r1* (ref. 35) and *Tnfrsf25* (ref. 36) were the genes with the greatest induction in matV2 cells (**Fig. 5b** and **Supplementary Tables 7** and **11**). This super-induction was accompanied by enhanced expression of RORγt and the signaling molecule Blk dedicated to the production of IL-17 (**Fig. 5c**). Receptors for four counter-regulators or modulators of IL-17 production had divergent expression after maturation, with matV2 cells having higher expression of those encoding IL-2 and IL-27, whereas the expression of *Il17rb* (which encodes part of the receptor for IL-25) and *Il12rb2* (which encodes part of the receptor for IL-12R) in matV2 cells was 10% and 25%, respectively, of the expression in immature cells (**Fig. 5b** and **Supplementary Table 11**). Thus, matV2 cells are programmed to respond to autocrine IL-17 turned on by paracrine IL-23, with multiple additional receptors for fine-tuning the response. The main gene upregulated by V1 cells after maturation was *Il18r1*, and this pattern was shared among the mature γδ T cell subsets. Expression of *Il17rb* was selectively higher in V6 cells, but lower in V1 and V2 cells, after maturation (**Fig. 5b** and **Supplementary Table 11**). IL-25 drives the population expansion of innate lymphocytes that produce T helper type 2 cytokines (IL-5 and IL-13), including αβ NKT cells and TCR$^-$ type 2 cells[34]. The selective induction of *Il17rb*, which encodes part of the receptor for IL-25, in IL-4-producing V6 cells indicated that like the acquisition of *Il23r* expression by matV2 cells, the quintessential cytokine responsiveness of γδ effector subsets is intrathymically programmed. Coincident with the overall modulation of cytokine-receptor expression as the subsets matured, the expression of genes encoding effector cytokines was also initiated specifically at the mature stage. The best illustration of this pattern was matV6 cells with high expression of PLZF that correlated with abundant *Il4* transcription (**Supplementary Fig. 5e**) and matV2 cells that produced IL-17A (**Fig. 5c**).

In sum, our results indicated that the unfolding gene programs that dictated γδ subset–specific effector functions occurred by the following two steps: an establishment phase at the immature stage; and a primed phase at the mature stage that equipped the thymus-exiting γδ thymocytes with tissue-migratory cues as well as sensors that, once engaged at tissue sites, will rapidly induce effector molecules (**Supplementary Fig. 6a**). The phase transition was accompanied by an overall dampening of the activities of transcription factors that established the effector gene circuits as γδ thymocytes first arose. This overall pattern of a dynamic transcription-factor expression program fits well with the established gene-regulatory networks that control cell differentiation in other systems[37].

Acquisition of the *i*NKT cell signature by matV6 cells

We predicted that matV6 cells would overlap αβ *i*NKT cells at the gene-expression level, given their shared expression of PLZF and similar functional properties. To determine the extent of this similarity, we first derived the gene signature of the αβ *i*NKT cell lineage by identifying genes with different expression in thymic *i*NKT cells versus other mature conventional αβ thymocyte subsets (**Fig. 6a**,**b** and **Supplementary Table 12**). PCA, Euclidian distances and correlation scores of the 540 *i*NKT signature genes showed that matV6 cells shared this expression signature, as matV6 and *i*NKT populations were nearly identical in their placement along principal components 1 and 2 (**Fig. 6c** and **Supplementary Fig. 6b**,**c**). This indicated that acquisition of the *i*NKT gene signature was independent of TCR type (αβ versus γδ). Genetic studies also supported the relatedness of V6 cells and αβ *i*NKT cells: *in vivo*, the loss of *Zbtb16* (which encodes PLZF), *Sh2d1a* (which encodes SAP), *Itk* (which encodes the tyrosine kinase Itk), *Id3* (which encodes the transcriptional inhibitor Id3), *Ctsl* (which encodes cathepsin L) and *Cd74* (which encodes the invariant chain) has a negative effect on the development of αβ *i*NKT cells[38]. The same gene deficiencies also specifically affected V6 cells, but not all led to the production of fewer V6 cells[39,40] (**Fig. 6d**,**e**). Together these results indicated that common gene networks and cellular processes regulated the generation of γδ V6 and αβ *i*NKT cells. Hence, either molecular programming by similar transcription-factor networks or a similarity in TCR-signaling features independent of the TCR type mainly specifies effector-cell fate.

