HHS Public Access

Author manuscript

Nat Immunol. Author manuscript; available in PMC 2012 November 01.

Published in final edited form as:

Nat Immunol. ; 13(5): 511–518. doi:10.1038/ni.2247.

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

Kavitha Narayan[1,*], **Katelyn E. Sylvia**[1,*], **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], and **The Immunological Genome Project Consortium**

[1]Department of Pathology, University of Massachusetts Medical School, Worcester, MA 01655, USA

[2]Department of Medicine, University of Massachusetts Medical School, Worcester, MA 01655, USA

[3]Department of Veterinary and Biomedical Sciences, The Penn State University, University Park, PA 16802, USA

[4]Brigham and Women's Hospital and Department of Medicine, Division of Rheumatology, Immunology and Allergy, Harvard Medical School, Boston, Massachusetts 02115, USA

Abstract

γδ T cells function in the early phase of immune responses. Although innate γδ T cells have primarily been studied as one homogenous population, they can be functionally classified into effector subsets based on the production of signature cytokines, analogous to adaptive T helper subsets. Unlike adaptive T cells, however, γδ T effector function correlates with genomically encoded TCR chains, suggesting that clonal TCR selection is not the primary determinant of γδ effector differentiation. A high resolution transcriptome analysis of all emergent γδ thymocyte subsets segregated based on TCRγ/δ chain usage indicates the existence of three separate subtypes of γδ effectors in the thymus. The immature γδ subsets are distinguished by unique transcription factor modules that program effector function.

The Immunological Genome Project (ImmGen) is a consortium of immunologists and computational biologists constituted to comprehensively define gene regulatory networks in the immune system of the mouse [1]. In this context, we determined global gene expression

---

Correspondence: joonsoo.kang@umassmed.edu, Phone: 508-856-2759, Fax: 508-856-4556.
*These authors contributed equally to this work

Author contributions

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

Accession code

GEO: microarray data GSE 15907.

profiles of intrathymic γδTCR$^+$ cell subsets to ascertain the heterogeneity of the T cell lineage and to identify gene networks controlling innate effector subset production distinct from those that operate in conventional adaptive T cells. T cells in vertebrates are separated into two lineages based on the expression of either γδTCRs or αβTCRs on their cell surface. The adaptive αβ T cell lineage is subdivided into T helper 1 (T_H1), T_H2, T_H17, T_H follicular, regulatory T and cytotoxic effectors that are often considered distinct cell lineages. There are indications that the innate γδ T cell lineage is also composed of distinct subsets that are programmed to secrete a discrete cluster of effector cytokines [2, 3], but the mechanism of innate effector specialization is unclear. γδ T cells are sentinels of the immune system, mostly localized to the mucosal epithelia where pathogens are first encountered [4] and as such, 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. Upon infection, γδ T cells rapidly produce effector cytokines and growth factors similar to memory αβ T cells [7–9].

While the genes encoding 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 distinguishing them, other than the TCR and the transcription factor *Sox13* [10], that can be classified as bona-fide γδ T cell lineage-specific markers.. Identification of the genetic circuits underpinning two unique features of γδ T cell development is particularly critical. First, the segregation of adult γδ T cell effector functions is based on V_γ and V_δ gene usage [11]: IL-17, IFN-γ, and IL-4 production are associated with $V_\gamma2^+$ (designated as V2, Tγδ17), $V_\gamma1.1^+$ $V_\delta6.3^-$ (V1) and $V_\gamma1.1^+V_\delta6.3^+$ (V6) γδ T cell subsets, respectively. In comparison, conventional αβ T cells are classified into functional subsets based on the repertoire of effector cytokines produced, not by TCR repertoire, which is vastly diverse within each subset. Second, γδ T cell subset function appears to be programmed in the thymus [12–14]. In contrast, conventional αβ T cells differentiate into effector subsets upon encounter with pathogens in peripheral tissues. How and when γδ subsets are programmed and/or selected toward distinct effectors in the thymus is not well understood. To date, γδ T cell development has been studied mainly by analysis of total γδTCR$^+$ cells as one uniform population, not as distinct subsets defined by Vγ/Vδ usage. We compared the gene expression profiles of emergent adult γδTCR$^+$ thymocytes segregated based on Vγ/Vδ gene usage and show that the major subsets are as distinct from each other as they are from αβTCR$^+$ thymocyte subsets. Most important, the profiles of the emergent immature γδ T cell subsets were already embedded with unique gene programs directing subset-specific effector function, indicating that γδ function is molecularly programmed in the thymus prior to, or concurrent with, TCR expression.

RESULTS

Distinct subtypes among emerging γδTCR$^+$ thymocytes

A correlation between effector function and TCR Vγ/Vδ chain usage [11] raised the possibility that the earliest identifiable γδ T cells in the thymus are composed of molecularly heterogeneous cell subtypes distinguishable by the expression of uniquegermline-encoded

γδ TCRs. To test this possibility, we determined gene expression signatures of four adult γδ T cell subsets based on TCRγ and/or δ expression and maturational stages based on CD24 (HSA) expression: V2 ($V_\gamma2^+$, see On-line methods for partner δ chain repertoire and frequency among total γδTCR$^+$ thymocytes), V1 ($V_\gamma1.1^+$), V6 ($V_\gamma1.1^+V_\delta6.3^+$) and V5 ($V_\gamma5^+$, intraepithelial lymphocytes, IELs) thymocyte subsets were compared in relation to other thymic subsets of the T cell lineage profiled within ImmGen. The paucity of mature (mat, CD24^{lo}) V5 thymocytes prohibited their analysis. In addition, we established gene expression profiles of three immature (imm, CD24^{hi}) fetal γδ cell subsets: V3 ($V_\gamma3^+$, dendritic epidermal T cells, DETCs), V4 ($V_\gamma4^+$) and V2 ($V_\gamma2^+$).

Pair-wise comparisons showed that the earliest identifiable adult immV2 subset, the largest population within the γδ T cell lineage, was distinct in its global gene expression profile compared to all other adult γδ T cell subsets (Fig. 1). At an arbitrary threshold of 2-fold change in expression value, the number of genes distinguishing the immV2 subset from other immature γδ subsets was roughly equivalent to the difference observed between the immV2 subset and αβ lineage CD8$^+$ thymocytes (Fig. 1). To contextualize this finding, differentially expressed genes between immature αβ lineage DP (CD4$^+$CD8$^+$CD69$^-$) and MHC-selected DP69$^+$ cells relative to total γδTCR$^+$ thymocytes numbered ~1600 and ~900, respectively. Within theαβ T cell lineage, FOXP3$^+$ regulatory versus conventional CD4$^+$ T cells in the spleen differed in the expression of ~400genes (Fig. 1). The distinct gene expression signature of immV2 cells was in sharp contrast to the similarity between V1, V5 and V6 cells, which differed at most in the expression of 54 genes. However, the higher expression of the NKT lineage marker PLZF (encoded by *Zbtb16*) and the IEL-enriched Granzyme A [15] in immV6 and immV5 cells, respectively, suggests that despite their similarity, they were already marked with the molecular features of their fully differentiated peripheral counterparts (Fig. 2a, **data not shown**). Confirmation of the relative amounts of protein products of select differentially expressed transcription factors (TFs) in adult γδ T cell subsets is shown (Fig. 2a) with confirmatory data for additional differentially expressed genes (Supplementary Fig. 1), which captures the impressive heterogeneity in the phenotype of γδ subtypes. Fetal immV2 and ImmV4 cells (the fetal subset programmed for IL-17 production [11, 16]) were highly similar in global gene expression profiles, correlating with their common Tγδ17 effector phenotype. In contrast, fetal immV3 cells, the DETC precursors, were distinct from other fetal γδ T cell subsets, consistent with their known distinct developmental origin and requirements [17, 18]. Thus, the γδ T cell lineage can be divided into three distinct subtypes based on unique gene expression profiles at the immature developmental stage: V3, V4/V2 and V1/V6/V5.

