nature immunology

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] & The Immunological Genome Project Consortium[6]

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

The Immunological Genome (ImmGen) Project is a consortium of immunologists and computational biologists who aim, through the use of 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 cell (*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 cell lineages of the innate and adaptive immune systems. A subset of αβ T cells, *i*NKT cells express a semi-invariant T cell antigen receptor (TCR) that recognizes 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 antimicrobial immunity, tumor rejection and inflammation[3]. Like major histocompatibility complex (MHC)-restricted T cells, *i*NKT cells undergo thymic differentiation with somatic recombination[4], recognize self and foreign antigens[2,3], secrete cytokines characteristic of the T_H1, T_H2 and T_H17 subsets of helper T cells[5], and provide help to B cells[6].

The designation 'NKT' was coined to reflect expression of the NK cell marker NK1.1 (ref. 7). Although many other receptors expressed by NK cells ('NK cell receptors' (NKRs)) can also be expressed by *i*NKT cells[8], the validity of the designation 'NKT' has been called into question[9] because *i*NKT cells are developmentally more closely related to the T cell lineage than to the NK cell lineage and because NKRs are neither specific to *i*NKT cells nor expressed on all *i*NKT cells. Although a central role for TCR-mediated activation in *i*NKT cell biology is clear, notable parallels in the activity of NK cells and *i*NKT cells have nevertheless emerged. The homeostatic distribution and survival requirements of *i*NKT cells are similar to those of NK cells[10–13]. Also similar are their trafficking and activation kinetics. Both cell types constitutively express receptors for inflammatory chemokines, 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]. Furthermore, *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 cells and *i*NKT cells engage in similar bidirectional interactions with antigen-presenting cells during which inflammatory cytokines derived from the antigen-presenting cells potentiate the responses of NK cells and *i*NKT cells to surface ligands, and NK cells or *i*NKT cells, in turn, promote the maturation of antigen-presenting cells[23–28].

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

[1]Department of Medicine, Division of Rheumatology, Immunology and Allergy, Brigham and Women's Hospital, 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. [5]These authors contributed equally to this work. [6]A list of members and affiliations appears at the end of the paper. Correspondence should be addressed to M.B.B. (mbrenner@rics.bwh.harvard.edu).

Received 19 June; accepted 6 November; published online 2 December 2012; doi:10.1038/ni.2490

RESULTS

Developmental programs specific to *i*NKT cells

Like other T lymphocytes, *i*NKT cells mature in the thymus, where they diverge from MHC-restricted αβ T cells at the CD4$^+$CD8$^+$ double-positive (DP) stage. During maturation, *i*NKT cells subsequently undergo sequential stages of differentiation characterized by differences in expression of the activation markers CD44 and NK1.1 (ref. 4; **Supplementary Fig. 1**). To characterize the developmental transcriptional programs in *i*NKT cells relative to those that operate 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**). We found that 1,850, 155 and 697 genes had a change in expression of at least twofold from the DP stage to stage 1, from stage 1 to stage 2, and from stage 2 to stage 3, respectively (**Fig. 1b**,**c**). Transcripts modulated between thymic *i*NKT cell populations encoded molecules known to be involved in *i*NKT cell maturation[4], as well as several factors of unknown function (**Supplementary Tables 1–6**). At the DP branch point, we identified a subset of genes modulated selectively by *i*NKT cells at stage 1 but not by early-stage CD4$^+$ adaptive