DISCUSSION

Systematic global gene-expression profiling of *ex vivo* γδTCR$^+$ thymocyte subsets distinguished by TCRγ and TCRδ repertoires demonstrated that the early divergence in gene-expression programs of emergent γδ cell subsets underpins the effector diversification of innate T cells. We have identified the following three distinct cell types in the γδ T cell lineage on the basis of TCR repertoire and shared gene-expression profiles: IL-17-producing γδ T cells, which include V2 cells from adult and fetal mice and V4 cells from fetal mice; V3 cells from fetal mice[17,18]; and V1, V5 and V6 cells from adult mice. We could not directly examine the only two remaining γδ T cell subsets on the basis of TCR repertoire in the adult thymus (V_γ1.2$^+$ and V_γ4$^+$ cells that make up the residual ~5% of total γδTCR$^+$ thymocytes), as

antibodies to sort these cells are not available. However, analysis of immature $V_\gamma1.1^-V_\gamma2^-V_\gamma5^-$ γδ thymocytes from adult mice (consisting of ~85% $V_\gamma1.2^+$ and 15% $V_\gamma4^+$ cells, as estimated on the basis of the observed functional *Tcrg* gene-rearrangement frequencies[41]) has suggested that immature $V_\gamma1.2^+$ cells resemble immV1, immV5 and immV6 thymocytes in gene expression (data not shown).

The biological processes that govern the generation of these three distinct immature γδ cell subsets are unknown, but the list of possibilities can now be narrowed down. Data indicate that αβ TCR^+ thymocytes do not appreciably alter the effector differentiation of γδ T cells in *trans* (K.N. *et al.*, data not shown). The most notable gene modules that distinguished the emergent immature γδ thymocyte subsets encoded transcription factors. The distinct responsiveness of precursor cells to cytokines and growth factors that control the expression of the transcription factors that define effector subtypes is one possible mechanism for the generation of programmed effector cells. The high-mobility-group transcription factor SOX13, its interacting partners TCF1-LEF1, and their gene targets are one central gene circuit involved in the differentiation of IL-17-producing cells from adult mice, as $Sox13^{-/-}$ mice lack V2 IL-17-producing cells (N.M. *et al.*, data not shown). Given the modulatory effect of Wnt signaling on TCF1-LEF1, canonical Wnt ligands are another potential regulator of the diversification of γδ effector cells. Alternatively, genomically encoded TCR V_γ chains may transmit distinct signals, either by unique interactions with selecting ligand(s)[13] or on the basis of intrinsic differences in TCR conformation and/or membrane localization, independently of selecting ligands presented in *trans*[42]. For αβ T cells, the pre-TCR-mediated transition to the immature DP stage occurs in a ligand-independent manner. Similarly, there is an *in vitro* precedent that the $V_\gamma2V_\delta5$ TCR can also signal in a ligand-independent, TCR dimerization–dependent manner[13]. The finding of similar expression of a gene cluster encoding molecules involved in cellular metabolism in V2 cells and DP cells of the αβ lineage supported the possibility that unique TCR signaling is one potential discriminator of effector subset diversification.

Finally, it is possible that the distinct timing of the generation different V_γ TCRs during precursor maturation leads to alternative cell fates by inheritance and fixation of distinct precursor gene-expression programs. The most clear-cut example of this is the programmed $V_\gamma3$ gene rearrangement exclusively in early T cell progenitors from fetal mice that generates dendritic epidermal T cells[18]. It has also been well documented that the timing of TCR expression during precursor maturation, rather than TCR type, is critical for the definition of the resultant T cell properties[43,44]. The proposal that γδ T cell subsets may originate at different stages of precursor maturation is supported by several studies indicating that the three functional C_γ loci are regulated distinctly and independently[45,46]. Most directly, we found that the earliest c-Kit^+ T cell progenitors were skewed toward the generation of $CCR6^+$ V2 cells in cultures of OP9 mouse bone marrow stromal cells expressing Delta-like ligand 1 (N.M. *et al.*, data not shown), which suggested that the difference in the onset of functional TCR γ-chain expression with the attendant inheritance of distinct collective gene activities of the precursors may underpin the molecular heterogeneity of γδ subsets.