Lineage relatedness among diverse subsets can be visualized by principal component analysis (PCA), a mathematical transformation that reduces the dimensionality of gene expression data to reveal the main components of variability in comparators. We focused on adult γδ subsets as the data set is complete. We first performed PCA to compare the gene expression profiles of non-segregated total γδTCR$^+$ thymocytes to three αβ thymocyte subsets ranging in maturity and lineages using the 15% most differentially expressed genes in the comparators, based on expression levels and variance across the populations (Fig. 2b, see On-line methods). Matureαβ thymocyte subsets were positioned together, increasingly

displaced from precursor subsets (ISP and DP) as a function of their relative maturity. Innate-like invariant αβNKT cells (*i*NKT) 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). Similar results were obtained when subset relatedness was measured by Euclidean distance and Pearson's correlation coefficients on the 15% most differentially regulated genes used for PCA (Supplementary Fig. 2a, b). When the analysis was performed using the segregatedγδ T cell subsets, however, it was clear that the immV2 subset was distinct from all other γδ thymocyte subsets, positioned uniquely in relation to other γδ and αβ subsets. (Fig. 2c, Supplementary Fig. 2c, d). Second, using genes differentially expressed in immV2 subsets compared to other immature γδ cell subsets (1006 genes, Supplementary Fig. 3a, Supplementary Tables 1, 2), PCA, Euclidian distances, and correlation scores supported a close similarity between immV2 and αβ DP thymocytes, which shared >50% of the variance and segregated distantly from T precursor subsets (Fig. 3a, Supplementary Fig. 3b, c). When DP populations were segregated, PCA indicated that immV2 cells were most similar to $CD69^{+}$DP cells (Supplementary Fig. 3d). These results revealed that immV2 cells distinguished by germline-encoded TCR domains are radically different from other immature γδ T cells, and conversely, that γδ T cells expressing the $V_{\gamma}2$-$C_{\gamma}1$ TCR, despite the extensive TCR clonal diversity generated from their TCRδ partners and VJ/VDJ junctional diversity, nevertheless represent a unique cell type.

Distinct regulatory circuitsin immature subtypes

About one-third of the genes expressed at lower levels in immV2 cells versus V1/V6 cells encode for proteins involved in DNA synthesis, RNA biogenesis, oxidative phosphorylation and protein translation (~270 genes), suggestive of decreased metabolism and energy production in immV2 cells (Supplementary Tables 1, 3, 4) compared to other γδ subsets. The expression of metabolic genes in immV2 cells corresponded to the pattern seen in immature αβ lineage DP cells, providing the primary 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 BrdU similarly, had comparable frequencies of cells expressing Ki-67, a marker of non-G0 state of cell cycle (Supplementary Fig. 4a, b), and Annexin V, a marker of apoptotic cells (data not shown). These results suggest that immV2 cells develop distinctly with unique maturational transitions.

Lineage specific TFs propel the lineage commitment process by turning on lineage-associated genes and turning off lineage-mismatched genes. Distinct programming of the dominant TFs in subset-specific effector function was already evident in emerging γδ T cell subsets. Hierarchical clustering analysis of TFs differentially expressed in immV2 and immV1/V5/V6 cells indicated that most TFs were expressed at strikingly higher levels in the immV2 subset, the most prominent belonging to the SOX-TCF1-TOX High Mobility Group (HMG) box TF family (Fig. 3c and Supplementary Tables 2, 5, 6). Among immune cell types, *Sox13* expression is restricted to developing γδ T cells. It interacts with the nuclear effectors of WNT signaling, TCF1 (encoded by *Tcf7*) and LEF1, to establish, in part, the pan-γδ lineage gene expression signature [10]. Conversely, expression of *Id2* and *Id3*, the negative regulators of E Helix-Loop-Helix TFs (E47 (encoded by *Tcfe2a*) and HEB

(encoded by *Tcf12*)), along with their upstream regulator *Egr2* were elevated in immV1/V5/V6 relative to immV2 cells [19, 20]. Further, a negative regulator of NK cell development *Bcl11b* [21] was decreased in expression in immV1/V5/V6 relative to immV2 cells (Fig. 3c, Supplementary Table 6).

In the periphery, V2 $CCR6^{+}CD27^{-}$ γδ T cells are biased to produce IL-17 while V1 $CD27^{+}$ T cells secrete IFN-γ [8, 14]. We further verified this effector pattern during the γδ T cell response to *Mycobacterium tuberculosis* infection [7]. Fourteen days post infection in the lung, all γδ subsets synthesized IFN-γ, but V2 and V4 cells were the predominant IL-17 producers (Supplementary Fig. 4c–f). The TFs *Rorc*, *Eomes* and *Zbtb16* regulate the expression of IL-17, IFN-γ and dual IL-4-IFN-γ, respectively, in αβ T cells [22–25]. Emergent immature V2, V1/V5 and V6 thymic subsets were precociously distinguished by selective expression of these three TF genes, correlating with their peripheral γδ T cell subset-specific functions (Fig. 2a, 3c). As further support for the thymic programming of immV2 cells, the genes encoding the three markers of Tγδ17 cells, B lymphocyte kinase (BLK [26] and Supplementary Fig. 4f) and the scavenger receptors SCART1 and SCART2 [27], were elevated in expression in immV2 cells compared to other subsets (Supplementary Table 2, 5, 6). Other factors required for αβ T_H effector differentiation, including *Maf* [28], *Gata3* [29, 30] and *Runx3* [31], were also expressed more highly in immV2 cells relative to other γδ subsets (Fig. 2, 3c). These results demonstrate that prior to or upon expression of different TCRs, γδ thymocyte subsets are regulated by divergent TF 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. The gene expression profiles of mature ($CD24^{lo}$) thymic γδ T cell subsets (V2, V1 and V6) were determined to examine the duration and impact 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 regulatedupon maturation of all γδ T cells (Supplementary Fig. 5a, b, Supplementary Table 7). PCA, hierarchical clustering, Euclidian distances and correlation scores of these γδ maturation-dependent genes in other thymocyte subsets clearly 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, irrespective of T cell lineage (Fig. 4a, b, Supplementary Fig. 5c, d). Of the 495 γδ maturation genes identified, ~78% were coordinately regulated upon maturation of αβ T cells subsets (matCD4, matCD8, or *i*NKT vs. DPs), while ~22% were uniquely modulated during γδ T cell maturation only (data not shown). These results clearly establish the loss of CD24 as a demarcation of the transition to maturity in the γδ T cell lineage.

There were four overriding themes that characterized γδ subset-specific maturation. First, the expression pattern of metabolic/RNA-DNA biogenesis genes in matV2 cells and matV1 cells converged, while matV6 cells remained distinct (Fig. 4c and Supplementary Table 8). Second, upon maturation, V2 cells down-regulated subset-specific TFs to the level associated with immV1 cells (Fig. 4d, Supplementary Table 9). Third, mature γδ T cell

subsets modulated expression of proteins directly involved in the execution of their peripheral subset-specific effector functions, including chemokine and cytokine receptors. Fourth, V1 and V6 cells diverged upon transit to the mature state, with the latter acquiring gene expression profile of PLZF^{+}αβ *i*NKT cells. The latter two features of γδ T cell maturation are discussed in greater detail.

The transition to the mature stage endows adult γδ subsets with unique homing properties as evidenced by the acquisition and loss of chemokine and integrin receptors. There were two patterns of chemokine receptor expression: those expressed in immature subsets and then extinguished or downregulated as γδ T cells mature (highlighted in red), and those induced precipitously during maturation and maintained in fully differentiated subsets (blue) (Fig. 5a, Supplementary Table 10). The expression pattern of CCR10 was confirmed by analysis of *Ccr10-Gfp* reporter mice, with the highest frequency of CCR10^{+} immature thymocytes found in the V6 subset (Supplementary Fig. 1). CCR6 was strongly and uniquely upregulated in the V2 subset upon transit to the mature stage, whereas CXCR3 induction was associated with V6 and V1 subset maturation (Fig. 5a, Supplementary Fig. 1, 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 non-lymphoid tissues. Interestingly, matV2 cells are *Itgb7*^{hi}*Cxcr6*^{hi}CCR6^{+}RORγt^{+}IL-7R^{hi}, reminiscent of fetal liver-derived RORγt^{+} CD3^{-} innate lymphoid cells (ILCs), some with the capacity to secrete IL-17 or IL-22 [34]. The dynamic subset-specific alterations in the tissue tropism upon maturation further underscore the separation of V2 cells from other γδ T cells and suggest that the Tγδ17 differentiation program may overlap with that of TCR negative ILCs.

