Title
Shared and distinct transcriptional programs underlie the hybrid nature of iNKT cells

Journal
Nature Immunology, 14(1)

ISSN
1529-2908

Authors
Cohen, Nadia R
Brennan, Patrick J
Shay, Tal
et al.

Publication Date
2013

DOI
10.1038/ni.2490

Peer reviewed

Published in final edited form as:
Nat Immunol. 2013 January ; 14(1): 90–99. doi:10.1038/ni.2490.

Shared and distinct transcriptional programs underlie the hybrid nature of *i*NKT cells

Nadia R. Cohen[1,5], **Patrick J. Brennan**[1,5], **Tal Shay**[2], **Gerald F. Watts**[1], **Manfred Brigl**[1,3], **Joonsoo Kang**[4], **Michael B. Brenner**[1], and **The Immunological Genome Project Consortium**

[1]Department of Medicine, Division of Rheumatology, Immunology and Allergy, Harvard Medical School, Boston, Massachusetts, USA

[2]Broad Institute of MIT and Harvard, Cambridge, Massachusetts, USA

[3]Department of Pathology, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts, USA

[4]Department of Pathology, University of Massachusetts Medical School, Worcester, Massachusetts, USA

Abstract

Invariant natural killer T (*i*NKT) cells are innate-like T lymphocytes that act as critical regulators of the immune response. To better characterize this population, we profiled *i*NKT cell gene expression during ontogeny and in peripheral subsets as part of the Immunological Genome Project (ImmGen). High-resolution comparative transcriptional analyses defined developmental and subset-specific *i*NKT cell gene expression programs. In addition, *i*NKT cells were found to share an extensive transcriptional program with natural killer (NK) cells, similar in magnitude to that shared with major histocompatibility complex (MHC)-restricted T cells. Strikingly, the NK-*i*NKT program also operated constitutively in γδT cells and in adaptive T cells following activation. Together, our findings highlight a core effector program regulated distinctly in innate and adaptive lymphocytes.

Introduction

The Immunological Genome Project (ImmGen) is a consortium of immunologists and computational biologists who aim, using rigorously standardized experimental and analysis pipelines, to generate a high-resolution, comprehensive definition of gene expression and regulatory networks in the mouse immune system[1]. In this context, we determined global gene expression profiles for thymic and peripheral invariant natural killer T (*i*NKT) cell subsets to gain insight into the *i*NKT cell transcriptional landscape, its unique features, and the relationships of *i*NKT cells to other innate and adaptive cell lineages.

*i*NKT cells are a subset of αβ T cells with an semi-invariant T cell antigen receptor (TCR) recognizing lipid antigens presented by CD1d[2]. By rapidly producing cytokines, these cells modulate both the innate and adaptive arms of the immune system, critically affecting biological processes in anti-microbial immunity, tumor rejection, and inflammation[3]. Like MHC-restricted T cells, *i*NKT cells undergo thymic differentiation with somatic

Corresponding author: Michael B. Brenner, Division of Rheumatology, Immunology and Allergy, Department of Medicine, Brigham and Women's Hospital, Smith Building, Room 552, One Jimmy Fund Way, Boston, MA 02115, Tel: 617-525-1000, Fax: 617-525-1001, mbrenner@rics.bwh.harvard.edu.
[5]These authors contributed equally to this work

recombination[4], recognize self and foreign antigens[2,3], secrete T_H1, T_H 2, and T_H17 cytokines[5], and provide help to B cells[6].

The term “NK T” was coined to reflect the expression of the natural killer (NK) cell marker NK1.1[7]. While a number of other NK receptors (NKRs) can also be expressed by *i*NKT cells[8], the validity of the term “NKT” has been called into question[9] because *i*NKT cells are developmentally more closely related to the T than the NK cell lineage, and because NKRs are neither specific to *i*NKT cells nor expressed on all *i*NKT cells.

While a central role for TCR-mediated activation in *i*NKT cell biology is clear, striking parallels in the behavior of both NK and *i*NKT cells have nevertheless emerged. The homeostatic distribution and survival requirements of *i*NKT cells are similar to those of NK cells[10,11,12,13]. Also comparable are their trafficking and activation kinetics. Both cell types constitutively express inflammatory chemokine receptors, accumulate at sites of infection within 24–72 h, and exert their effector functions without a priming requirement[14,15]. In addition, evidence suggests that *i*NKT cells, like NK cells, can use activating NKRs to sense stress-induced ligands[15–19]. *i*NKT cells also detect cellular stress via their TCRs, which are reactive to inflammation-induced alterations in CD1d-presented self-lipid antigens[20–22]. Finally, NK and *i*NKT cells engage in comparable bidirectional interactions with antigen presenting cells (APCs) during which APC-derived inflammatory cytokines potentiate NK and *i*NKT responses to surface ligands, and NK or *i*NKT cells, in return, promote APC maturation[23–28].

In this report, we shed light on the transcriptional programs operating over the course of *i*NKT cell development and in peripheral $CD4^+$ and $CD4^-$ *i*NKT cell subsets. Utilizing the ImmGen compendium, which allows direct comparison of gene expression in developing and mature *i*NKT, NK, and T cell subsets, we assess the transcriptional basis for NK- *i*NKT similarities. Our data demonstrate that shared NK- *i*NKT transcriptional programs are more extensive than currently appreciated and comparable in breadth to those shared by *i*NKT and MHC-restricted T cells. Finally, we show that the transcriptional patterns expressed constitutively by both NK and *i*NKT cells represent a core effector program also operational in other innate lymphocytes and induced in adaptive lymphocyte populations following activation.

Results

*i*NKT cell-specific developmental programs

*i*NKT cells, like other T lymphocytes, mature in the thymus, where they diverge from MHC-restricted αβ T cells at the $CD4^+$ $CD8^+$ double-positive (DP) stage. *i*NKT cells subsequently undergo sequential stages of differentiation characterized by differential expression of CD44 and NK1.1[4] (Supplementary Fig. 1). To characterize the developmental transcriptional programs in *i*NKT cells in relation to those operating in maturing adaptive αβ T cells, we profiled $CD44^-$ $NK1.1^-$ (stage 1), $CD44^+NK1.1^-$ (stage 2) and $CD44^+NK1.1^+$ (stage 3) thymic *i*NKT cells in the context of the ImmGen Project (Fig. 1a). 1850, 155 and 697 genes were differentially expressed from DP to stage 1, stage 1 to stage 2, and stage 2 to stage 3, respectively, using an arbitrary fold change (FC) threshold of FC > 2 (Fig. 1b, c). Transcripts modulated between thymic *i*NKT cell populations included a number of genes involved in *i*NKT cell maturation[4] as well as several genes of unknown function (Supplementary Tables 1–6). At the DP branch point, we identified a subset of genes modulated selectively by stage 1 *i*NKT cells but not by early stage $CD4^+$ adaptive T cells ($CD4^+8^{int}$) (Fig. 1d). Although some of these genes such as *Zbtb16* (PLZF) and *Vdr* (vitamin D receptor) are known to play important functions in *i*NKT cell lineage specification[29–31], many have not yet been linked to *i*NKT cell development

(Supplementary Tables 7, 8). At the last *i*NKT cell maturation stage, many of the most strongly upregulated genes belonged to the killer lectin receptor family, which encodes activating and inhibitory NK receptors (NKRs) (Fig. 1c, bottom panel). By the final stage of ontogeny, *i*NKT cells expressed several NKR mRNAs at levels comparable to those of splenic NK cells (Supplementary Fig. 2a). In addition, NKR upregulation occurs selectively in developing *i*NKT cells but not in maturing MHC-restricted T cells or at any stage in early T cell development, with a few exceptions (Fig. 1e). Flow cytometric analysis of thymic *i*NKT cells confirmed the acquisition of surface NKRs over the course of maturation (Fig. 1f and Supplementary Fig. 2b). K-means clustering analysis determined that a large number of genes follow the same expression kinetics as NKRs over the course of the DP to stage 3 transition (Fig. 1g) suggesting that the progressive upregulation of NKRs may be part of a broader gene program. Together, these data comprehensively characterize shared and distinct gene expression changes occurring in maturing *i*NKT cells in the broader context of αβ T cell development.

Peripheral *i*NKT cell transcriptional signatures

In both mouse and human, differential cytokine production has been reported between $CD4^+$ and $CD4^-$ *i*NKT cells[5,32,33]. In addition, *i*NKT cell subsets from the liver have also been suggested to be functionally distinct[34]. To characterize the transcriptional basis that may underlie subset-specific functional differences, we assessed the gene expression profiles of $CD4^+$ and $CD4^-$ *i*NKT cells sorted from the spleen, liver and lungs of mice. 159 and 261 genes were differentially expressed between $CD4^+$ and $CD4^-$ *i*NKT cells from the spleen or lung, respectively, while only 17 transcripts were differentially expressed between the liver subsets (Fig. 2a, Supplementary Fig. 3 and Supplementary Tables 9–14), suggesting that splenic and pulmonary *i*NKT cell subsets may be more functionally distinct from each other than are liver *i*NKT cell subsets. In all three tissues, we found that a number of NKRs were among the most differentially expressed genes, with elevated levels in $CD4^-$ compared to $CD4^+$ subsets (Fig. 2a and Supplementary Fig. 3). In liver and spleen, this reflected a larger percentage of NKR-positive cells in $CD4^-$ compared to $CD4^+$ *i*NKT cells, as determined by flow cytometry (Fig. 2b). NKR mRNA expression in peripheral *i*NKT cells, although reduced compared to stage 3 thymic *i*NKT cells, was maintained at higher levels than most MHC-restricted αβ T cell subsets (Fig. 2c and data not shown). In addition, comparison of liver *i*NKT cell subsets to their splenic counterparts revealed that $CD4^-$ populations are more transcriptionally similar to one another than are $CD4^+$ *i*NKT cells across tissues. A small number of genes were expressed differentially between tissues regardless of CD4 expression (Supplementary Fig. 4), supporting the idea that *i*NKT cells subsets may perform organ-specific functions.