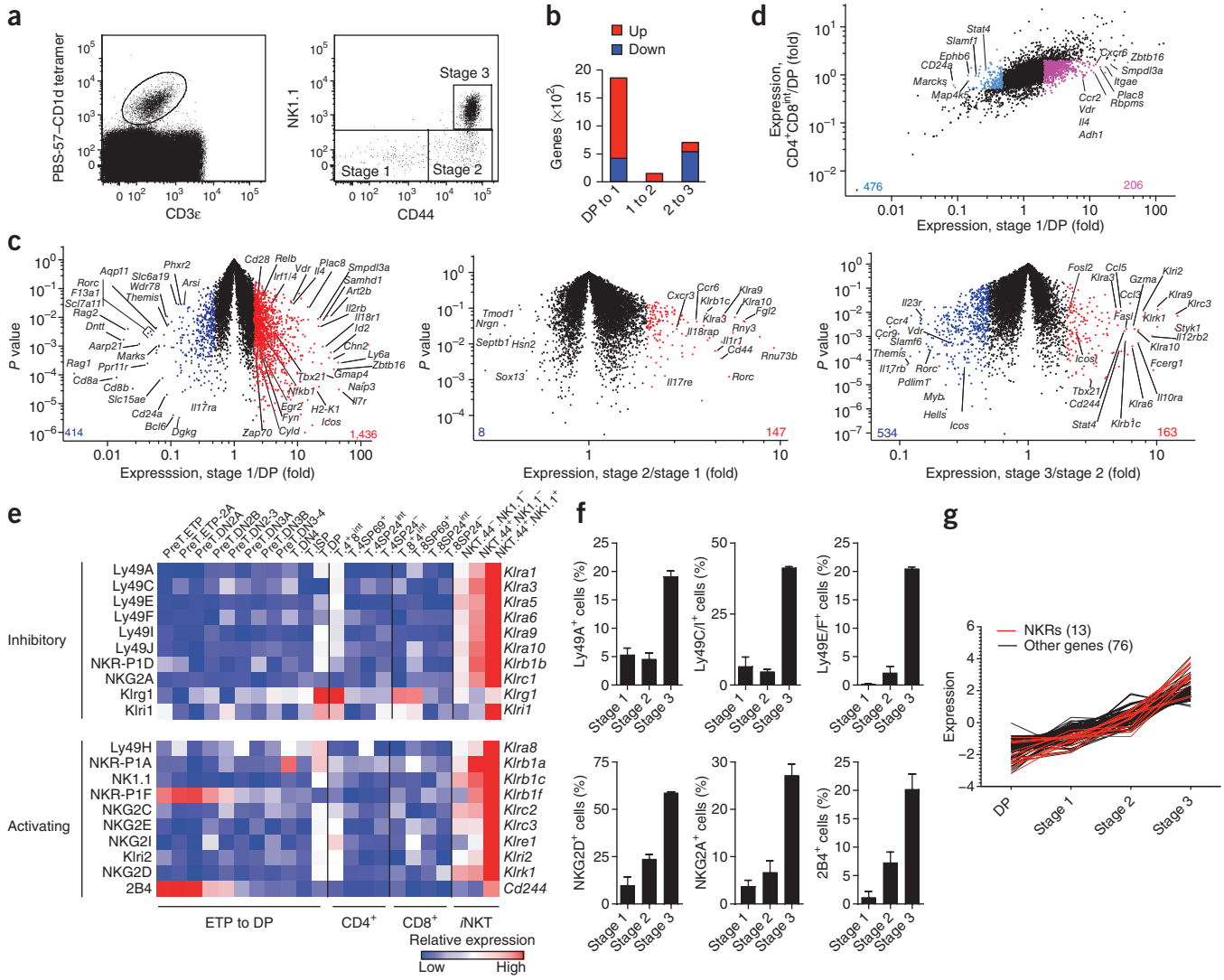


Figure 1 Upregulation of the expression of NKRs by *i*NKT cells at the end of thymic differentiation. (**a**) Identification of thymic *i*NKT cells by flow cytometry. PBS-57, α-galactosylceramide analog. (**b**) Quantification of genes upregulated (Up) or downregulated (Down) during the developmental transitions of *i*NKT cells. (**c**) Genes with a difference in expression (upregulated, red; downregulated, blue) in *i*NKT cells at stage 1 versus DP thymocytes (left), in *i*NKT cells at stage 2 versus *i*NKT cells at stage 1 (middle) or in *i*NKT cells at stage 3 versus *i*NKT cells at stage 2 (right); complete lists, **Supplementary Tables 1–6**. (**d**) Genes with a difference in expression in *i*NKT cells at stage 1 versus DP thymocytes; colors indicate genes regulated differently in *i*NKT cells but not in CD4$^+$CD8^{int} T cells relative to their regulation in DP thymocytes (upregulated, pink; downregulated, light blue; complete lists, **Supplementary Tables 7** and **8**). For **b**–**d**, only genes with expression above the detection limit in one or more subsets and a coefficient of variation of less than 0.5 in all subsets were considered. Each symbol represents a single gene. (**e**) Expression of genes encoding NKRs in differentiating thymocytes (subset nomenclature (top), **Supplementary Table 16**); results were $\log_2$-transformed, row-centered and locally color-scaled. (**f**) Surface expression of NKRs, determined by flow cytometry. (**g**) Expression of a cluster of genes encoding NKRs (*Klra3*, *Klra5*, *Klra6*, *Klra9*, *Klra10*, *Klrb1b*, *Klrb1c*, *Klrb1f*, *Klrc1*, *Klrc2*, *Klrc3*, *Klrd1*, *Klri2* and *Klrk1*; red) and other genes (black), derived from a Euclidean distance–based K-means clustering analysis (mean correlation, 0.958); genes were prefiltered for expression and for a change in expression of over twofold between any two subsets, and the results were $\log_2$-transformed. Data are combined from at least three independent experiments for each population, with cells pooled from three or more mice in each experiment (**a**–**e**,**g**), or are representative of at least two independent experiments (**f**; mean and s.e.m. of three mice).