Maturation stage-dependent alterations in cytokine receptor expression were also γδ T cell subset-specific and endow γδ subsets with the capacity to sense environmental cues that fully engage their programmed effector functions. Upon thymic maturation, V2 cells entered the effector-poised phase by abrupt super-induction of cytokine receptor genes that are dedicated for IL-17 production and responsiveness: *Il23r* (the most highly induced gene, ~100 fold increase in CD24^{lo} relative to CD24^{hi}), *Il17re, Il1r1, IL-18r*[35] and *Tnfrsf25* [36] were the strongest induced genes in matV2 cells (Fig. 5b, Supplementary Table 7, 11). This super-induction was accompanied by the enhanced expression of RORγt and the signaling molecule BLK that is dedicated to IL-17 production (Fig. 5c). Receptors for four counter-regulators or modulators of IL-17 production were divergently expressed upon maturation, with those encoding IL-2 and IL-27 expressed higher in matV2 cells, whereas *Il17rb* (part of the IL-25R) and *Il12rb2* (part of the IL-12R) were expressed at 1/10 and the amounts observed, respectively, in immature cells (Fig. 5b, 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. V1 cells mainly upregulated *Il18r* upon maturation and this pattern was shared among the mature γδ T cell subsets. *Il17rb* was selectively increased in expression in V6 cells, while depressed in V1 and V2 cells, upon maturation (Fig. 5b, Supplementary Table 11). IL-25 drives the expansion of Th2 cytokine (IL-5 and IL-13) producing innate lymphocytes, includingαβNKT and TCR^{-} Type II cells [34]. The selective induction of *Il17rb*, which encodes part of the IL-25R, on IL-4

producing V6 cells indicates that, like *Il23r* acquisition 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 mature, effector cytokine gene expression was also specifically initiated at the mature stage. The best illustration of this pattern was observed in matV6 cells expressing high levels of PLZF that correlateed with increased *Il4* transcription (Supplementary Fig. 5e) and matV2 cells which produced IL-17A (Fig. 5c).

In sum, the unfolding gene programs that dictate γδ subset-specific effector functions occurs in two steps: the establishment phase at the immature stage and the primed phase at the mature stage that equips the thymus-exiting γδ thymocytes with tissue migratory cues as well as sensors that once engaged at tissues sites will rapidly induce effector molecules (see summary model, Supplementary Fig. 6a). The phase transition is accompanied by an overall dampening in the activities of TFs that established the effector gene circuits as γδ thymocytes first arose. This overall pattern of dynamic TF expression program fits well with the established gene regulatory networks controlling cell differentiation in other systems [37].

MatV6 cells acquire the *i*NKT signature

MatV6 cells were predicted to overlap with αβ *i*NKT cells at the gene expression level given their shared expression of PLZF and similar functional properties. To determine the extent of similarity, we first derived the αβNKT cell lineage gene signature by identifying genes differentially expressed in thymic *i*NKT cells versus other mature conventional αβ thymocyte subsets (Fig. 6a, b, Supplementary Table 12). PCA, Euclidian distances, and correlation scores of the 540 *i*NKT signature genes showed that matV6 cells shared the expression of this signature, as matV6 and *i*NKT populations were nearly identical intheir placement along PC1 and PC2 (Fig. 6c, Supplementary Fig. 6b, c). This indicates that the acquisition of the *i*NKT gene signature is independent of TCR type (αβ versus γδ). Genetic studies also supported the relatedness of V6 and αβ *i*NKT cells: *in vivo*, the loss of *Zbtb16*, *Sh2d1a* (encodes for SAP), *Itk*, *Id3*, *Ctsl* (Cathepsin L), and *Cd74* (the invariant chain) negatively impact the development of αβ *i*NKT cells [38]. The same mutations also specifically impacted V6 cells, but not all led to decreased production of V6 cells ([39, 40] and Fig. 6d, e). Together, these results indicate that common gene networks and cellular processes regulate the generation of γδ V6 and αβ *i*NKT cells. Hence, it is either the molecular programming by similar TF networks or a similarity in TCR signaling features independent of TCR type that primarily specifies effector cell fate.

DISCUSSION

Systematic global gene expression profiling of *ex vivo*γδTCR$^+$ thymocyte subsets distinguished by TCRγ/δ repertoire demonstrated that the precocious divergence in gene expression programs of emergent γδ cell subsets underpins innate T cell effector diversification. We have identified three distinct cell types in the γδ T cell lineage based on TCR repertoire and shared gene expression profiles: Tγδ17 cells consisting of V2 and fetal V4 cells; fetal V3 cells [17, 18]; and adult V1/V5/V6 cells. The only two remaining γδ T cell subsets based on TCR repertoire in adult thymus, $V_\gamma1.2^+$ and $V_\gamma4^+$ cells that make up the

residual ~5% of total γδTCR$^+$ thymocytes, could not be directly examined as antibodies to sort these cells are lacking. However, analysis of adult immature V_γ1.1, V_γ2 and V_γ5 negative γδ thymocytes (consisting of ~85% V_γ1.2$^+$ and 15% V_γ4$^+$ cells, estimates based on the observed functional *Tcrg* gene rearrangement frequencies [41]) suggests that immV_γ1.2$^+$ cells resemble immV1/V5/V6 thymocytes in gene expression (data not shown).

The biological processes governing the generation of these three distinct immature γδ cell subsets are unknown, but the list of possibilities can now be narrowed down. αβ TCR$^+$ thymocytes do not appreciably impact γδ T effector subset differentiation *in trans* (Narayan *et al.*, manuscript submitted). The most striking gene modules distinguishing the emergent immature γδ thymocyte subsets consist of TFs. Distinct responsiveness of precursors to cytokines and growth factors that control the expression of the TFs defining effector subtypesis one possible way to generate programmed effectors. The HMG TF SOX13, its interacting partners TCF1-LEF1, and their gene targets are one central gene circuit involved in adult Tγδ17 cell differentiation, as *Sox13*$^{-/-}$ mice lack V2 Tγδ17 cells (Malhotra *et al.*, manuscript in preparation). Given the modulatory effect of WNT signaling on TCF1-LEF1, canonical WNT ligands are another potential regulator of γδ effector diversification. Alternatively, genomically encoded TCR Vγ chains may transmit distinct signals, either by unique interactions with selecting ligand(s) [13] or based on intrinsic differences in TCR conformation and/or membrane localization, independent of selecting ligands presented *in trans* [42]. For αβ T cells, the preTCR-mediated transition to the immDP stage occurs in a ligand-independent manner. Similarly, there is an *in vitro* precedent that the V_γ2V_δ5TCR can also signal in a ligand-independent, TCR dimerization-dependent manner [13]. Co-segregation of V2 cells with αβ lineage DP cells for the gene cluster involved in cellular metabolism support unique TCR signaling property as one potential discriminator of effector subset diversification.

Lastly, it is possible that the distinct timing of generation different VγTCRs during precursor maturation leads to alternate cell fates by inheritance and fixation of distinct precursor gene expression programs. The most clear-cut example of this concept is the programmed V_γ3 gene rearrangement exclusively in early fetal T cell progenitors that generates DETCs [18]. It has also been well documented that the timing of TCR expression during precursor maturation, rather than TCR type, is critical for defining resultant T cell properties [43, 44]. That γδ T cell subsets may originate at different stages of precursor maturation is supported by several studies indicating that the three functional Cγ loci are regulated distinctly and independently [45, 46]. Most directly, we have found that the earliest c-Kit$^+$ T cell progenitors were skewed toward the generation of CCR6$^+$ V2 cells in OP9-DL1 cultures (Malhotra *et al.*, manuscript in preparation), suggesting that the differing onset of functional TCRγ chain expression with the attendant inheritance of distinct collective gene activities of the precursors may underpin the molecular heterogeneity of γδ subsets.

Identification of gene networks specifying effector functions of innate γδ T cell subsets established at the earliest stage of γδ thymocyte differentiation and fixed during thymic maturation suggests that γδ T cells are unlikely to respond homogeneously to external cues. Discovery of the upstream regulators of these gene networks will identify the mechanism of thymic programming of innate γδ effector subset diversification. Moreover, developmentally

hard-wired programs to produce IL-17 family cytokines by other recently discovered innate lymphocyte subsets in the gut are predicted to utilize genetic networks similar to those operational in Tγδ17 cells [34, 47], providing a possible route to the identification of a unified origin of innate lymphoid effector subsets.

METHODS

Mice

5 week-old male C57BL/6J mice (Jackson Labs) were used for microarray analysis within 1 week of arrival. *Ccr10-Gfp* (N. Xiong, Penn State) and *Il4-Gfp* reporter (M. Mohrs, Trudeau Institute) mice have been described previously. *Axin2* (H. Birchmeier, MDC Berlin), *Cenpk* (SOLT, IGTC), and *Sox5* (V. Lefebvre, Cleveland Clinic) *LacZ* reporter mice have been published. *Ctsl*$^{-/-}$ mice were provided by Kenneth Rock (UMMS, Worcester, MA) and *Cd74*$^{-/-}$ mice were provided by Eric Huseby (UMMS, Worcester, MA). Mice were housed in a specific pathogen free rodent barrier facility. Experiments were approved by the UMMS IACUC (Worcester, MA).