The identification of gene networks that specify the effector functions of subsets of innate γδ T cells that are established at the earliest stage of γδ thymocyte differentiation and fixed during thymic maturation suggests that γδ T cells probably do not respond homogeneously to external cues. Elucidation of the upstream regulators of these gene networks will identify the mechanism of the thymic programming of effector subset diversification of innate γδ T cells. Moreover, developmentally hardwired programs for the production of cytokines of the IL-17 family by other innate lymphocyte subsets in the gut are predicted to use genetic networks similar to those that operate in IL-17-producing γδ T cells[34,47], which indicates a possible route for the identification of a unified origin of innate lymphoid effector subsets.

METHODS

Methods and any associated references are available in the online version of the paper at http://www.nature.com/natureimmunology/.

Accession code. GEO: microarray data, GSE15907.

Note: Supplementary information is available on the Nature Immunology website.

ACKNOWLEDGMENTS

We thank M. Mohrs (Trudeau Institute) for IL-4–GFP reporter mice; H. Birchmeier (Max-Delbrück-Center for Molecular Medicine Berlin) for *Axin2* reporter mice; V. Lefebvre (Cleveland Clinic) for *Sox5* reporter mice; K. Rock (University of Massachusetts Medical School) for $Ctsl^{-/-}$ mice; E. Huseby (University of Massachusetts Medical School) for $Cd74^{-/-}$ mice; S. Davis (Harvard Medical School) for the ConsolidateProbeSets module and PopulationDistances PCA program; A. Hayday (King's College, London) for anti-$V_\delta1$ (17D1); members of the ImmGen Consortium for discussions; the ImmGen core team (M. Painter, J. Ericson and S. Davis) for help with data generation and processing; and eBioscience, Affymetrix and Expression Analysis for support of the ImmGen Project. Core resources supported by the Diabetes Endocrinology Research Center (DK32520) were used. Supported by the National Institute of Allergy and Infectious Diseases of the US National Institutes of Health (R24 AI072073 to the ImmGen group, and CA100382 to J.K.).

AUTHOR CONTRIBUTIONS

K.E.S. sorted cell subsets; K.N., N.M., K.E.S. and C.C.Y. did follow-up experiments and analyzed data; G.M., T.V. and K.N. did the studies of *M. tuberculosis*; H.K. supervised the studies of *M. tuberculosis*; N.X. provided reagents and cells from a mutant strain; N.R.C. and M.B.B. generated the gene-expression profiles of NKT cell subsets; L.J.B. provided reagents and shared data used in the interpretation of some results; K.N. analyzed gene-expression data; and J.K. and K.N. designed studies and wrote the paper.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at http://www.nature.com/natureimmunology/.
Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