An extensive NK- *i*NKT shared transcriptional program

Because *i*NKT cells share broad functional features with NK cells, we hypothesized that NKR expression by *i*NKT cells might reflect a much larger NK-NKT shared transcriptional program than currently appreciated. To assess the extent of transcriptional relatedness between *i*NKT cells and NK cells at a global level, and to compare the NK- *i*NKT relationship and with that of *i*NKT to T cells, we calculated Euclidian distances between subsets of steady-state NK, T and *i*NKT cells. Euclidian distance, a measure of the similarity between the gene expression patterns of pairwise compared subsets, was determined using the 15% of gene probes with the highest variability among the subsets analyzed, and is displayed as a matrix (Fig. 3). This revealed a strong degree of similarity between mature *i*NKT and NK cell gene expression (Fig. 3, top panel, area i), contrasting greater distances observed between all MHC-restricted T cells and NK cells (Fig. 3, top panel, area ii). Compared to the *i*NKT -NK cell distance, *i*NKT cells exhibited a somewhat closer relationship to memory $CD8^+$ T cells and certain memory $CD4^+$ T cell, however, the latter

were sorted using markers expressed at similar levels by *i*NKT cells ($CD44^+$ and $CD62L^{low}$) and likely thus themselves contain a significant percentage of *i*NKT cells. The average Euclidian distance between *i*NKT and NK cells was only slightly larger than that separating *i*NKT cells from naive $CD4^+$ and $CD8^+$ T cells (Fig. 3, top panel, area iii). These relative relationships were maintained when the analysis was performed using all genes and not just the 15% most variable (data not shown). To determine whether the NK- *i*NKT relationship is dependent on the shared expression of NKRs, we recalculated the distance matrix after removing NKRs and related molecules from the data (see on-line methods). The outcome of the analysis remained essentially unchanged (Supplementary Fig. 5). Thus, the transcriptional relationship between NK and *i*NKT cells is not limited to the shared expression of NKRs. Together, these data support an unexpectedly substantial transcriptional relationship between *i*NKT and NK cells that is close in magnitude to that occurring between *i*NKT and naive T cells.

Shared and distinct *i*NKT cells programs

We next sought to identify the specific concordantly-regulated genes among *i*NKT, NK and T cells. For this purpose, we performed one-way analysis of variance (ANOVA) comparing the transcriptomes of peripheral steady-state *i*NKT cells, NK cells and naive and memory $CD4^+$ and $CD8^+$ T cells. Genes with low variability over the entire ImmGen dataset were excluded from the analysis. 20.2% (1192 genes) of the remaining genes were differentially expressed (Bonferroni corrected P value < 0.05) between the three groups. The modulated genes were further classified into 6 categories (see on-line methods): genes expressed similarly in NK and *i*NKT cells compared to T cells (A_1 or A_2, higher or lower than in T cells, respectively), genes expressed similarly in *i*NKT and T cells compared to NK cells (B_1 or B_2, higher or lower than in NK cells, respectively), and genes expressed differentially in *i*NKT cells (C_1 or C_2, higher or lower than T and NK cells, respectively) (Fig. 4a). A number of transcription factors, TCR signaling components, and cytokine or chemokine-related molecules that are known to be expressed differentially by NK, *i*NKT, and T cells partitioned as expected in the ANOVA categories (Fig. 4b–d). For example, the transcription factor PLZF (*Zbtb16*, Fig. 4b), the chemokine receptor *Cxcr6* (Fig. 4c), and a component of the high-affinity IL-12 receptor (*Il12rb1*) (Fig. 4d), all known to be expressed at higher levels in *i*NKT cells than in resting T or NK cells[29,30,35,36], partitioned to category C_1 (higher in *i*NKT cells relative to NK and T cells). The transcription factor T-bet (*Tbx21*, Fig. 4b) as well as *Il12rb2* (Fig. 4d), known to be upregulated in NK and *i*NKT cells[37], partitioned to category A_1 (higher in *i*NKT and NK cells relative to T cells). TCR signaling components such as *Cd3e*, *Itk*, *Plcg1*, and *Zap70* (Fig. 4c) segregated as anticipated to category B_1 (higher in *i*NKT and T compared to NK, Fig. 4c, d). Functional biological process enrichment analysis using DAVID software[38] of category A_1 genes showed a statistically significant enrichment for effector functions including NK cell-mediated immunity, chemokine or cytokine responses, signal transduction, and cell motility (Table 1). Distinct biological processes were also enriched in the group of genes upregulated in *i*NKT cells as compared to NK and MHC-restricted T cells (category C_1), and indicated a role for these genes in proliferation and survival, likely reflecting the uniquely activated phenotype of steady state *i*NKT cells (Table 2).

We found that a comparable number of genes were expressed similarly in *i*NKT and NK lineages (36.91%, 440 genes, categories A_1 and A_2) as in *i*NKT and T cells (43.88%, 523 genes, categories A_2 and B_2). In addition, about one-fifth of the differentially expressed genes were regulated uniquely in *i*NKT cells (19.21%, 229 genes, categories C_1 and C_2) (Fig. 4a, Supplementary Table 15). Thus, in addition to expressing a distinctive genetic program, *i*NKT cells share a transcriptional program transcriptional with steady-state NK

cells that is extensive and similar in magnitude to that shared between *i*NKT cells and other αβ T cells, consistent with the Euclidian distance analysis.

Thymic induction of NK- *i*NKT shared programs

To determine at what point during thymic development the transcriptional program shared by mature *i*NKT and NK cells is induced, we analyzed the expression patterns of ANOVA category A_1 genes over the course of MHC-restricted T and *i*NKT cell differentiation by hierarchical clustering. Although approximately 75% of the shared, upregulated NK- *i*NKT gene program is expressed in early thymic precursors (ETP), these genes are then largely shut-down or downregulated by the DP thymocyte stage, and are not reactivated in differentiating $CD4^+$ or $CD8^+$ T cells. In contrast, differentiating *i*NKT cells upregulated these same genes, such that stage 3 thymic *i*NKT cells expressed more than 90% of the program (Fig. 5a). Consistent with this observation, the FC distribution of category A_1 genes calculated by comparing the expression of each gene in developing *i*NKT cells and their DP progenitor showed an increasing shift towards higher FC values with *i*NKT maturation. All three distributions were significantly different from the baseline FC distribution for all expressed genes as determined by a Kolmogorov-Smirnov (K-S) test, with *P* values for enrichment of 3.97×10^{-15}, 1.28×10^{-19} and 9.14×10^{-48} for stage 1, 2 and 3 *i*NKT cells respectively (Fig. 5b). In contrast, the FC distribution calculated by comparing naïve splenic $CD4^+$ or $CD8^+$ T cells to DP was not significantly different between the category A_1 geneset and all expressed genes ($P = 0.79$ and $P = 0.65$, respectively). Conversely to category A_1 genes, the genes that were significantly downregulated in mature NK and *i*NKT cells compared to T cells (ANOVA category A_2) were not as strongly elicited in differentiating thymic *i*NKT cells as in $CD4^+$ and $CD8^+$ T cells (Supplementary Fig. 6). These data extend our earlier observation that *i*NKT cells acquire NKR expression at the end of development, and indicate that a large part of the transcriptional program *i*NKT cells share with NK cells is acquired late in thymic maturation.

*i*NKT cell programs expressed in γδ T cells

We next asked if other lymphocytes with innate features might also express this transcriptional program. γδ T cells can be categorized on the basis of their TCR Vγ and Vδ chain usage. The subset bearing the Vγ2 chain tends to be IL-17 polarized, whereas the $V\gamma2^-$ subsets (including $V\gamma1.1^+V\delta6.3^+$ and $V\gamma1.1^+V\delta6.3^-$ subsets) predominantly secrete T_H1 or T_H2 cytokines[39]. Distinct intraepithelial lymphocyte (IEL) $CD8\alpha\alpha^+$ γδ T cell populations posses an effector phenotype similar to that of *i*NKT cells[40,41].

In splenic γδ T cells negative for the activation marker CD44, NK- *i*NKT shared genes (ANOVA category A_1) were expressed at relatively low levels, similar to those observed in resting T cells. A subset of these genes was upregulated in splenic $CD44^+$ $V\gamma2^+$ and $V\gamma2^-$ cells. Strikingly, we found that all tissue-resident IEL γδ T cells expressed a large cluster of category A_1 genes at levels surpassing those of the reference NK and *i*NKT populations (Fig. 6a). FC distributions comparing the expression of category A_1 genes in splenic or IEL γδ T cells and T cells revealed an enrichment of over-expressed genes (*P* value 4.67×10^{-41} and 3.90×10^{-34}, respectively) (Fig. 6b). IEL γδ T cells expressed an even smaller proportion of ANOVA category A_2 genes (downregulated in NK and *i*NKT compared to T) than did *i*NKT cells, although most of these genes were expressed in splenic γδ T cells (data not shown). Although less prominently than for the shared NK- *i*NKT cell program (category A_1) a portion of the genes upregulated in *i*NKT cells but not T or NK subsets (category C_1) were also relatively highly expressed in γδ T cell subsets (Supplementary Fig. 7a). Together, these data suggest that most of the gene program shared by NK and *i*NKT cells, as well as part of the program differentially upregulated in *i*NKT cells, is also utilized by populations of γδ T cells at steady state.

NK-*i*NKT shared program induction in CD8$^+$ T cells

We next investigated the expression of the shared NK- *i*NKT transcriptional program in CD8$^+$ effector T cells, a cell population that shares functional characteristics with NK and *i*NKT cells, including the expression of certain NKRs[42]. We examined the expression of ANOVA categories A_1 and A_2 in CD8$^+$ T cells from the spleen of ovalbumin (OVA)-reactive αβ TCR transgenic (OT-I) mice following infection with *Listeria monocytogenes* expressing OVA. Only a small fraction of ANOVA category A_1 genes became upregulated in effector CD8$^+$ T cells at 12, 24, and 48 h following Listeria infection. By day 6, however, the NK- *i*NKT shared program was dramatically upregulated, with a FC distribution significantly enriched for higher expression, and this genetic program was maintained in CD8$^+$ memory T cells as late as day 100 following infection (Fig. 7a, b). In contrast, only a limited portion of category C_1 genes (differentially expressed by *i*NKT cells) followed a similar pattern in effector CD8$^+$ T cell populations (Supplementary Fig. 7b). The genes downregulated in NK and *i*NKT cells compared to T cells (ANOVA category A_2, which by definition are widely expressed in naive OT-I transgenic CD8$^+$ T cells), were partially repressed following infection (Supplementary Fig. 8). Similar results were obtained when the expression of the NK- *i*NKT shared transcriptional programs were examined in CD8$^+$ T cells from OT-I transgenic mice infected with OVA-expressing vesicular stomatitis virus (VSV) (Supplementary Fig. 9). We found that the homologous human genes comprising the NK- *i*NKT shared program (ANOVA category A_1) were also expressed at significantly higher levels in human peripheral blood effector memory CD8$^+$ T cell populations as compared to naive CD8$^+$ T cells[43] (Fig. 7c, d). Thus, a large proportion of the genetic program shared by NK and *i*NKT cells is elicited in effector αβ T cells, but only several days following their activation.