Sample preparation for microarray analysis

Pooled thymocytes from 4–30 mice were enriched for γδ T cells by depletion of CD8$^+$ cells using magnetic beads and an autoMACS, stained, and sorted directly into Trizol (Invitorgen) using a FACSAria (~2–3 × 10^4 cells, >99% pure). Independent triplicates were sorted unless noted otherwise (complete sorting details at Immgen.org). As antibodies to V$_\gamma$4 are not available, fetal immV4 cells were sorted by gating on TCRδ$^+$ cells that were V$_\gamma$1.1$^-$V$_\gamma$2$^-$V$_\gamma$3$^-$V$_\gamma$5$^-$ (estimated purity ~95%). Approximate frequencies of TCRδ chains associated with sorted TCRγ subsets are as follows: V2 (V$_\gamma$2$^+$, 50% V$_\delta$4$^+$, 40% V$_\delta$5$^+$, all V$_\delta$6.3$^-$, ~45% of total γδ cells); V1 (V$_\gamma$1.1$^+$, diverse Vδs including 25% V$_\delta$4$^+$, 15% V$_\delta$5$^+$, others at lower frequencies, all V$_\delta$6.3$^-$, 30% of total γδ cells); V6 (V$_\gamma$1.1$^+$, 100% V$_\delta$6.3$^+$, ~8% of total γδ cells); V5 (V$_\gamma$5$^+$, 40% V$_\delta$5$^+$ and various others at lower frequencies, ~5% of total γδ cells); and fetal V3 and V4 thymocytes co-express V$_\delta$1.

Data analysis and visualization

RNA processing and microarray analysis using the Affymetrix MoGene 1.0 ST array was performed at the ImmGen processing center (SOP at Immgen.org). Data analysis was performed using GenePattern (Genepattern.org) analysis modules. Raw data were RMA normalized with quantile normalization and background correction (ExpressionFileCreator). The ConsolidateProbeSets module (Scott Davis, Harvard Medical School, Boston, MA) was used to consolidate multiple probe sets into a single mean probe set value for each gene. Identification of differentially regulated genes was performed using Multiplot. Unless otherwise indicated, genes were considered differentially regulated if they differed in expression by more than 2-fold, had a coefficient of variation (cv) among replicates of less than 0.5, had a *t*-test *P* value of less than 0.05, and had a mean expression value (MEV) of greater than 120 in at least one subset in the comparison. Heatmaps were generated by hierarchical clustering (HierarchicalClustering module) of data based on gene (row) and subset (column) using the Pearson correlation for distance measurement. Data were log transformed and clustered using pair-wise complete linkage. Data were row centered prior to

visualization using the HeatmapViewer module. Principal component analysis was performed using the PopulationDistances PCA program (Scott Davis, Harvard Medical School, Boston, MA). Where indicated, the PCA program was used to identify the 15% most differentially expressed genes among subsets by filtering based on a variation of ANOVA analysis using the geometric standard deviation of populations to weight genes that vary in multiple populations. Data were log transformed, gene and subset normalized, and filtered for genes that had a MEV>120 prior to visualization. Euclidian distance and Pearson's correlation coefficients were calculated using the "dist" and "cor" commands in R to generate distance and correlation matrixes for a given set of genes and subsets. Heatmaps of Euclidian distance and Pearson's correlation coefficients were generated by hierarchical clustering (HierarchicalClustering module, Pearson's correlation for row and column, complete-linkage). Data were visualized using a global color scheme in the HeatmapViewer module. Pathway analysis was performed using Ingenuity software (Ingenuity.com) and by manual inspection. Some functional classifications were performed using AmiGO (Amigo.geneontology.org) and KEGG pathways (Genome.kp.kegg).

Flow cytometry

The following cell surface antibodies were purchased from BD Biosciences: CD25 (PC61), CD27 (LG.3A10), $V_\gamma 2$ (UC3-10A6), $V_\gamma 3$ (536), $V_\delta 6.3$ (8F4H7B7), Ly49C/I (5E6), CCR6 (140706), Thy1.2 (53-2.1), CD9 (KMC8), streptavidin APC; eBioscience: CD3 (145-2C11), CD4 (RM4-5), CD8 (53-6.7), CD44 (IM7), CD122 (5H4), CD127 (IL-7Rα A7R34), CD24 (HSA, M1/69), c-Kit (2B8), TCRδ (GL3), NKG2D (CX5), CD62L (MEL-15), PD1 (J43), CD73 (TY/11.8), CD199 (CCR9, CW-1.2), CD197 (CCR7, 4B12), CD183 (CXCR3, 174), CD184 (CXCR4, 2B11), CD244.2 (2B4, 244F4), NKG2ACE (20D5), GITR (DTA-1), CD48 (HM48-1), CD28 (37.51), streptavidin PE-Cy7; or other vendors: IL-6R (D7715A7, BioLegend), CD119 (IFNγrα, RDI/Fitzgerald Industries). $V_\gamma 1.1$ Ab was purified by Bio-XCell and biotinylated using the FluoReporter Mini-Biotin-XX Labeling Kit (Invitrogen). $V_\delta 1$ (17D1) antibody was provided by A.Hayday (King's College, London). Intracellular proteins (IL-17, 17B7, eBioscience; IFN-γ, XMG1.2, BD Biosciences) were detected using the Cytofix/Cytoperm kit (BD Biosciences). Intranuclear staining was performed using the FOXP3 staining kit (eBioscience) for the following antibodies: TOX (TXRX10), EOMES (Dan11mag), RORγt (AFKJS-9), GATA-3 (L50-823), from eBioscience; BLK (cat# 3262), LEF-1 (C12A5), from Cell Signaling; Ki-67 from BD Biosciences; PLZF (D-9) and SMO (N-19), from Santa Cruz. FDG (Invitrogen) staining was performed according to standard protocols. Data were acquired on a BD LSRII cytometer and analyzed using FlowJo (Treestar). When indicated, data were concatenated from multiple independent samples in FlowJo for visualization in histograms.

Ex vivo stimulation

Ex vivo stimulations were performed by culturing total cells (2×10^6/well) with PMA (10ng/ml) and Ionomycin (1ug/ml) for 4 hours at 37°C, with Golgi Stop and Golgi Plug (BD Biosciences) added after 1 hour, according to the manufacturers' protocols.

Mtb infection and analysis

WT *Mtb* bacterial strains were used to infect 4–10 male 6–8 week old C57BL/6 mice (Jackson Labs) with 2×10^3 CFU of *Mtb* by the aerosol route. Two weeks post infection, mice were sacrificed 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 differentially expressed genes, *t*-test *P* values were generated using Multiplot (Genepattern). For statistical analysis of flow cytometry data, Prism (GraphPad Software) was used. Data were tested for normality using F tests and then analyzed using unpaired two-tailed *t*-tests. Pathway analysis was performed using Ingenuity Pathway Analysis and statistical significance was determined using the program's built-in Fisher's exact test.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank the members of the ImmGen Consortium for discussions, the ImmGen core team (M. Painter, J. Ericson, 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 grant DK32520 were used. Supportedby R24 AI072073 from NIH/NIAID to the ImmGen group, and CA100382 to J.K.

References

1. Heng TS, Painter MW. The Immunological Genome Project: networks of gene expression in immune cells. Nat Immunol. 2008; 9:1091–1094. [PubMed: 18800157]
2. Ferrick DA, 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. 1995; 373:255–257. [PubMed: 7816142]
3. Stark MA, et al. Phagocytosis of apoptotic neutrophils regulates granulopoiesis via IL-23 and IL-17. Immunity. 2005; 22:285–294. [PubMed: 15780986]
4. Hayday A, Tigelaar R. Immunoregulation in the tissues by γδ T cells. Nature Rev Immunol. 2003; 3:233–242. [PubMed: 12658271]
5. Berg LJ. Signalling through TEC kinases regulates conventional versus innate CD8(+) T-cell development. Nature Rev Immunol. 2007; 7:479–485. [PubMed: 17479128]
6. Lee YJ, Jameson SC, Hogquist KA. Alternative memory in the CD8 T cell lineage. Trends Immunol. 2011; 32:50–56. [PubMed: 21288770]
7. Lockhart E, Green AM, Flynn JL. IL-17 production is dominated by γδ T cells rather than CD4 T cells during Mycobacterium tuberculosis infection. J Immunol. 2006; 177:4662–4669. [PubMed: 16982905]
8. Martin B, Hirota K, Cua DJ, Stockinger B, Veldhoen M. Interleukin-17-producingγδ T cells selectively expand in response to pathogen products and environmental signals. Immunity. 2009; 31:321–330. [PubMed: 19682928]
9. Sutton CE, et al. Interleukin-1 and IL-23 induce innate IL-17 production from γδ T cells, amplifying Th17 responses and autoimmunity. Immunity. 2009; 31:331–341. [PubMed: 19682929]
10. Melichar HJ, et al. Regulation of γδ versus αβ T lymphocyte differentiation by the transcription factor SOX13. Science. 2007; 315:230–233. [PubMed: 17218525]
11. O'Brien RL, Born WK. γδ T cell subsets: a link between TCR and function? Semin Immunol. 2010; 22:193–198. [PubMed: 20451408]