1. Heng, T.S. & Painter, M.W. The Immunological Genome Project: networks of gene expression in immune cells. *Nat. Immunol.* **9**, 1091–1094 (2008).
2. Ferrick, D.A. *et al.* Differential production of interferon-γ and interleukin-4 in response to Th1- and Th2-stimulating pathogens by γδ T cells *in vivo*. *Nature* **373**, 255–257 (1995).
3. Stark, M.A. *et al.* Phagocytosis of apoptotic neutrophils regulates granulopoiesis via IL-23 and IL-17. *Immunity* **22**, 285–294 (2005).
4. Hayday, A. & Tigelaar, R. Immunoregulation in the tissues by γδ T cells. *Nat. Rev. Immunol.* **3**, 233–242 (2003).
5. Berg, L.J. Signalling through TEC kinases regulates conventional versus innate $CD8^+$ T-cell development. *Nat. Rev. Immunol.* **7**, 479–485 (2007).
6. Lee, Y.J., Jameson, S.C. & Hogquist, K.A. Alternative memory in the CD8 T cell lineage. *Trends Immunol.* **32**, 50–56 (2011).
7. Lockhart, E., Green, A.M. & Flynn, J.L. IL-17 production is dominated by γδ T cells rather than CD4 T cells during *Mycobacterium tuberculosis* infection. *J. Immunol.* **177**, 4662–4669 (2006).
8. Martin, B., Hirota, K., Cua, D.J., Stockinger, B. & Veldhoen, M. Interleukin-17-producing γδ T cells selectively expand in response to pathogen products and environmental signals. *Immunity* **31**, 321–330 (2009).
9. Sutton, C.E. *et al.* Interleukin-1 and IL-23 induce innate IL-17 production from γδ T cells, amplifying Th17 responses and autoimmunity. *Immunity* **31**, 331–341 (2009).
10. Melichar, H.J. *et al.* Regulation of γδ versus αβ T lymphocyte differentiation by the transcription factor SOX13. *Science* **315**, 230–233 (2007).
11. O'Brien, R.L. & Born, W.K. γδ T cell subsets: a link between TCR and function? *Semin. Immunol.* **22**, 193–198 (2010).
12. Azuara, V., Levraud, J.P., Lembezat, M.P. & Pereira, P. A novel subset of adult γδ thymocytes that secretes a distinct pattern of cytokines and expresses a very restricted T cell receptor repertoire. *Eur. J. Immunol.* **27**, 544–553 (1997).