Discussion

*i*NKT cells do not fit the classical paradigm of adaptive T cell immunity. Indeed, innate features are at the core their physiological functions[3,14]. The transcriptional basis for these features, however, remains incompletely defined.

The transcriptional programs and regulatory factors we found to be operating during *i*NKT cell maturation were consistent with previous reports, and included *Zbtb16* (PLZF), *Vdr*, *Tbx21* (Tbet), components of the NF-κB and Ras-MapK pathways, and NKRs[4,37,44,45]. By comparing developing *i*NKT cells to differentiating MHC-restricted T cells, we define transcriptional programs expressed specifically by stage 1 *i*NKT cells shortly after the DP branch point. We thus highlight a large number genes not previously known to affect *i*NKT cell biology but likely modulating *i*NKT cell thymic maturation specifically.

A number of genes reported to be involved in *i*NKT cell development were not expressed in an *i*NKT cell specific-manner. For instance, NF-κB and Ras/MapK pathway members such as *NFkb1, RelB*, Ras/MapK and Egr2 were upregulated in both *i*NKT and non-*i*NKT thymocyte populations[4] (data not shown). Other genes known to affect *i*NKT cells ontogeny, Bcl11b[46] and the chromatin modifier Med1[47] for example, exhibited little or no transcriptional variation over the course of *i*NKT cell differentiation. The functional regulation of these and other genes not detected in our analyses may thus be controlled at post-transcriptional levels.

Mature *i*NKT cells share several innate functional features with NK cells, developmentally distant relatives. We hypothesized that NK and *i*NKT cells shared a broader transcriptional program than is currently appreciated. Our analyses revealed that the transcriptional patterns *i*NKT cells share with NK cells make up nearly as large a part of the *i*NKT transcriptome as those shared with MHC-restricted T cells. Further, we found that the NK- *i*NKT cell shared

program was also active in steady-state γδ T cells subsets similarly poised for rapid, innate-like responsiveness. The NK- *i*NKT program can in addition be induced in adaptive αβ T cells, but only days following antigen-specific stimulation. These data suggest that the NK-*i*NKT shared program represents a core effector program operational in lymphocytes of distinct lineages, consistent with the many functional capabilities *i*NKT, NK and activated MHC-restricted T cells have in common.

The factors that regulate expression of this shared effector program remain to be fully defined. Nevertheless, our data provides several leads. Acquired largely at the end of thymic *i*NKT cell maturation, the shared program mirrors the cells' acquisition of NKRs. NKR expression is driven by the transcription factor T-bet both in NK and *i*NKT cells[37]. Thus, it is likely that T-bet, which we found to be among the genes significantly upregulated in both NK and *i*NKT compared to resting T cells, plays an important role in eliciting and maintaining at least part of the shared effector lymphocyte program. T-bet-deficient *i*NKT cells express reduced mRNA and protein levels of IFN-γ, granzyme B, Fas ligand, CCR5, and CD38, all molecules that are part of the shared program[44]. IL-15, recently reported to act upstream of T-bet in maturing *i*NKT cells[45], is also likely to be involved in *i*NKT cell acquisition of this program. Furthermore, IL-15 and T-bet are important for regulating $CD8^+$ effector T cell responses[48] as well as for the homeostasis of IEL γδ T cells[49]. Thus, IL-15 and T-bet may play an important role in inducing the expression of the shared effector program in several cell types.

We also highlight a number of previously unappreciated transcriptional regulators exhibiting similar expression patterns in NK and *i*NKT cells, and that may help regulate the core effector program in innate lymphocytes. For example, the transcription factor Bhlhe40, a circadian rhythm regulator[50] possessing immune-modulatory functions in $CD4^+$ T cells[51,52], as well as Smad3, a transcription factor important for tuning TGF-β-mediated lymphocyte activation, were upregulated in both NK and *i*NKT cells compared to resting T cells. Also, a number of factors were downregulated in NK and *i*NKT cells. Among these genes were *Lef1* and *Tcf7*, transcription factors downstream of the Wnt signaling pathway that are important for the establishment and maintenance of T cell identity[53]. Wnt signaling, which is typically repressed upon T cell activation[54,55]. The relatively low expression of *Tcf7* and *Lef1* in NK and *i*NKT cells during development is consistent with the acquisition their terminally differentiated effector phenotype at steady state.

Approximately 20% of the significantly differentially expressed genes between NK, *i*NKT, and T cells exhibited transcriptional patterns specific to *i*NKT cells. As expected, PLZF (*Zbtb16*) and Gata-3 were among these factors. The genes preferentially upregulated in *i*NKT cells (category C_1 genes) were enriched for cellular activation and survival programs. For instance, several AP-1 family members (*Jun*, *Junb*, *Jund*, *Fos*, *Fosb* and *Cebpb*) normally induced in cells only after activation, were expressed relatively highly in *i*NKT cells. Constitutive expression of AP-1 transcription factors by *i*NKT cells is consistent with their poised effector phenotype.

Together, these data offer a new view of *i*NKT cells and their relationships to other lymphocyte lineages. Using the Immunological Genome consortium database, we have uncovered gene expression programs that both link to and differentiate *i*NKT cells from other innate and adaptive lymphocytes. Our data highlight extensive genetic modules that are shared with NK cells and other innate-like T cells, and that are also elicited several days following activation in adaptive T cells. By defining both distinct and shared transcriptional programs in *i*NKT cells, our data opens avenues for future research and brings into clearer focus how lymphocyte populations that differ markedly in their ontogeny can ultimately

carry out similar effector functions through modular expression of similar transcriptional programs.

Online Methods

Mice

6 week old C57BL/6 male mice shipped from Jackson Laboratories one week prior to organ harvest were used. Mice were maintained under specific pathogen free conditions. All studies were approved by the Animal Care and Use Committee of the Dana-Farber Cancer Institute.

Antibodies

Anti-CD19 (MB19-1), anti-B220 (RA3–6B2), anti-Ter119 (TER119), anti-CD11b (M1/70), anti-CD11c (N418), anti-Ly6G/Gr1 (A18), anti-CD8α (53/6.7), anti-TCRβ (H58–597), anti-NK1.1 (PK136), anti-CD44 (IM7), anti-CD4 (GK1.5), anti-CD45 (102), anti-CD3ε (145–2C11), anti-Ly49e,f (CM4), anti-NKG2D (CX5), anti-NKG2A (16a11), anti-2B4 (ebio244F4), and anti-CD16,32 (2.4G2) were from eBioscience. Anti-Ly49a (A1) and anti-Ly49c,1 (5F6) were from BD biosciences. CD1d-tetramers loaded with PBS-57 (an α-galactosylceramide analog) were provided by the NIH tetramer facility.

Cell isolation, microarray analysis and subset nomenclature

All immune cells purification were performed in strict adherence to the ImmGen standard Operating Procedure guidelines (available at www.immgen.org). Further details on thymic, NK cell, γδ T cell, and CD8+ T cell populations used for comparison to *i*NKT cells can be found at the ImmGen website. For *i*NKT cells, thymocytes, splenocytes, liver and lung mononuclear cells were isolated from 5–10 mice per sample. Thymocytes were disaggregated, blocked with anti-CD16 and CD32 (clone 2.4G2), then stained with fluorophore-labeled antibodies for depletion of non- *i*NKT cell populations, and separated with anti-fluorophore magnetic beads (Miltenyi). Spleens were disaggregated, treated with ACK lysis buffer (Lonza) to remove red blood cells, and then depleted of non- *i*NKT cells as above. Lungs and livers were harvested after perfusion with cold PBS and mechanically homogenized. Lungs were digested for 15 minutes at 37° C in 7 U/ml of Liberase III enzyme (Roche) in DMEM, then filtered and washed. Livers were homogenized and liver mononuclear cells (LMNC) were isolated by Ficoll density gradient centrifugation. *i*NKT cell-enriched thymocytes, splenocytes, LMNCs and lung cells were stained for cell surface markers (see www.immgen.org for complete staining procedure including gating strategy) and double sorted directly into Trizol (Invitrogen) at a purity of >99%. By staining thymocytes from CD1d-deficient mice that lack *i*NKT cells, a false positive rate of 2.7 ± 0.9% was estimated in the case of $CD44^{-}$ $NK1.1^{-}$ (Stage 1) *i*NKT cells, the rarest thymic subset. Contamination was negligible for other *i*NKT cell populations. Two to four replicates were obtained for each sample using a FACSAria, with the exception of the $CD4^{-}$ *i*NKT subset from the lung, for which only a single replicate passed quality control (see below). RNA extraction, microarray hybridization (Affymetrix MoGene 1.0 ST array) and data processing were performed at the ImmGen processing center. For further details, please see Supplementary Table 16 (subset nomenclature key), the Data Generation and Quality Control pipeline documentation, the ImmGen Quality Control Statistics or the ImmGen website (www.immgen.org). Non-ImmGen human and mouse datasets were downloaded from NCBI Geo datasets[43,56]. Human homologs of mouse genes were determined using NCBI HomoloGene.

Data analysis and visualization

Data from the March 2011 ImmGen release (802 arrays with 22,268 probesets) were used. Probesets associated with the same gene symbol were consolidated by selecting the probeset with the highest mean expression overall. For heatmaps, data were log2-transformed and a relative color scale with row centering (subtraction of the mean) and normalization was used. Heatmaps were produced by using the HeatmapViewer module of GenePattern (www.broadinstitute.org/cancer/software/genepattern/). When indicated, Pearson correlation with pair-wise complete linkage was applied to rows for clustering analysis. Volcano plots were produced by using the Multiplot module of GenePattern.

For K-means clustering analysis, genes were pre-filtered for mean expression value ≥ 120 (cut-off above which genes have a 95% chance of expression) and for FC>2 between any two subsets analyzed. Clustering was performed using the ExpressCluster application (Scott Davis, Harvard Medical School, Boston MA – application and documentation available at http://cbdm.hms.harvard.edu) with K=10.