12. Azuara V, Levraud JP, Lembezat MP, 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. 1997; 27:544–553. [PubMed: 9045929]
13. Jensen KD, et al. Thymic selection determines γδ T cell effector fate: antigen-naive cells make interleukin-17 and antigen-experienced cells make interferon gamma. Immunity. 2008; 29:90–100. [PubMed: 18585064]
14. Ribot JC, et al. CD27 is a thymic determinant of the balance between Interferon-γ- and interleukin 17-producing gammadelta T cell subsets. Nat Immunol. 2009; 10:427–436. [PubMed: 19270712]
15. Shires J, Theodoridis E, Hayday AC. Biological insights into TCRγδ+ and TCRαβ+ intraepithelial lymphocytes provided by serial analysis of gene expression (SAGE). Immunity. 2001; 15:419–434. [PubMed: 11567632]
16. Shibata K, et al. Identification of CD25+ γδ T cells as fetal thymus-derived naturally occurring IL-17 producers. J Immunol. 2008; 181:5940–5947. [PubMed: 18941182]
17. Ikuta K, et al. A developmental switch in thymic lymphocyte maturation potential occurs at the level of hematopoietic stem cells. Cell. 1990; 62:863–874. [PubMed: 1975515]
18. Xiong N, Kang C, Raulet DH. Positive selection of dendritic epidermal γδ T cell precursors in the fetal thymus determines expression of skin-homing receptors. Immunity. 2004; 21:121–131. [PubMed: 15345225]
19. Bain G, et al. Regulation of the helix-loop-helix proteins, E2A and Id3, by the Ras-ERK MAPK cascade. Nat Immunol. 2001; 2:165–171. [PubMed: 11175815]
20. Haks MC, et al. Attenuation of γδTCR signaling efficiently diverts thymocytes to the αβ lineage. Immunity. 2005; 22:595–606. [PubMed: 15894277]
21. Rothenberg EV, Zhang J, Li L. Multilayered specification of the T-cell lineage fate. Immuno Rev. 2010; 238:150–168.
22. Ivanov II, et al. The orphan nuclear receptor RORγt directs the differentiation program of proinflammatory IL-17+ T helper cells. Cell. 2006; 126:1121–1133. [PubMed: 16990136]
23. Pearce EL, et al. Control of effector CD8+ T cell function by the transcription factor Eomesodermin. Science. 2003; 302:1041–1043. [PubMed: 14605368]
24. Savage AK, et al. The transcription factor PLZF directs the effector program of the NKT cell lineage. Immunity. 2008; 29:391–403. [PubMed: 18703361]
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. 2008; 9:1055–1064. [PubMed: 18660811]
26. Laird RM, Laky K, Hayes SM. Unexpected role for the B cell-specific Src family kinase B lymphoid kinase in the development of IL-17-producing gammadelta T cells. J Immunol. 2010; 185:6518–6527. [PubMed: 20974990]
27. Kisielow J, Kopf M, Karjalainen K. SCART scavenger receptors identify a novel subset of adult γδ T cells. J Immunol. 2008; 181:1710–1716. [PubMed: 18641307]
28. Bauquet AT, 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. 2009; 10:167–175. [PubMed: 19098919]
29. Zheng W, Flavell RA. The transcription factor GATA-3 is necessary and sufficient for Th2 cytokine gene expression in CD4 T cells. Cell. 1997; 89:587–596. [PubMed: 9160750]
30. Maruyama T, et al. Control of the differentiation of regulatory T cells and T(H)17 cells by the DNA-binding inhibitor Id3. Nat Immunol. 2010; 12:86–95. [PubMed: 21131965]
31. Djuretic IM, et al. Transcription factors T-bet and Runx3 cooperate to activate Ifng and silence Il4 in T helper type 1 cells. Nat Immunol. 2007; 8:145–153. [PubMed: 17195845]
32. Acosta-Rodriguez EV, et al. Surface phenotype and antigenic specificity of human interleukin 17-producing T helper memory cells. Nat Immunol. 2007; 8:639–646. [PubMed: 17486092]
33. Weinreich MA, et al. KLF2 transcription-factor deficiency in T cells results in unrestrained cytokine production and upregulation of bystander chemokine receptors. Immunity. 2009; 31:122–130. [PubMed: 19592277]

34. Spits H, Di Santo JP. The expanding family of innate lymphoid cells: regulators and effectors of immunity and tissue remodeling. Nat Immunol. 2011; 12:21–27. [PubMed: 21113163]
35. Andrews DM, et al. Homeostatic defects in interleukin 18-deficient mice contribute to protection against the lethal effects of endotoxin. Immunol Cell Bio. 2011; 89:739–746. [PubMed: 21263463]
36. Pappu BP, et al. TL1A-DR3 interaction regulates Th17 cell function and Th17-mediated autoimmune disease. J Exp Med. 2008; 205:1049–1062. [PubMed: 18411337]
37. Yosef N, Regev A. Impulse control: temporal dynamics in gene transcription. Cell. 2011; 144:886–896. [PubMed: 21414481]
38. Bendelac A, Savage PB, Teyton L. The biology of NKT cells. Annual Rev Immunol. 2007; 25:297–336. [PubMed: 17150027]
39. Felices M, Yin CC, Kosaka Y, Kang J, Berg LJ. Tec kinase Itk in γδ T cells is pivotal for controlling IgE production in vivo. Proc Natl Acad Sci USA. 2009; 106:8308–8313. [PubMed: 19416854]
40. Verykokakis M, et al. Inhibitor of DNA binding 3 limits development of murine slam-associated adaptor protein-dependent “innate” gammadelta T cells. PloS One. 2010; 5:e9303. [PubMed: 20174563]
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. 2004; 173:3261–3270. [PubMed: 15322188]
42. Kuhns MS, Davis MM. Disruption of extracellular interactions impairs T cell receptor-CD3 complex stability and signaling. Immunity. 2007; 26:357–369. [PubMed: 17368054]
43. Bruno L, Fehling HJ, von Boehmer H. The ab T cell receptor can replace the γδ receptor in the development of gd lineage cells. Immunity. 1996; 5:343–352. [PubMed: 8885867]
44. Baldwin TA, Sandau MM, Jameson SC, Hogquist KA. The timing of TCR alpha expression critically influences T cell development and selection. J Exp Med. 2005; 202:111–121. [PubMed: 15998791]
45. Passoni L, et al. Intrathymic d selection events in γδ cell development. Immunity. 1997; 7:83–95. [PubMed: 9252122]
46. Livák F, Tourigny M, Schatz DG, Petrie HT. Characterization of TCR gene rearrangements during adult murine T cell development. J Immunol. 1999; 162:2575–2580. [PubMed: 10072498]
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. Nature Immunol. 2010; 11:945–952. [PubMed: 20818394]

Members of the Immunological Genome Project Consortium

Natasha Asinovski, Angelique Bellemare-Pelletier, Christophe Benoist, Adam Best, Michael Bevan, Natalie Bezman, David Blair, Harald von Boehmer, Milena Bogunovic, Patrick Brennan, Michael Brenner, Andrew Chow, Nadia Cohen, Jim Collins, James Costello, Scott Davis, Michael Dustin, Kutlu Elpek, Ayla Ergun, Jeff Ericson, Anne Fletcher, Emmanuel Gautier, Roi Gazit, Ananda Goldrath, Daniel Gray, Melanie Greter, Richard Hardy, Daigo Hashimoto, Kimie Hattori, Julie Helft, Tracy Heng, Jonathan Hill, Gordon Hyatt, Claudia Jakubzick, Daniel Jepson, Radu Jianu, Vladimir Jojic, Joonsoo Kang, Charlie Kim, Jamie Knell, Francis Kim, Daphne Koller, Taras Kreslavsky, David Laidlaw, Lewis Lanier, Catherine Laplace, Hu Li, Deepali Malhotra, Diane Mathis, Miriam Merad, Jennifer Miller, Paul Monach, Kavitha Narayan, Adriana Ortiz-Lopez, Henry Paik, Michio Painter, Jeremy Price, Gwendalyn Randolph, Aviv Regev, Derrick Rossi, Ravi Sachidanandam, Tal Shay, Susan Shinton, Joseph Sun, Katelyn Sylvia, Nageswara Tata, Shannon Turley, Amy Wagers, Ei Wakamatsu, Linda Wakim, Yan Zhou.