13. Jensen, K.D. *et al.* Thymic selection determines γδ T cell effector fate: antigen-naive cells make interleukin-17 and antigen-experienced cells make interferon gamma. *Immunity* **29**, 90–100 (2008).
14. Ribot, J.C. *et al.* CD27 is a thymic determinant of the balance between interferon-γ- and interleukin 17-producing γδ T cell subsets. *Nat. Immunol.* **10**, 427–436 (2009).
15. Shires, J., Theodoridis, E. & Hayday, A.C. Biological insights into TCRγδ$^+$ and TCRαβ$^+$ intraepithelial lymphocytes provided by serial analysis of gene expression (SAGE). *Immunity* **15**, 419–434 (2001).
16. Shibata, K. *et al.* Identification of CD25$^+$ γδ T cells as fetal thymus-derived naturally occurring IL-17 producers. *J. Immunol.* **181**, 5940–5947 (2008).
17. Ikuta, K. *et al.* A developmental switch in thymic lymphocyte maturation potential occurs at the level of hematopoietic stem cells. *Cell* **62**, 863–874 (1990).
18. Xiong, N., Kang, C. & Raulet, D.H. Positive selection of dendritic epidermal γδ T cell precursors in the fetal thymus determines expression of skin-homing receptors. *Immunity* **21**, 121–131 (2004).
19. 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).
20. Haks, M.C. *et al.* Attenuation of γδTCR signaling efficiently diverts thymocytes to the αβ lineage. *Immunity* **22**, 595–606 (2005).
21. Rothenberg, E.V., Zhang, J. & Li, L. Multilayered specification of the T-cell lineage fate. *Immunol. Rev.* **238**, 150–168 (2010).
22. Ivanov, I.I. *et al.* The orphan nuclear receptor RORγt directs the differentiation program of proinflammatory IL-17$^+$ T helper cells. *Cell* **126**, 1121–1133 (2006).
23. Pearce, E.L. *et al.* Control of effector CD8$^+$ T cell function by the transcription factor Eomesodermin. *Science* **302**, 1041–1043 (2003).
24. Savage, A.K. *et al.* The transcription factor PLZF directs the effector program of the NKT cell lineage. *Immunity* **29**, 391–403 (2008).
25. Kovalovsky, D. *et al.* The BTB-zinc finger transcriptional regulator PLZF controls the development of invariant natural killer T cell effector functions. *Nat. Immunol.* **9**, 1055–1064 (2008).
26. Laird, R.M., Laky, K. & Hayes, S.M. Unexpected role for the B cell-specific Src family kinase B lymphoid kinase in the development of IL-17-producing γδ T cells. *J. Immunol.* **185**, 6518–6527 (2010).
27. Kisielow, J., Kopf, M. & Karjalainen, K. SCART scavenger receptors identify a novel subset of adult γδ T cells. *J. Immunol.* **181**, 1710–1716 (2008).
28. Bauquet, A.T. *et al.* The costimulatory molecule ICOS regulates the expression of c-Maf and IL-21 in the development of follicular T helper cells and TH-17 cells. *Nat. Immunol.* **10**, 167–175 (2009).
29. Zheng, W. & Flavell, R.A. The transcription factor GATA-3 is necessary and sufficient for Th2 cytokine gene expression in CD4 T cells. *Cell* **89**, 587–596 (1997).
30. Maruyama, T. *et al.* Control of the differentiation of regulatory T cells and T_H17 cells by the DNA-binding inhibitor Id3. *Nat. Immunol.* **12**, 86–95 (2011).
31. Djuretic, I.M. *et al.* Transcription factors T-bet and Runx3 cooperate to activate *Ifng* and silence *Il4* in T helper type 1 cells. *Nat. Immunol.* **8**, 145–153 (2007).
32. Acosta-Rodriguez, E.V. *et al.* Surface phenotype and antigenic specificity of human interleukin 17-producing T helper memory cells. *Nat. Immunol.* **8**, 639–646 (2007).
33. Weinreich, M.A. *et al.* KLF2 transcription-factor deficiency in T cells results in unrestrained cytokine production and upregulation of bystander chemokine receptors. *Immunity* **31**, 122–130 (2009).
34. Spits, H. & Di Santo, J.P. The expanding family of innate lymphoid cells: regulators and effectors of immunity and tissue remodeling. *Nat. Immunol.* **12**, 21–27 (2011).
35. Andrews, D.M. *et al.* Homeostatic defects in interleukin 18-deficient mice contribute to protection against the lethal effects of endotoxin. *Immunol. Cell Biol.* **89**, 739–746 (2011).
36. Pappu, B.P. *et al.* TL1A–DR3 interaction regulates Th17 cell function and Th17-mediated autoimmune disease. *J. Exp. Med.* **205**, 1049–1062 (2008).
37. Yosef, N. & Regev, A. Impulse control: temporal dynamics in gene transcription. *Cell* **144**, 886–896 (2011).
38. Bendelac, A., Savage, P.B. & Teyton, L. The biology of NKT cells. *Annu. Rev. Immunol.* **25**, 297–336 (2007).
39. Felices, M., Yin, C.C., Kosaka, Y., Kang, J. & Berg, L.J. Tec kinase Itk in γδ T cells is pivotal for controlling IgE production in vivo. *Proc. Natl. Acad. Sci. USA* **106**, 8308–8313 (2009).
40. Verykokakis, M. *et al.* Inhibitor of DNA binding 3 limits development of murine slam-associated adaptor protein-dependent "innate" γδ T cells. *PLoS ONE* **5**, e9303 (2010).
41. Pereira, P. & Boucontet, L. Rates of recombination and chain pair biases greatly influence the primary γδ TCR repertoire in the thymus of adult mice. *J. Immunol.* **173**, 3261–3270 (2004).
42. Kuhns, M.S. & Davis, M.M. Disruption of extracellular interactions impairs T cell receptor-CD3 complex stability and signaling. *Immunity* **26**, 357–369 (2007).
43. Bruno, L., Fehling, H.J. & von Boehmer, H. The ab T cell receptor can replace the γδ receptor in the development of γδ lineage cells. *Immunity* **5**, 343–352 (1996).
44. Baldwin, T.A., Sandau, M.M., Jameson, S.C. & Hogquist, K.A. The timing of TCR alpha expression critically influences T cell development and selection. *J. Exp. Med.* **202**, 111–121 (2005).
45. Passoni, L. *et al.* Intrathymic d selection events in γδ cell development. *Immunity* **7**, 83–95 (1997).
46. Livák, F., Tourigny, M., Schatz, D.G. & Petrie, H.T. Characterization of TCR gene rearrangements during adult murine T cell development. *J. Immunol.* **162**, 2575–2580 (1999).
47. Aliahmad, P., de la Torre, B. & Kaye, J. Shared dependence on the DNA-binding factor TOX for the development of lymphoid tissue-inducer cell and NK cell lineages. *Nat. Immunol.* **11**, 945–952 (2010).