For Euclidian distance matrices, the 15% most variable genes were identified using the PopulationDistances PCA application (Scott Davis), which filters probes based on a variation of ANOVA analysis using the geometric standard deviation of populations to weight genes that vary in multiple populations. The selected genes were log2-transformed, filtered for probes with a mean expression value ≥ 120, and mean centered prior to visualization. Activating and inhibitory NKRs, adaptors and signaling partners (*Klra1*, *Klra2*, *Klra3*, *Klra5*, *Klra6*, *Klra8*, *Klra9*, *Klra10*, *Klra17*, *Klrb1a*, *Klrb1b*, *Klrb1c*, *Klrb1f*, *Klrc1*, *Klrc2*, *Klrc3*, *Klrd1*, *Klre1*, *Klri1*, *Klri2*, *Klrk1*, *Klrg2*, *Ncr1*, *Cd244*, *Fcer1g*, *Tyrobp* and *Hcst*) were present in the initial gene list used in the first analysis (Fig. 3, top panel) and were manually removed as indicated for the second analysis (Fig. 3, bottom panel). *Klra1*, *Klra3*, *Klra5*, *Klra6*, *Klra8*, *Klra9*, *Klra10*, *Klrb1a*, *Klrb1b*, *Klrb1c*, *Klrb1f*, *Klrc1*, *Klrc2*, *Klrc3*, *Klrd1*, *Klre1*, *Klri1*, *Klri2*, *Klrk1*, *Ncr1*, *Cd244*, *Fcer1g*, *Tyrobp* (but not *Klra2, Klra17, Klrg2 or Hcst*) were present after filtering for 15% most variable gene list and were contributed to the first, but not the second Euclidian distance analysis.

For ANOVA, data were log2-transformed, low-variability genes with a standard deviation of ≤ 0.5 across all ImmGen samples were removed and only genes with expression values ≥ 120 in 2 or more arrays were considered, leaving 5,900 genes. The ANOVA was performed using the Matlab (MathWorks) function “anova1” to compare *i*NKT, NK and T cell populations Bonferroni correction was applied to the resulting list of *P* values, and 1192 genes passed the ANOVA ($p < 0.05$ out of 5,900). A secondary test was used to determine which of the three populations significantly differed from the other two, as indicated by the “multcompare” function of Matlab. Based on the results of this analysis, the 1,192 genes were then separated into 14 gene categories. By comparing the average expression levels in NK, *i*NKT, and T cells for each gene, these categories were then sorted into 6 groups of genes: genes expressed most similarly by NK and *i*NKT cells (either A_1, up- or A_2, downregulated compared to T), genes expressed most similarly in T and *i*NKT (either B_1, up- or B_2, downregulated compared to NK) and genes expressed uniquely in *i*NKT (either C_1, up- or C_2, downregulated compared to NK and T). Functional geneset enrichment analysis was performed using DAVID software (Version 6.7, National Institutes of Health, National Institute for Allergy and Infectious Diseases)[57]. Panther biological processes[38] are shown in Tables 1 and 2. *P* values calculated by DAVID represent a modified Fisher Exact test. Biological processes with *P* values less than 0.01 are shown.

For estimation of the significance of enrichment for the FC distributions associated with the heatmaps in Figs. 5–7, Kolmogorov-Smirnov *P* values were calculated with JMP (SAS

Institute) comparing the selected geneset to all genes meeting the criteria for expression (>120) in the samples tested.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank the NIH tetramer facility for their ongoing support. We thank M. Painter, C. Benoist, S. Raychaudhuri, X. Hu, J. Erricson, S. Davis, H. Li, T. Kreslavsky, L. Lanier and A. Goldrath for advice, discussions and technical assistance. The work was supported by R01AI063428 to M.B.B., T32AI007306 to P.J.B, and R24AI072073 to the ImmGen consortium.

References

1. Heng TS, Painter MW, Immunological Genome Project C. The Immunological Genome Project: networks of gene expression in immune cells. Nat Immunol. 2008; 9:1091–1094. [PubMed: 18800157]
2. Cohen NR, Garg S, Brenner MB. Antigen Presentation by CD1 Lipids, T Cells, and NKT Cells in Microbial Immunity. Adv Immunol. 2009; 102:1–94. [PubMed: 19477319]
3. Bendelac A, Savage PB, Teyton L. The biology of NKT cells. Annu Rev Immunol. 2007; 25:297–336. [PubMed: 17150027]
4. Godfrey DI, Stankovic S, Baxter AG. Raising the NKT cell family. Nat Immunol. 2010; 11:197–206. [PubMed: 20139988]
5. Coquet JM, et al. Diverse cytokine production by NKT cell subsets and identification of an IL-17-producing CD4-NK1.1- NKT cell population. Proc Natl Acad Sci U S A. 2008; 105:11287–11292. [PubMed: 18685112]
6. Cerundolo V, Silk JD, Masri SH, Salio M. Harnessing invariant NKT cells in vaccination strategies. Nat Rev Immunol. 2009; 9:28–38. [PubMed: 19079136]
7. Makino Y, Kanno R, Ito T, Higashino K, Taniguchi M. Predominant expression of invariant V alpha 14+ TCR alpha chain in NK1.1+ T cell populations. Int Immunol. 1995; 7:1157–1161. [PubMed: 8527413]
8. Bendelac A. Mouse NK1+ T cells. Curr Opin Immunol. 1995; 7:367–374. [PubMed: 7546402]
9. Godfrey DI, MacDonald HR, Kronenberg M, Smyth MJ, Van Kaer L. NKT cells: what's in a name? Nat Rev Immunol. 2004; 4:231–237. [PubMed: 15039760]
10. Gregoire C, et al. The trafficking of natural killer cells. Immunol Rev. 2007; 220:169–182. [PubMed: 17979846]
11. Matsuda JL, et al. Tracking the response of natural killer T cells to a glycolipid antigen using CD1d tetramers. J Exp Med. 2000; 192:741–754. [PubMed: 10974039]
12. Cooper MA, et al. In vivo evidence for a dependence on interleukin 15 for survival of natural killer cells. Blood. 2002; 100:3633–3638. [PubMed: 12393617]
13. Matsuda JL, et al. Homeostasis of V alpha 14i NKT cells. Nat Immunol. 2002; 3:966–974. [PubMed: 12244311]
14. Brigl M, Brenner MB. How invariant natural killer T cells respond to infection by recognizing microbial or endogenous lipid antigens. Semin Immunol. 2010; 22:79–86. [PubMed: 19948416]
15. Lanier LL. Evolutionary struggles between NK cells and viruses. Nat Rev Immunol. 2008; 8:259–268. [PubMed: 18340344]
16. Kuylenstierna C, et al. NKG2D performs two functions in invariant NKT cells: direct TCR-independent activation of NK-like cytolysis and co-stimulation of activation by CD1d. Eur J Immunol. 2011; 41:1913–1923. [PubMed: 21590763]
17. Kawamura T, et al. NKG2A inhibits invariant NKT cell activation in hepatic injury. J Immunol. 2009; 182:250–258. [PubMed: 19109156]

18. Maeda M, Lohwasser S, Yamamura T, Takei F. Regulation of NKT cells by Ly49: analysis of primary NKT cells and generation of NKT cell line. J Immunol. 2001; 167:4180–4186. [PubMed: 11591738]
19. Ota T, et al. IFN-gamma-mediated negative feedback regulation of NKT-cell function by CD94/NKG2. Blood. 2005; 106:184–192. [PubMed: 15746081]
20. Brennan PJ, et al. Invariant natural killer T cells recognize lipid self antigen induced by microbial danger signals. Nat Immunol. 2011; 12:1202–1211. [PubMed: 22037601]
21. Paget C, et al. Activation of invariant NKT cells by toll-like receptor 9-stimulated dendritic cells requires type I interferon and charged glycosphingolipids. Immunity. 2007; 27:597–609. [PubMed: 17950005]
22. Salio M, et al. Modulation of human natural killer T cell ligands on TLR-mediated antigen-presenting cell activation. Proc Natl Acad Sci U S A. 2007; 104:20490–20495. [PubMed: 18077358]
23. Reschner A, Hubert P, Delvenne P, Boniver J, Jacobs N. Innate lymphocyte and dendritic cell cross-talk: a key factor in the regulation of the immune response. Clin Exp Immunol. 2008; 152:219–226. [PubMed: 18336590]
24. Andrews DM, Scalzo AA, Yokoyama WM, Smyth MJ, Degli-Esposti MA. Functional interactions between dendritic cells and NK cells during viral infection. Nat Immunol. 2003; 4:175–181. [PubMed: 12496964]
25. Brigl M, Bry L, Kent SC, Gumperz JE, Brenner MB. Mechanism of CD1d-restricted natural killer T cell activation during microbial infection. Nat Immunol. 2003; 4:1230–1237. [PubMed: 14578883]
26. Fernandez NC, et al. Dendritic cells directly trigger NK cell functions: cross-talk relevant in innate anti-tumor immune responses in vivo. Nat Med. 1999; 5:405–411. [PubMed: 10202929]
27. Vincent MS, et al. CD1-dependent dendritic cell instruction. Nat Immunol. 2002; 3:1163–1168. [PubMed: 12415264]
28. Walzer T, Dalod M, Vivier E, Zitvogel L. Natural killer cell-dendritic cell crosstalk in the initiation of immune responses. Expert Opin Biol Ther. 2005; 5(Suppl 1):S49–S59. [PubMed: 16187940]
29. Savage AK, et al. The transcription factor PLZF directs the effector program of the NKT cell lineage. Immunity. 2008; 29:391–403. [PubMed: 18703361]
30. 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]
31. Yu S, Cantorna MT. The vitamin D receptor is required for iNKT cell development. Proc Natl Acad Sci U S A. 2008; 105:5207–5212. [PubMed: 18364394]
32. Gumperz JE, Miyake S, Yamamura T, Brenner MB. Functionally distinct subsets of CD1d-restricted natural killer T cells revealed by CD1d tetramer staining. J Exp Med. 2002; 195:625–636. [PubMed: 11877485]
33. Watarai H, et al. Development and function of invariant natural killer T cells producing T(h)2- and T(h)17-cytokines. PLoS Biol. 2012; 10:e1001255. [PubMed: 22346732]
34. Crowe NY, et al. Differential antitumor immunity mediated by NKT cell subsets in vivo. J Exp Med. 2005; 202:1279–1288. [PubMed: 16275765]
35. Brigl M, et al. Innate and cytokine-driven signals, rather than microbial antigens, dominate in natural killer T cell activation during microbial infection. J Exp Med. 2011; 208:1163–1177. [PubMed: 21555485]
36. Johnston B, Kim CH, Soler D, Emoto M, Butcher EC. Differential Chemokine Responses and Homing Patterns of Murine TCRalphabeta NKT Cell Subsets. J Immunol. 2003; 171:2960–2969. [PubMed: 12960320]
37. Townsend MJ, et al. T-bet regulates the terminal maturation and homeostasis of NK and Valpha14i NKT cells. Immunity. 2004; 20:477–494. [PubMed: 15084276]
38. Thomas PD, et al. PANTHER: a browsable database of gene products organized by biological function, using curated protein family and subfamily classification. Nucleic Acids Res. 2003; 31:334–341. [PubMed: 12520017]