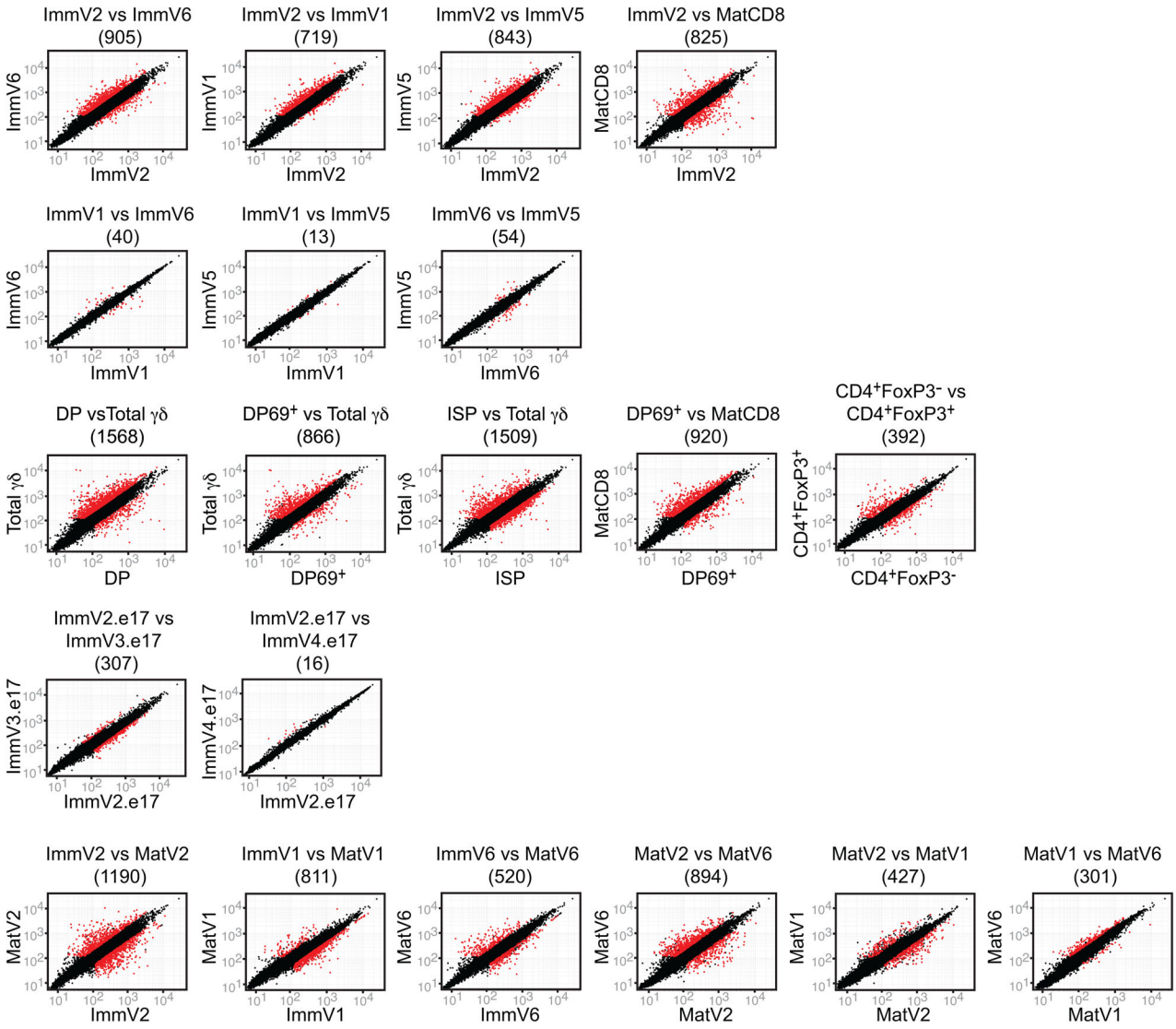

Figure 1.

Distinct global gene expression profiles of γδ cell subsets defined by TCR repertoire. The mean expression of sample replicates for consolidated probe sets was plotted to compare populations of immature adult thymocytes, immature fetal thymocytes, and mature thymocytes from C57BL/6 mice using Multiplot. Each dot represents one gene (mean of all probe sets), and dots highlighted in red represent genes whose expression is changed by greaterthan 2 fold, $P<0.05$, coefficient of variation (cv)<0.5, mean expression value (MEV)>120 in one subset. The total number of highlighted genes is listed in parenthesis at the top of each scatter plot. $CD4^+Foxp3^-$ and $CD4^+Foxp3^+$ samples were isolated from the spleen. Similar results were obtained when $CD8^+CD24^{int}$ thymocytes were compared with immV2 cells (615 genes). The following abbreviations were used that correlate with the indicated ImmGen populations: 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−); $CD69^{+}$DP=T.DP69+.Th; Totalγδ=Tgd.Th; ISP=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).

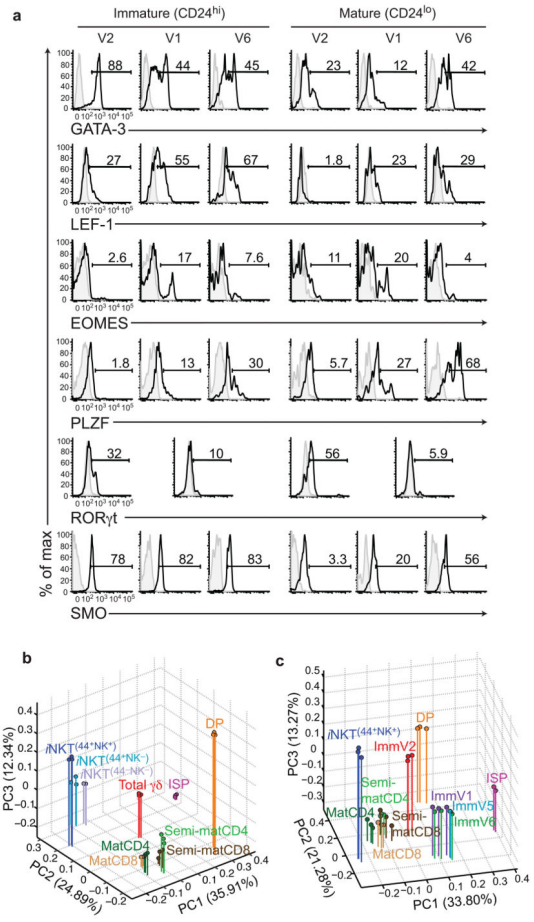

Figure 2.

Distinct TF protein expression and divergence of γδ T cell subsets. (**a**) Histograms of the expression of transcription factors in immature (CD24^{hi}) and mature (CD24^{lo}) V2, V1, and V6 γδ thymocytes are shown. Histograms of TF expression were generated by gating on total TCRδ$_+$ cells, subsetting cells based on V$_γ$2, V$_γ$1.1, and V$_δ$6.3 expression, and gating on CD24^{hi} or CD24^{lo} cells within each subset. Plots straddling V1 and V6 columns represent histograms gated on total V$_γ$1.1$^+$ cells. Isotype control staining is shown for each TF (grey histograms). For GATA-3 staining, a FACS minus one (FMO) control was used, and for RORγt, a negative control was used (TCRβ^{hi} cells that are negative for RORγt expression). SMO is shown as a representative TF expressed similarly among immature subsets and downregulated upon maturation. In some cases, gates were drawn on the “high” expressers for a given TF to best show the relative difference in expression among the γδ subsets. A minimum of 3 mice were analyzed per experiment, and a minimum 2 experiments were performed per marker. (**b**) PCA on the populations shown was performed using the 15% most variable genes (MEV>120 in at least one population, 1594 genes). The first three PCs are shown, along with the proportion of the total variability represented by each component. Expression of CD44 (44) and NK1.1 (NK) are designated for *i*NKT subsets. (**c**) PCA on the populations shown was performed using the 15% most variable genes (MEV>120 in at least one population, 1597 genes).