The complete list of authors is as follows:

Yan Zhou[7], Susan A Shinton[7], Richard R Hardy[7], Natalie A Bezman[8], Joseph C Sun[8], Charlie C Kim[8], Lewis L Lanier[8], Jennifer Miller[9], Brian Brown[9], Miriam Merad[9], Anne Fletcher[10], Kutlu Elpek[10], Angelique Bellemare-Pelletier[10], Deepali Malhotra[10], Shannon Turley[10], Kavitha Narayan[11], Katelyn Sylvia[11], Joonsoo Kang[11], Roi Gazit[12], Brian Garrison[12], Derrick J Rossi[12], Vladimir Jojic[13], Daphne Koller[13], Radu Jianu[14], David Laidlaw[14], James Costello[15], Jim Collins[15], Nadia Cohen[16], Patrick Brennan[16], Michael Brenner[16], Tal Shay[17], Aviv Regev[17], Francis Kim[18], Tata Nageswara Rao[18], Amy Wagers[18], Emmanuel L Gautier[9,19], Claudia Jakubzick[9,19], Gwendalyn J Randolph[9,19], Paul Monach[20], Adam J Best[21], Jamie Knell[21], Ananda Goldrath[21], Tracy Heng[22], Taras Kreslavsky[10], Michio Painter[22], Diane Mathis[22] & Christophe Benoist[22]

[7]Fox Chase Cancer Center, Philadelphia, Pennsylvania, USA. [8]Department of Microbiology & Immunology, University of California San Francisco, San Francisco, California, USA. [9]Icahn Medical Institute, Mount Sinai Hospital, New York, New York, USA. [10]Dana-Farber Cancer Institute and Harvard Medical School, Boston, Massachusetts, USA. [11]University of Massachusetts Medical School, Worcester, Massachusetts, USA. [12]Immune Disease Institute, Children's Hospital, Boston, Massachusetts, USA. [13]Computer Science Department, Stanford University, Stanford, California, USA. [14]Computer Science Department, Brown University, Providence, Rhode Island, USA. [15]Department of Biomedical Engineering, Howard Hughes Medical Institute, Boston University, Boston, Massachusetts, USA. [16]Division of Rheumatology, Immunology and Allergy, Brigham and Women's Hospital, Boston, Massachusetts, USA. [17]Broad Institute, Cambridge, Massachusetts, USA. [18]Joslin Diabetes Center, Boston, Massachusetts, USA. [19]Department of Pathology & Immunology, Washington University, St. Louis, Missouri, USA. [20]Department of Medicine, Boston University, Boston, Massachusetts, USA. [21]Division of Biological Sciences, University of California San Diego, La Jolla, California, USA. [22]Division of Immunology, Department of Microbiology & Immunobiology, Harvard Medical School, Boston, Massachusetts, USA.

ONLINE METHODS

Mice. Male C57BL/6J mice 5 weeks of age (Jackson Labs) were used for microarray analysis within 1 week of arrival. CCR10-GFP reporter mice[48], IL-4–GFP[49], Axin2-LacZ reporter mice[50] and Sox5-LacZ reporter mice[51] have been described. A *Cenpk* (encoding the centromere protein SOLT) LacZ reporter gene-trap embryonic stem cell clone (International Gene Trap Consotrium) was used for the generation of reporter mice. *Ctsl*$^{-/-}$ mice were provided by K. Rock, and *Cd74*$^{-/-}$ mice were provided by E. Huseby. Mice were housed in a specific pathogen–free rodent barrier facility. Experiments were approved by the Institutional Animal Care and Use Committee of the University of Massachusetts Medical School.