39. O'Brien RL, Born WK. gammadelta T cell subsets: a link between TCR and function? Semin Immunol. 2010; 22:193–198. [PubMed: 20451408]
40. Fahrer AM, et al. Attributes of gammadelta intraepithelial lymphocytes as suggested by their transcriptional profile. Proc Natl Acad Sci U S A. 2001; 98:10261–10266. [PubMed: 11526237]
41. Shires J, Theodoridis E, Hayday AC. Biological insights into TCRgammadelta+ and TCRalphabeta + intraepithelial lymphocytes provided by serial analysis of gene expression (SAGE). Immunity. 2001; 15:419–434. [PubMed: 11567632]
42. Vivier E, Anfossi N. Inhibitory NK-cell receptors on T cells: witness of the past, actors of the future. Nat Rev Immunol. 2004; 4:190–198. [PubMed: 15039756]
43. Gattinoni L, et al. A human memory T cell subset with stem cell-like properties. Nat Med. 2011; 17:1290–1297. [PubMed: 21926977]
44. Matsuda JL, et al. T-bet concomitantly controls migration, survival, and effector functions during the development of Valpha14i NKT cells. Blood. 2006; 107:2797–2805. [PubMed: 16357323]
45. Gordy LE, et al. IL-15 regulates homeostasis and terminal maturation of NKT cells. J Immunol. 2011; 187:6335–6345. [PubMed: 22084435]
46. Kastner P, et al. Bcl11b represses a mature T-cell gene expression program in immature CD4(+)CD8(+) thymocytes. Eur J Immunol. 2010; 40:2143–2154. [PubMed: 20544728]
47. Yue X, Izcue A, Borggrefe T. Essential role of Mediator subunit Med1 in invariant natural killer T-cell development. Proc Natl Acad Sci U S A. 2011; 108:17105–17110. [PubMed: 21949387]
48. Sullivan BM, Juedes A, Szabo SJ, von Herrath M, Glimcher LH. Antigen-driven effector CD8 T cell function regulated by T-bet. Proc Natl Acad Sci U S A. 2003; 100:15818–15823. [PubMed: 14673093]
49. Inagaki-Ohara K, Nishimura H, Mitani A, Yoshikai Y. Interleukin-15 preferentially promotes the growth of intestinal intraepithelial lymphocytes bearing gamma delta T cell receptor in mice. Eur J Immunol. 1997; 27:2885–2891. [PubMed: 9394814]
50. Honma S, et al. Dec1 and Dec2 are regulators of the mammalian molecular clock. Nature. 2002; 419:841–844. [PubMed: 12397359]
51. Miyazaki K, et al. The role of the basic helix-loop-helix transcription factor Dec1 in the regulatory T cells. J Immunol. 2010; 185:7330–7339. [PubMed: 21057086]
52. Sun H, Lu B, Li RQ, Flavell RA, Taneja R. Defective T cell activation and autoimmune disorder in Stra13-deficient mice. Nat Immunol. 2001; 2:1040–1047. [PubMed: 11668339]
53. Weber BN, et al. A critical role for TCF-1 in T-lineage specification and differentiation. Nature. 2011; 476:63–68. [PubMed: 21814277]
54. Willinger T, et al. Human naive CD8 T cells down-regulate expression of the WNT pathway transcription factors lymphoid enhancer binding factor 1 and transcription factor 7 (T cell factor-1) following antigen encounter in vitro and in vivo. J Immunol. 2006; 176:1439–1446. [PubMed: 16424171]
55. Gattinoni L, et al. Wnt signaling arrests effector T cell differentiation and generates CD8+ memory stem cells. Nat Med. 2009; 15:808–813. [PubMed: 19525962]
56. Marshall HD, et al. Differential expression of Ly6C and T-bet distinguish effector and memory Th1 CD4(+) cell properties during viral infection. Immunity. 2011; 35:633–646. [PubMed: 22018471]
57. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc. 2009; 4:44–57. [PubMed: 19131956]

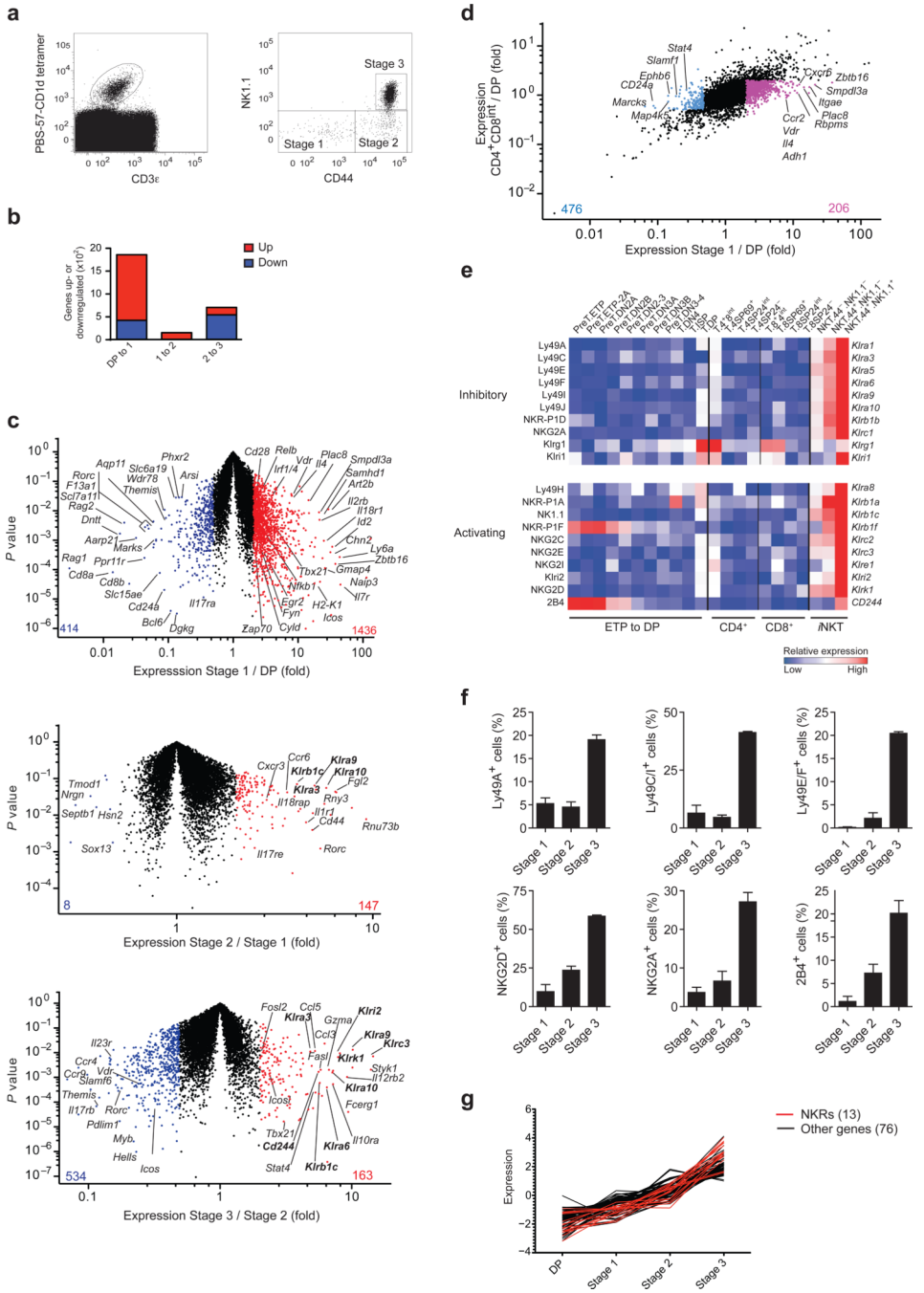

Fig. 1.

*i*NKT cells upregulate NKRs at the end of thymic differentiation. (**a**) Flow cytometric identification of thymic *i*NKT cells. (**b**) Number of genes differentially expressed during *i*NKT cell developmental transitions. (**c**) Genes differentially expressed between DP thymocytes and stage 1 (top), stage 1 and stage 2 (middle), and stage 2 and stage 3 *i*NKT cells (bottom). Genes upregulated are highlighted in red or blue, respectively. See Supplementary Tables 1–6 for complete lists. (**d**) Genes differentially expressed in stage 1 *i*NKT cells versus DP thymocytes. Colored highlights indicate genes differentially regulated in *i*NKT but not in $CD4^{+}8^{int}$ compared to DP thymocytes. See Supplementary Tables 7 and

8 for complete lists. For **b-d**, only genes with expression above the detection level one or more subsets and a coefficient of variation (CV) < 0.5 in all subsets were considered. The differential expression threshold used was FC > 2. (**e**) Expression of NKRs in differentiating thymocytes (see Supplementary Table 16 for key to subset nomenclature). Values were log2-transformed, row centered and locally color scaled. (**f**) NKR surface expression as determined by flow cytometry. Bars, mean percentage-positive ± s.e.m., n=3 mice. Data is representative of at least two independent experiments. (**g**) Expression of an NKR-containing gene cluster derived from a Euclidian distance-based K-means clustering analysis (mean correlation, 0.958). Genes were pre-filtered for expression and for FC > 2 between any 2 subsets, and data was log2-transformed. Klra3, 5, 6, 9, 10, Klrb1b, c, f, Klrc1, 2, 3, Klrd1, Klri2 and Klrk1 are labeled in red.