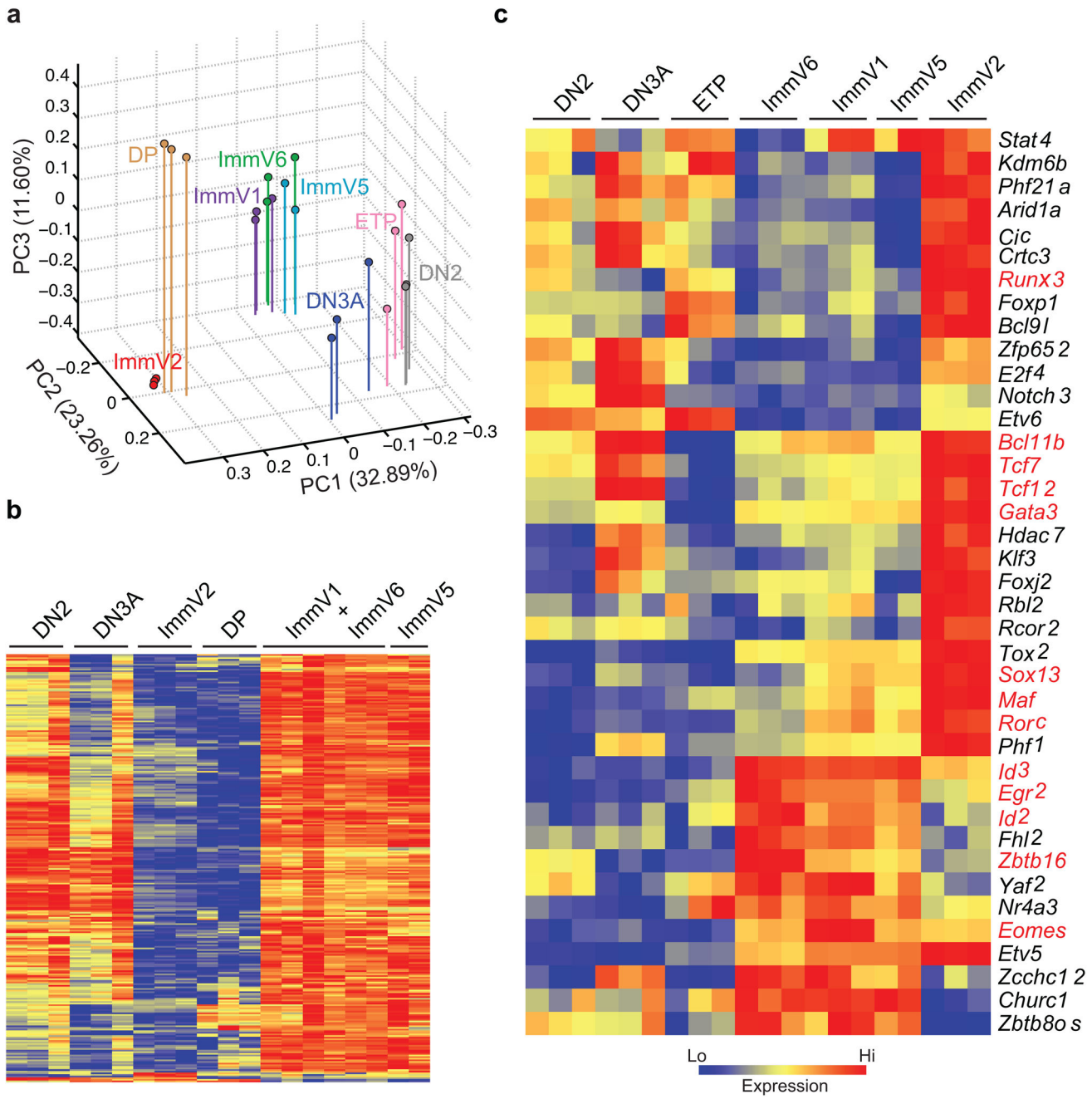

Figure 3.

Expression of TFs and metabolic genes distinguishes γδ thymocyte subsets. (**a**) PCA was performed on the populations shown using the genes that were differentially regulated among immV2 and immV1 or immV6 cells (1006 genes, see Supplementary Figure 3a, Supplementary Table 1, 2). The first three PCs are shown along with the proportion of the total variability represented by each component. (**b, c**) Heatmaps of relative expression of genes involved in metabolic processes (**b**) and of TFs (**c**) in thymocyte subsets that were differentially regulated among ImmV2 and immV1 or immV6 cells. For heatmaps, data were log transformed, gene row centered, and hierarchically clustered by gene and subset,

and genes are color coded (see legend) to display relative expression. TFs discussed in the text are in red font. In (**b**), interspersed immV1 and immV6 replicates are grouped together.

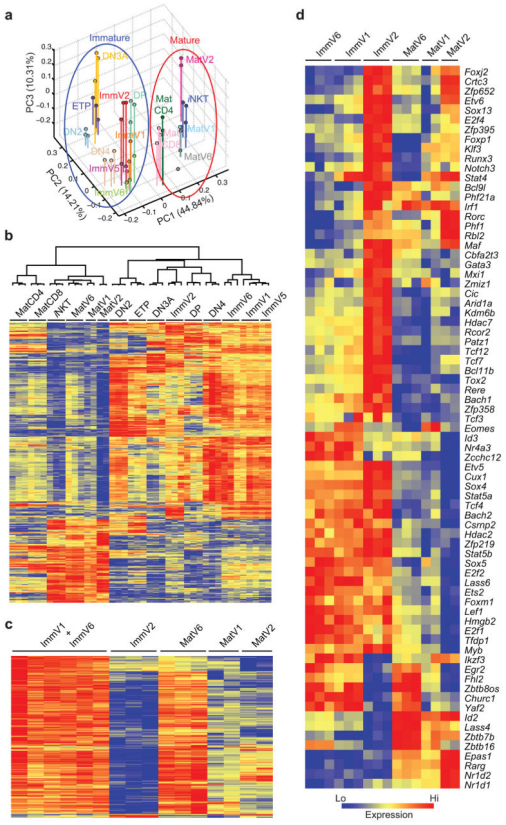

Figure 4.

Convergence of gene expression profiles of γδ subsets upon maturation. (**a**) PCA was performed on the populations shown using the 495 genes that were differentially regulated upon maturation of adult γδ T cells (see Supplementary Fig. 5, Supplementary Table 7). The first three PCs are shown along with the proportion of the total variability represented by each component. (**b**) Heatmap showing the expression of the 495 genes of the γδ maturation gene signature in precursor, αβ, and γδ T cell subsets. The dendrogram for sample clustering shows that immature and mature subsets form two distinct clusters irrespective of T cell lineage. (**c, d**) Heatmaps of relative expression of metabolic genes (**c**) and TFs (**d**) in immature and mature γδ subsets. For all heatmaps, data were log transformed, gene row centered, and hierarchically clustered by gene and subset, and genes are color coded (see legend) to display relative expression. In (**c**), interspersed immV1 and immV6 replicates are grouped together.