Sample preparation for microarray analysis. Pooled thymocytes from 4–30 mice were enriched for γδ T cells by depletion of CD8$^+$ cells with magnetic beads and an autoMACS, then were stained and sorted with a FACSAria (~2×10^4 to 3×10^4 cells; >99% pure) directly into TRIzol (Invitorgen). Independent triplicates were sorted unless noted otherwise (complete sorting details available from the ImmGen Project). As antibodies to V_γ4 are not available, immV4 cells were sorted from fetal mice by gating on TCRδ$^+$ cells that were $V_\gamma1.1^-V_\gamma2^-V_\gamma3^-V_\gamma5^-$ (estimated purity, ~95%). The approximate frequencies of TCR δ-chains associated with sorted TCRγ subsets were as follows: V2 (~45% of total γδ cells; $V_\gamma2$ paired with $V_\delta4$ (50%), $V_\delta5$ (40%), all $V_\delta6.3^-$); V1 (30% of total γδ cells; $V_\gamma1.1$ paired wth diverse Vδ chains including $V_\delta4$ (25%), $V_\delta5$ (15%), others at lower frequencies, all $V_\delta6.3^-$); V6 (~8% of total γδ cells; $V_\gamma1.1$ paired with $V_\delta6.3$ (100%)); V5 (~5% of total γδ cells; $V_\gamma5$ paired with $V_\delta5$ (40%) and various others at lower frequencies); and V3 and V4 thymocytes from fetal mice coexpressing $V_\delta1$.

Data analysis and visualization. RNA processing and microarray analysis with the Affymetrix MoGene 1.0 ST array was done at the ImmGen processing center (by ImmGen standard operating procedures; http://www.immgen.org/Protocols/ImmGen Cell prep and sorting SOP.pdf; **Supplementary Notes 1** and **2**). Data were analyzed with modules of the GenePattern genomic analysis platform. Raw data were normalized by the robust multi-array average method with quantile normalization and background correction (ExpressionFileCreator module of GenePattern). ConsolidatedProbeSets, a custom GenePattern module written by S. Davis (Harvard Medical School), was used for the consolidation of multiple probe sets into a single mean probeset value for each gene. Genes with differences in regulation were identified with the Multiplot module of GenePattern through the use of the average of all replicates. Unless indicated otherwise, genes were considered to be regulated differently if they differed in expression by more than twofold, had a coefficient of variation among replicates of less than 0.5, had a Student's *t*-test *P* value of less than 0.05, and had a mean expression value of more than 120 in at least one subset in the comparison. Heat maps were generated by hierarchical clustering (HierarchicalClustering module of GenePattern) of data on the basis of gene (row) and subset (column) with the Pearson correlation for distance measurement. Data were log transformed and clustered with pairwise complete linkage. Data were centered on rows before visualization with the HeatMapViewer module of GenePattern. Principal component analysis was done with the PopulationDistances PCA program (http://cbdm.hms.harvard.edu/LabMembersPges/SD.html). Where indicated, the PCA program was used for the identification of the 15% of genes with the greatest difference in expression among subsets by filtering on the basis of a variation of analysis of variance with the geometric standard deviation of populations to weight genes that vary in multiple populations. Data were log transformed, normalized by gene and subset, and filtered for genes with an MEV of >120 before visualization. Euclidian distance and Pearson's correlation coefficients were calculated with the 'dist' and 'cor' commands in R to generate distance and correlation matrixes for a given set of genes and subsets. Heat maps of Euclidian distance and Pearson's correlation coefficients were generated by hierarchical clustering (HierarchicalClustering module of GenePattern, Pearson's correlation for row and column, complete linkage). Data were visualized with a global color scheme in the HeatMapViewer module of GenePattern. Ingenuity software and manual inspection were used for pathway analysis. The AmiGO pathway (Gene Ontology Project) and KEGG pathway (Kyoto Encyclopedia of Genes and Genomes) were used for some functional classifications.