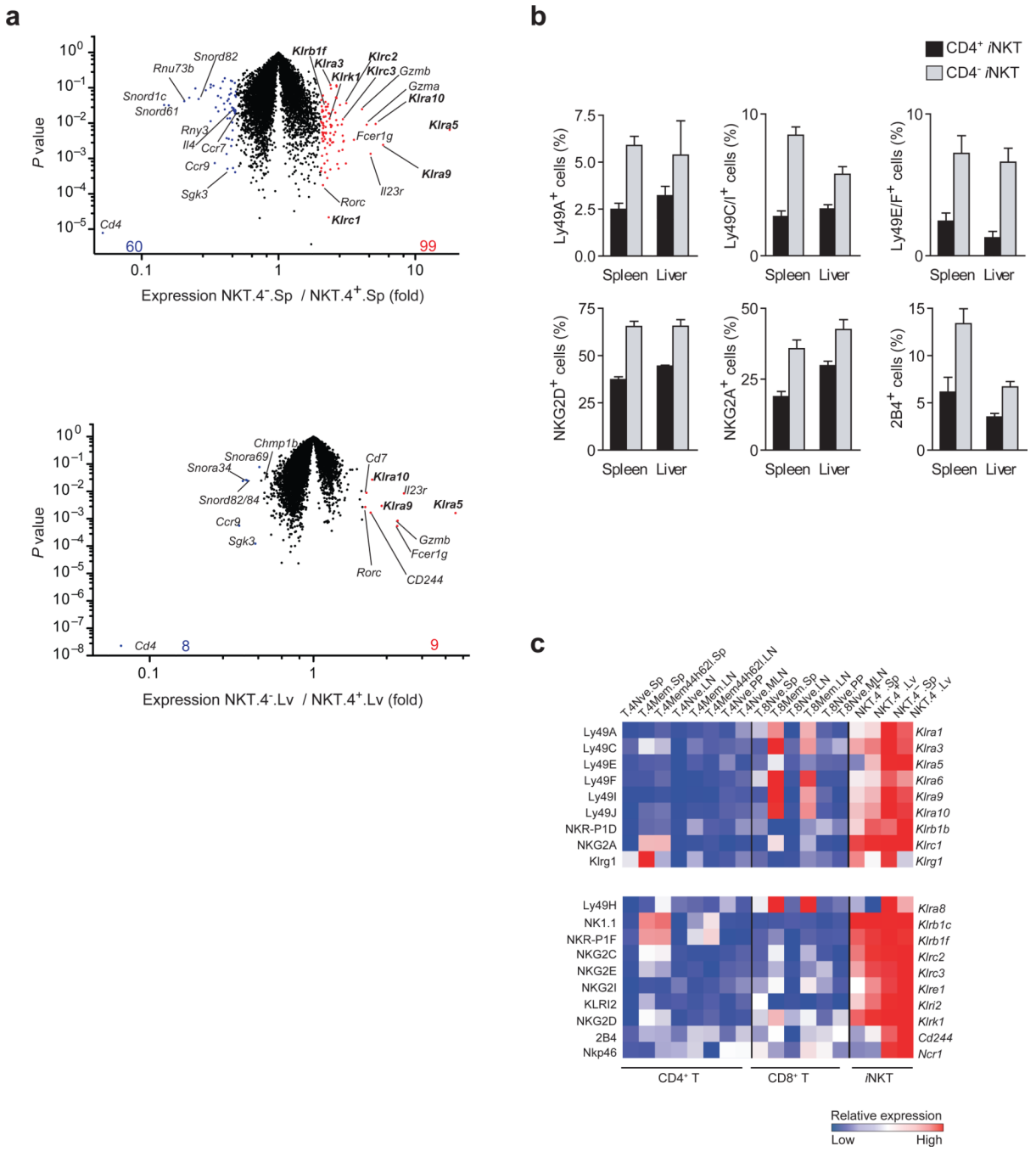

Fig. 2.

Expression of NKRs in peripheral CD4$^+$ and CD4$^-$ *i*NKT cells. (**a**) Top, differential gene expression between CD4$^+$ and CD4*i*NKT cells from the spleen; bottom, differential gene expression between CD4$^+$ and CD4$^-$ *i*NKT cells from the liver. Genes up- or downregulated with FC > 2 in CD4$^-$ compared to CD4$^+$ subsets are highlighted in red or blue, respectively. Only genes with expression values above the detection level in at least one subset and a CV < 0.5 in all subsets are displayed. (**b**) Flow cytometric quantification of NKRs on spleen and liver CD4$^+$ and CD4$^-$ *i*NKT cells. Bars represent mean percentage-positive for surface expression ± s.e.m., n=3 mice. Data is representative of at least two independent

experiments. (**c**) NKR expression in peripheral MHC-restricted T and *i*NKT cell subsets. Values were log2-transformed, gene row centered and local color scaling was used. See Supplementary Table 16 for key to subset nomenclature.

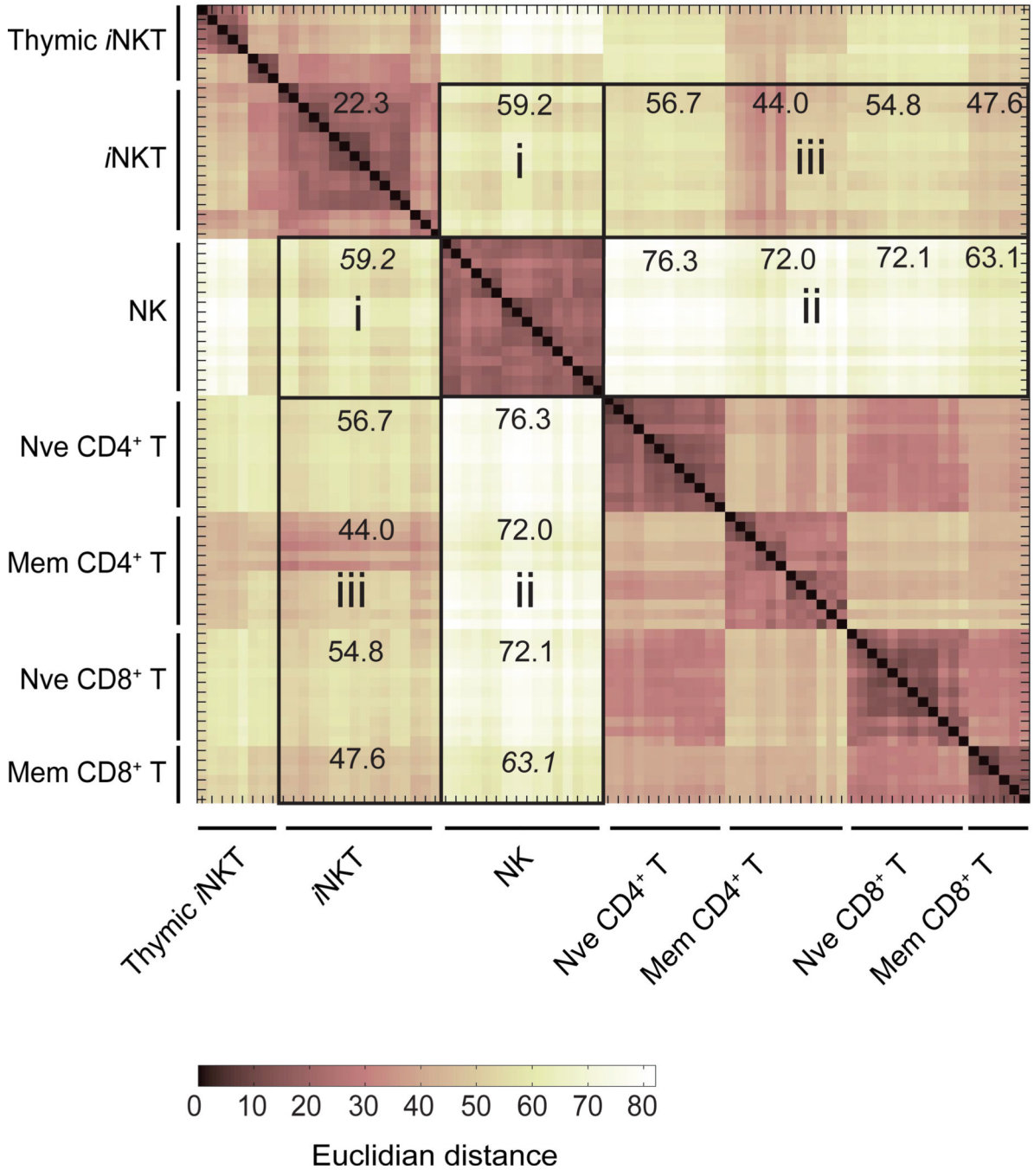

Fig. 3.

The global transcriptional relationship between NK and *i*NKT cells is similar in magnitude to the relationship between T and *i*NKT cells. Euclidian distance matrix calculated using the 15% most variable probes. Only genes with mean expression values above 120 were included in these analyses, and data was log2 transformed and mean centered. Numbers represent average Euclidian distances between intersecting subsets. Note that the matrix is symmetrical along the indicated diagonal. Areas of interest are marked with lowercase Roman numerals.