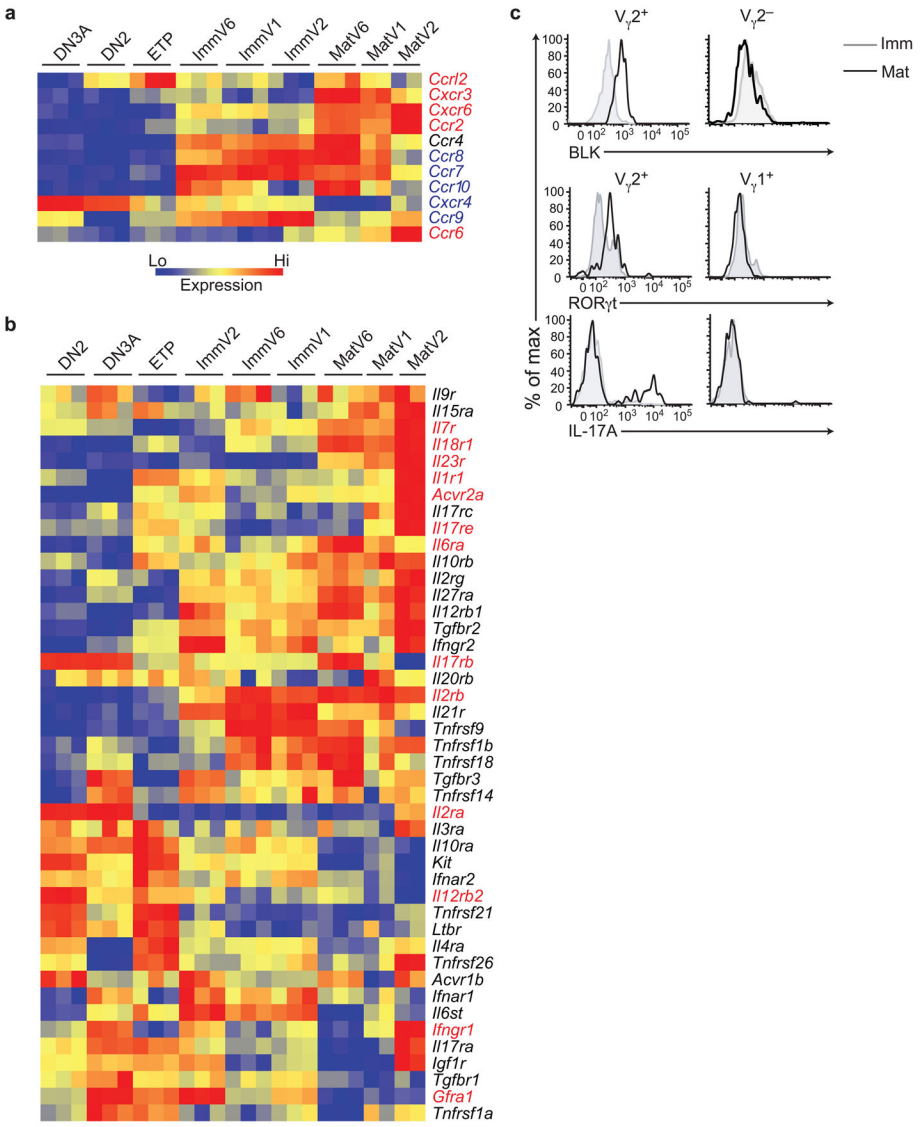

Figure 5.

Generation of mature γδ cell subsets poised for elaboration of effector function programs. (**a**) Heatmap of the relative expression of select chemokine receptors is shown. Chemokine receptors expressed in immature subsets and extinguished or downregulated as γδ T cells mature (blue) and chemokine receptors induced precipitously during maturation and maintained in fully differentiated subsets (red) are indicated. (**b**) Heatmap of relative expression of select cytokine and growth factor receptors is shown. For heatmaps, data were log transformed, gene row centered, and hierarchically clustered by gene and subset, and genes are color coded (see legend) to display relative expression. (**c**) Representative BLK, RORγt, and IL-17A expression in immature and mature γδ thymocyte subsets. Histograms were gated on total TCRδ$^+$ cells, separated into V$_\gamma$2$^+$ and V$_\gamma$2$^-$ populations, and gated on CD24^{hi} (light grey line) or CD24^{lo} (black line) to show expression of the indicated markers in overlays. BLK and IL-17A protein expression could only be discerned in mature V$_\gamma$2$^+$ thymocytes, whereas a low amount of RORγt was expressed in immV2 cells (Fig. 2). For

IL-17A expression, cells were stimulated for 4 hours with PMA-ionomycin followed by surface and intracellular cytokine staining. One of 6 experiments is shown, each with a minimum of three mice per experiment.

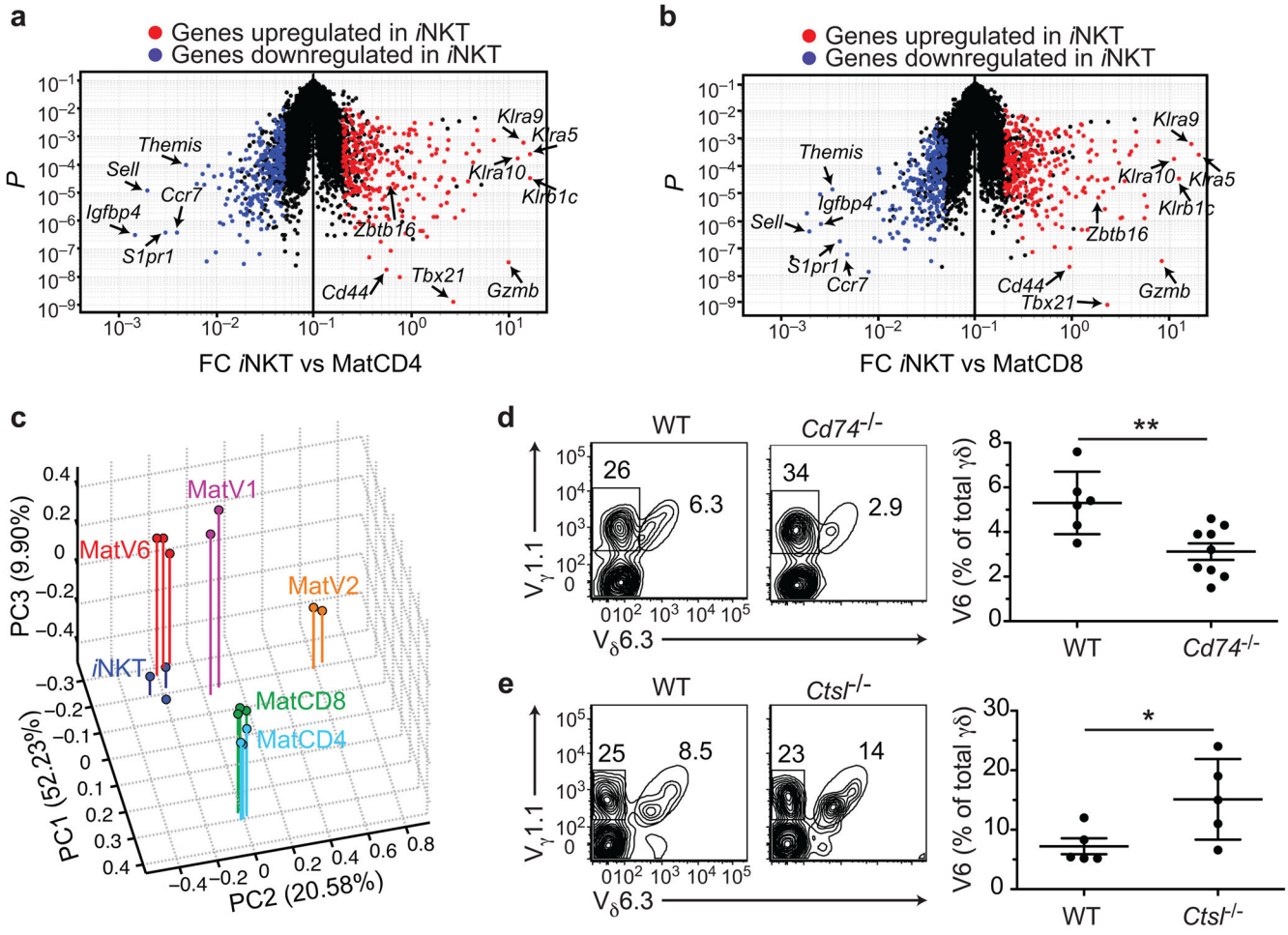

Figure 6.

Common features of αβ *i*NKT cells and γδ matV6 cells. (**a, b**) *i*NKT signature genes were identified using Multiplot based on being altered by 2-fold or more in expression in bothαβNKT versus matCD4 and αβNKT versus matCD8 comparisons, and having a cv<0.5, and MEV>120 in at least one subset. A total of 292 genes were increased in *i*NKT versus matCD4-CD8 cells, and 248 genes were decreased in *i*NKT versus matCD4-CD8 cells. Fold change (FC) versus *P*-value volcano plots of *i*NKT versus matCD4 (**a**) and *i*NKT versus matCD8 (**b**) are shown with the genes upregulated in *i*NKT cells in both comparisons in red, and the genes downregulated in *i*NKT cells in both comparisons in blue. The names and locations of select genes, including *Zbtb16* are indicated with arrows. (**c**) PCA was performed on the populations shown using the 540 genes that were differentially regulated between *i*NKT cells and matCD4 and matCD8 cells. The first three PCs are shown along with the proportion of the total variability represented by each component. (**d, e**) Representative FACS plots showing $V_\gamma 1.1$ and $V_\delta 6.3$ staining on C57BL/6 and *Cd74*$^{-/-}$ (**d**) or *Ctsl*$^{-/-}$ (**e**) γδ thymocytes (left). Frequencies of V6 cells among γδ T cells for all mice analyzed were graphed (right). Each symbol represents an individual mouse. Horizontal bars

represent the mean ±s.d. $^{**}P$=0.005 and $^{*}P$=0.04 (two-tailed Student's *t*-test). Data shown were combined from two independent experiments.