Flow cytometry. The following antibodies to cell surface molecules were used: anti-CD25 (PC61), anti-CD27 (LG.3A10), anti-$V_\gamma2$ (UC3-10A6), anti-$V_\gamma3$ (536), anti-$V_\delta6.3$ (8F4H7B7), anti-Ly49C/I (5E6), anti-CCR6 (140706), anti-Thy-1.2 (53-2.1), anti-CD9 (KMC8) and streptavidin-allophycocyanin (all from BD Biosciences); anti-CD3 (145-2C11), anti-CD4 (RM4-5), anti-CD8 (53-6.7), anti-CD44 (IM7), anti-CD122 (5H4), anti-CD127 (IL-7Rα; A7R34), anti-CD24 (heat-stable antigen; M1/69), anti-c-Kit (2B8), anti-TCRδ (GL3), anti-NKG2D (CX5), anti-CD62L (MEL-15), anti-PD1 (J43), anti-CD73 (TY/11.8), anti-CD199 (CCR9, CW-1.2), anti-CD197 (CCR7, 4B12), anti-CD183 (CXCR3, 174), anti-CD184 (CXCR4, 2B11), anti-CD244.2 (2B4, 244F4), anti-NKG2ACE (20D5), anti-GITR (DTA-1), anti-CD48 (HM48-1), anti-CD28 (37.51) and streptavidin-phycoerythrin-indotricarbocyanine (all from eBioscience); anti-IL-6R (D7715A7; BioLegend); and anti-CD119 (IFN-γRα; GR20; RDI–Fitzgerald Industries). Anti-$V_\gamma1.1$ (2.11) was purified by Bio-XCell and biotinylated with the FluoReporter Mini-Biotin-XX Labeling Kit (Invitrogen). Anti-$V_\delta1$ (17D1) was provided by A. Hayday. Intracellular IL-17 (17B7; eBioscience) and IFN-γ (XMG1.2; BD Biosciences) were detected with the Cytofix/Cytoperm kit (BD Biosciences). Intranuclear staining was done with the Foxp3 Staining Kit (eBioscience) with the following antibodies: anti-TOX (TXRX10), anti-EOMES (Dan11mag), anti-RORγt (AFKJS-9) and anti-GATA-3 (L50-823; all from eBioscience); anti-Blk (3262) and anti-LEF-1 (C12A5; both from Cell Signaling); anti-Ki67 (B56; BD Biosciences); and anti-PLZF (D-9) and anti-SMO (N-19; both from Santa Cruz). Staining for fluorescein di-β-D-galactopyranoside (F1179; Invitrogen) was done according to standard protocols. Data were acquired on a BD LSRII and analyzed with FlowJo (Treestar). In **Supplementary Figure 1**, data were concatenated from multiple independent samples in FlowJo for visualization in histograms.

***Ex vivo* stimulation.** Cells were stimulated *ex vivo* by culture of total cells (2×10^6 cells per well) for 4 h at 37 °C with PMA (phorbol 12-myristate 13-acetate; 10 ng/ml) and Ionomycin (1 ug/ml), with Golgi Stop and Golgi Plug (BD Biosciences) added after 1 h, according to the manufacturers' protocols.

Infection with *M. tuberculosis* and analysis. Wild-type strains of *M. tuberculosis* were used for the infection of four to ten male 6- to 8-week-old C57BL/6 mice (Jackson Labs) with 2×10^3 colony-forming units of *M. tuberculosis* by the aerosol route. Two weeks after infection, mice were killed and cells were isolated from the spleen and lungs (enzymatic digestion) and stained immediately, or after culture with PMA and ionomycin.

Statistical analysis. For the identification of genes with different expression, Student's *t*-test *P* values were generated with Multiplot (Genepattern). For statistical analysis of flow cytometry data, Prism (GraphPad Software) was used. Data were tested for normality with F-tests and then were analyzed with unpaired two-tailed *t*-tests. Pathway analysis was done with Ingenuity Pathway Analysis and statistical significance was determined with the program's built-in Fisher's exact test.

48. Jin, Y., Xia, M., Sun, A., Saylor, C.M. & Xiong, N. CCR10 is important for the development of skin-specific γδ T cells by regulating their migration and location. *J. Immunol* **185**, 5723–5731 (2010).
49. Mohrs, M., Shinkai, K., Mohrs, K. & Locksley, R.M. Analysis of type 2 immunity in vivo with a bicistronic IL-4 reporter. *Immunity* **15**, 303–311 (2001).
50. Lustig, B. *et al.* Negative feedback loop of Wnt signaling through upregulation of conductin/axin2 in colorectal and liver tumors. *Mol. Cell. Bio.* **22**, 1184–1193 (2002).
51. Smits, P. *et al.* The transcription factors L-Sox5 and Sox6 are essential for cartilage formation. *Dev. Cell* **1**, 277–290 (2001).