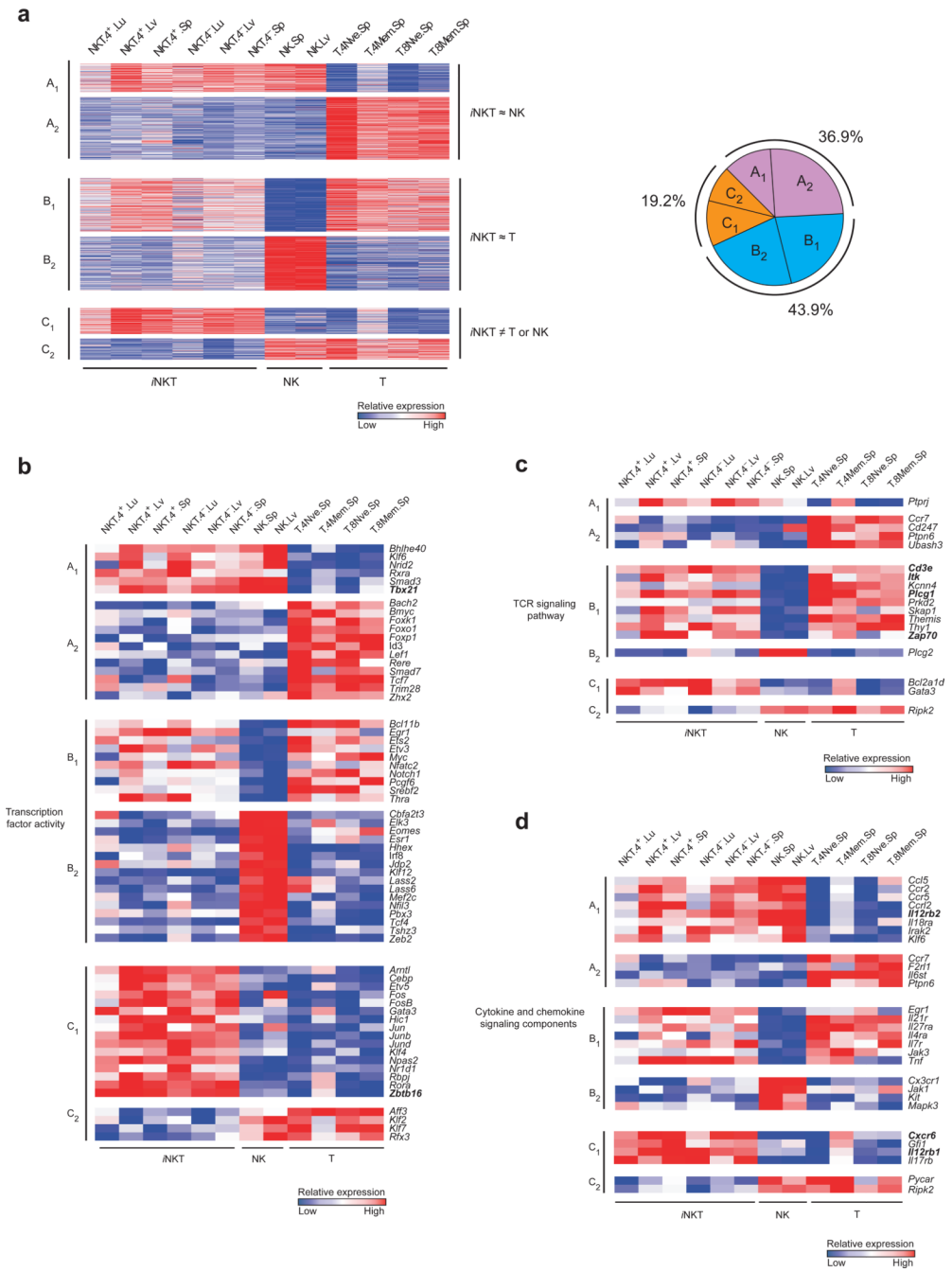

Fig. 4.

Characterization of shared and differential gene expression among *i*NKT, NK, and T cells. One-way ANOVA was performed to compare gene expression in *i*NKT, NK, and T cell populations. Differentially expressed genes were separated into six categories depending on common patterns in 2 of the 3 subsets. Genes up- (A_1) or down- (A_2) regulated in NK and *i*NKT compared to T cells; genes up- (B_1) or down- (B_2) regulated in *i*NKT and T compared to NK cells; genes up- (C_1) or down- (C_2) regulated in *i*NKT compared to T or NK cells. (**a**) *Left*, heatmap of genes expressed differentially in *i*NKT, NK, and T cells for the indicated cell subsets, separated according to category; *right*, proportion of differentially expressed

genes in each category. (**b-d**) Expression of ANOVA category genes in relevant manually selected functional gene groups (see methods). Genes reported in the literature to be highly expressed in each category shown in underlined bold. In all heatmaps, rows are mean-centered and normalized, and local scaling is used. See Supplementary Table 16 for key to subset nomenclature.

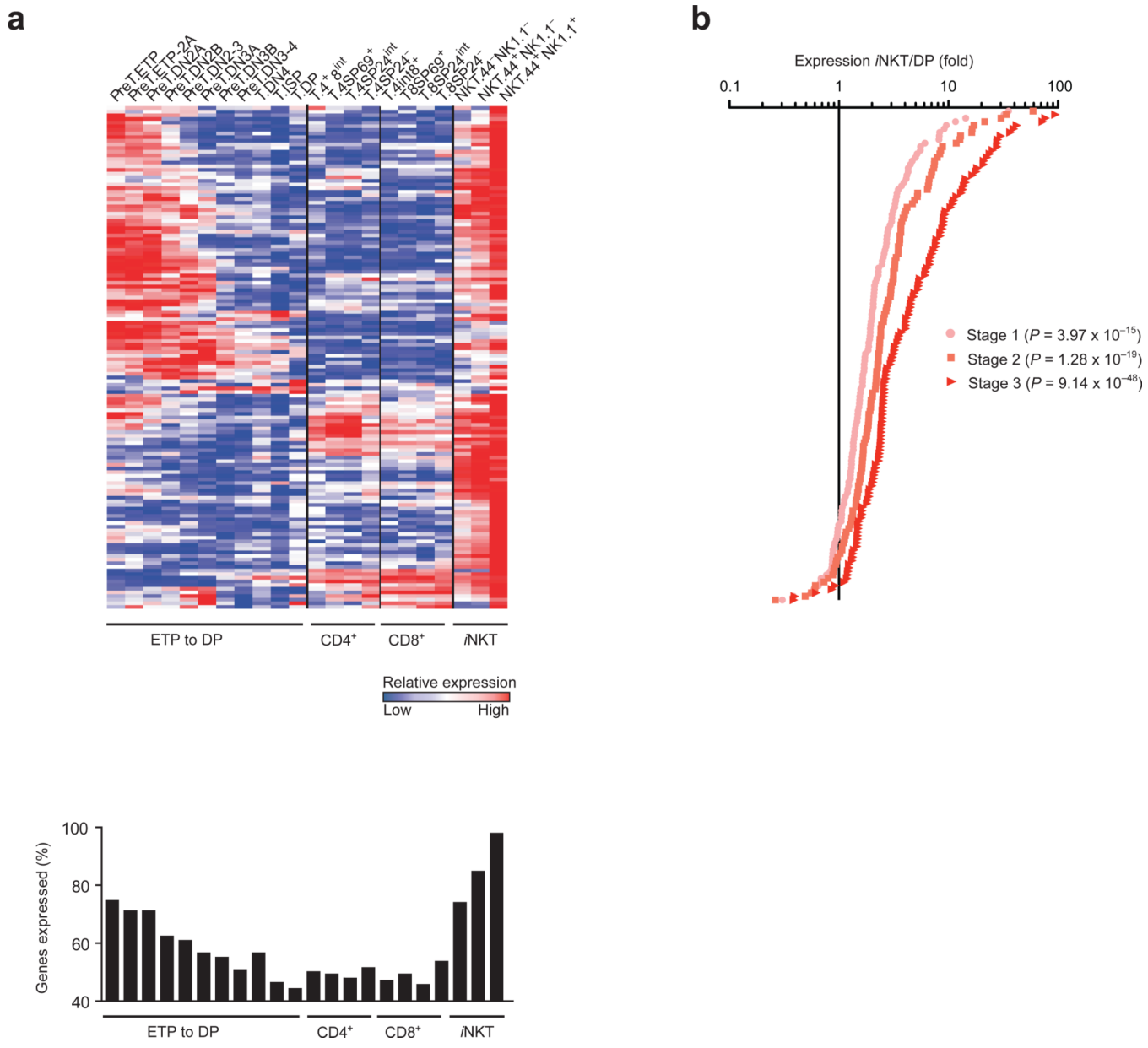

Fig. 5.

Transcriptional programs shared between NK and *i*NKT cells are acquired during thymic *i*NKT cell maturation. (**a**) Top, relative expression of genes significantly upregulated in peripheral NK and *i*NKT cells (ANOVA category A_1) over the course of thymic T and *i*NKT cell maturation. Genes were ordered by hierarchical clustering using Pearson correlation. Rows are mean-centered and normalized, and local scaling is used. Bottom, percentages of genes exceeding the threshold for expression over the course of thymic maturation. T cell populations are represented in the same order as in the heatmaps. (**b**) FC distribution of category A_1 genes in stage 1, 2 and 3 *i*NKT cells compared to DP thymocytes with individual genes displayed in rank-order along the y-axis. *P* values represent significance of K-S test comparing the FC distribution of the ANOVA category A1 geneset to that of all expressed genes. See Supplementary Table 16 for key to subset nomenclature.

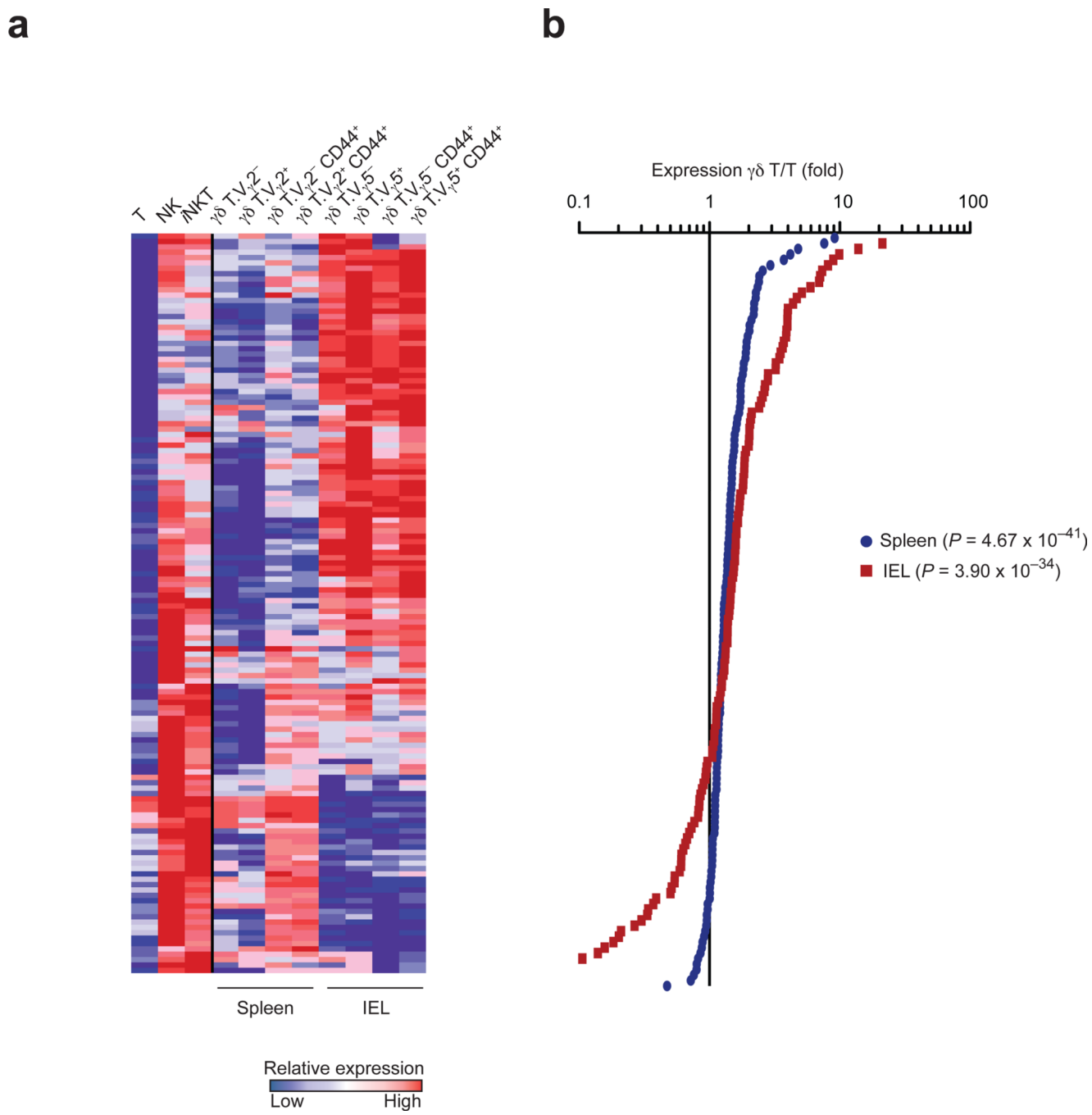

Fig. 6.

Activated splenic γδ T cells and IEL γδ T cells express NK- *i*NKT shared gene programs. (**a**) Relative expression of genes significantly upregulated in peripheral NK and *i*NKT compared to T cells (ANOVA category A_1) in the splenic and IEL γδ T cell subsets indicated. Averaged values for the ANOVA subsets (*i*NKT, NKT and T) are displayed for comparison. Genes are ordered by hierarchical clustering using Pearson correlation. Rows are mean-centered, normalized and local color scaling is used. (**b**) ANOVA category A_1 gene FC distributions for averaged splenic and IEL γδ T cell subsets compared to the averaged T cell subset displayed with individual genes displayed in rank-order along the y-

axis. *P* values represent significance of K-S test comparing the distribution of the ANOVA category A_1 geneset to that of all expressed genes. See Supplementary Table 16 for key to subset nomenclature.

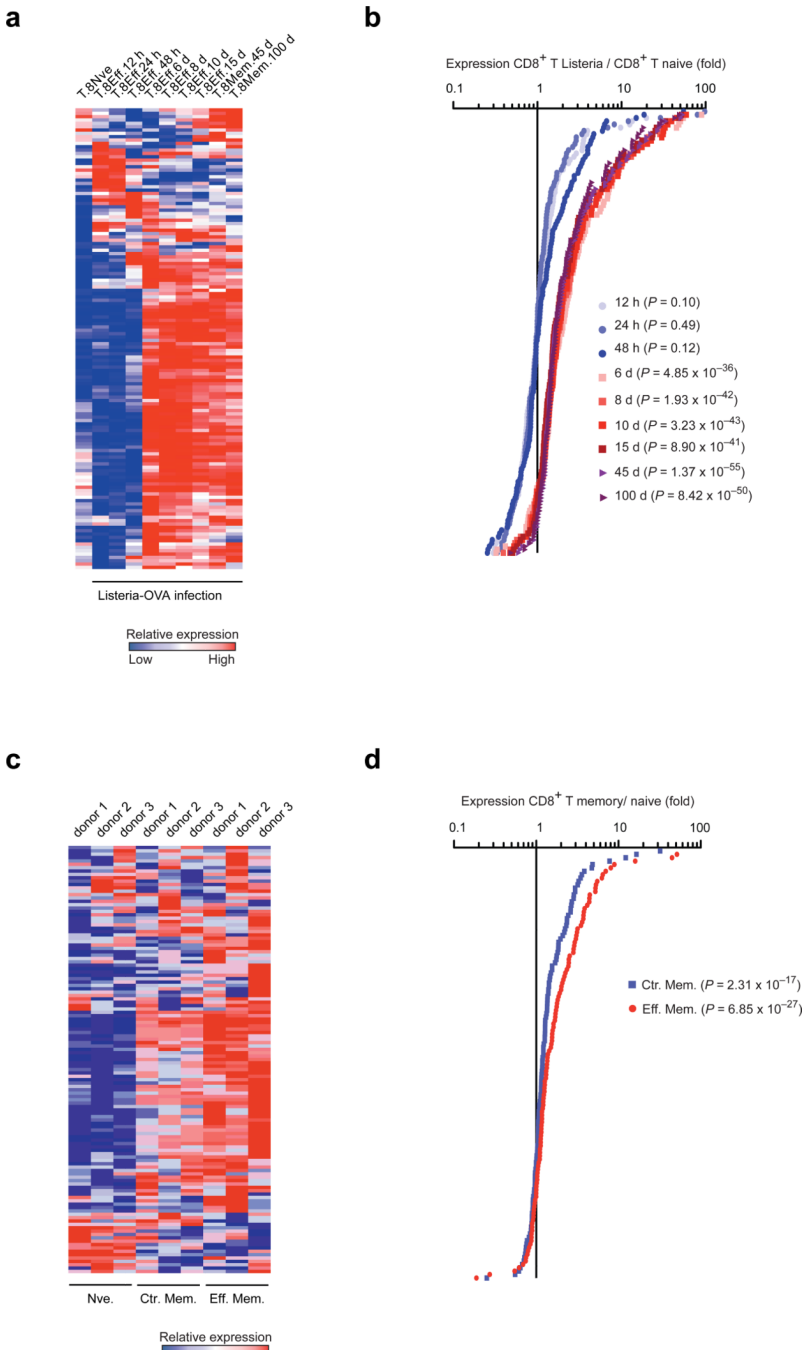

Fig. 7.

NK-*i*NKT shared gene programs are induced in activated CD8$^+$ T cells. (**a**) Relative expression of genes significantly upregulated in peripheral NK and *i*NKT cells (ANOVA category A_1) in antigen-specific CD8$^+$ T cells over the course of *Listeria*-OVA infection. See Supplementary Table 16 for key to subset nomenclature. (**b**) ANOVA category A_1 gene FC distributions for activated CD8$^+$ T cells compared to the naive control with individual genes displayed in rank-order along the y-axis. (**c**) Relative expression of the human gene homologs from ANOVA category A_1 in human peripheral blood CD8$^+$ T cell subsets from different donors[43]. (**d**) FC distributions for human memory CD8$^+$ T cells compared to the

naive control, using human homologs of the ANOVA category A_1 geneset. For all heatmaps, genes were ordered by hierarchical clustering using Pearson correlation. Rows are mean-centered and normalized, and relative scaling is used. For FC distribution plots, *P* values represent significance of K-S test comparing the distribution of the ANOVA category A_1 geneset to that of all expressed genes. Nve., naïve; Ctr. Mem., central memory; Eff. Mem., effector memory.

Table 1

Functional pathway enrichment for ANOVA category A_1 genes

Biological process	Genes present	Fold-enrichment	*P* value
BP00157: Natural killer cell mediated immunity	*Klrc2, Klrc3, Ifng, Klrb1f, Klrk1, Klrb1c, Klrc1*	13.70	1.11×10^{-5}
BP00255: Cytokine/chemokine mediated immunity	*Ccrl2, Ccr5, Ifng, Ccr2, Xcl1, Ccl5*	7.31	1.33×10^{-3}
BP00107: Cytokine and chemokine mediated signaling pathway	*Ccrl2, Il12rb2, Ccr5, Ifng, Ccr2, Inpp5d, Xcl1, Ccl5*	4.31	2.49×10^{-3}
BP00287: Cell motility	*Ccrl2, Coro2a, Dok2, Ccr5, Ccr2, Anxa1, Abi2, Dock5, Diap1*	3.68	2.98×10^{-3}
BP00148: Immunity and defense	*F2rl2, Klrc2, Klrc3, Fgr, Cysltr2, Adora2a, Tbx21, Klrk1, Ccl5, Ccrl2, Cd97, Il12rb2, Ifng, Irak2, Klf6, Il18rap, Lgals3, Gzmb, Sh2d2a, Ccr5, Ccr2, Klrb1f, Xcl1, Klrb1c, Klrc1, Sema4a*	2.11	3.57×10^{-4}
BP00102: Signal transduction	*F2rl2, Klrc2, S100a6, Gna15, Klrc3, Fgr, Adora2a, Cysltr2, Gpr65, Klrk1, Abi2, Itgb2, Ccl5, Prkx, Il12rb2, Cd97, Ccrl2, Coro2a, Zfp36l2, Plcb3, Ifng, Rhob, Fgl2, Rhoc, Fasl, Inpp5d, Ptprj, Irak2, Abr, Rxra, Anxa1, Ntng2, Smad3, Gem, Dock5, Arhgap26, Dusp5, Dok2, Dusp2, Ccr5, Rgs2, Rgs3, Ccr2, Chn2, Xcl1, Sema4a, Klrc1*	1.55	6.50×10^{-4}

Table 2

Functional pathway enrichment for ANOVA category C_1 genes

Biological process	Genes present	Fold-enrichment	*P* value
BP00116: JNK cascade	*Vav3, Dusp1, Jun, Jund, Junb, Cxcl10*	14.49	5.45×10^{-5}
BP00265: Oncogene	*Fos, Vav3, Lmo4, Jun, Jund, Etv5, Junb*	10.35	5.53×10^{-5}
BP00263: Inhibition of apoptosis	*Il4, bcl2a1d, bcl2a1c, cebpb, socs2, bcl2a1b, bcl2a1a*	8.11	2.13×10^{-4}
BP00179: Apoptosis	*Il4, bcl2a1d, bcl2a1c, cebpb, socs2, bcl2a1b, bcl2a1a, tnfsf14, nfkbia, gadd45b, emp1*	3.25	1.88×10^{-3}
BP00281: Oncogenesis	*Fos, vav3, lmo4, jun, jund, fam129a, etv5, junb, emp1*	3.13	7.59×10^{-3}
BP00111: Intracellular signaling cascade	*Il4, cap2, vav3, socs2, rab4a, nfkbia, junb, cxcl10, plk3, dusp1, jun, jund, gadd45b, tbkbp1, dusp6, rab27a*	2.78	5.12×10^{-4}
BP00040: mRNA transcription	*Dtx4, cebpb, lmo4, nfkbia, fosb, arntl, zbtb16, rora, junb, hic1, fos, npas2, nr1d1, gata3, jun, jund, dennd4c, thoc4, gfi1, rbpj, etv5, klf4*	1.76	9.97×10^{-3}