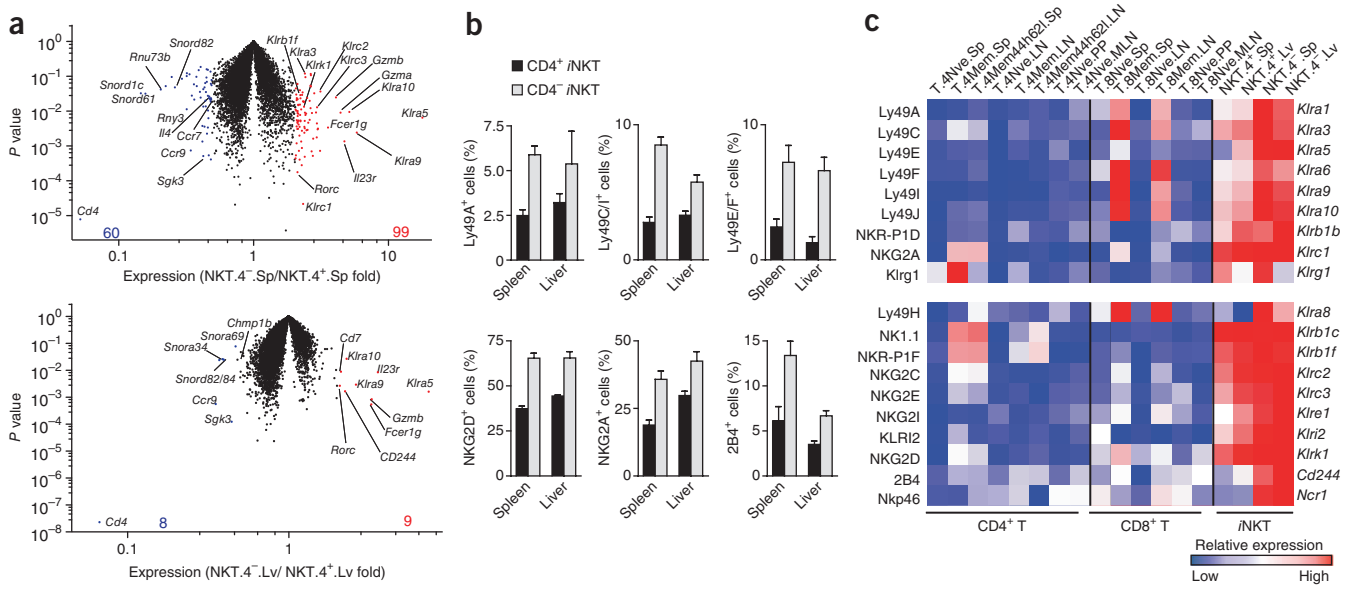
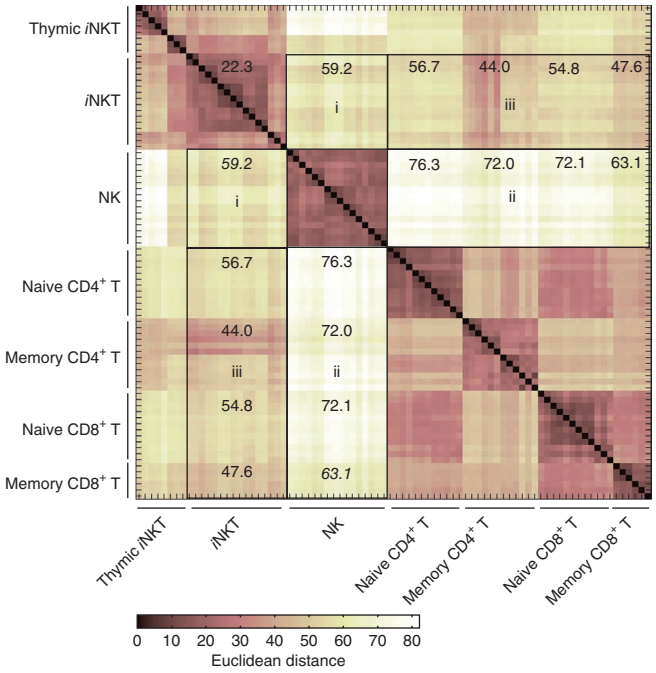


Figure 2 Expression of NKRs in peripheral CD4+ and CD4− *i*NKT cells. (**a**) Genes with a difference in expression in splenic CD4+ *i*NKT cells (NKT.4+.Sp) versus splenic CD4− *i*NKT cells (NKT.4−.Sp; top) or in liver CD4+ *i*NKT cells (NKT.4+.Lv) versus liver CD4− *i*NKT cells (NKT.4−.Lv; bottom); colors indicate genes upregulated (red) or downregulated (blue) in CD4− cells relative to their expression in CD4+ cells (change in expression of over twofold); only genes with expression above the detection limit in one or more subsets and a coefficient of variation of less than 0.5 in all subsets are included. Each symbol represents a single gene. (**b**) Surface expression of NKRs on spleen and liver CD4+ and CD4− *i*NKT cells, determined by flow cytometry. (**c**) Expression of genes encoding NKRs in peripheral MHC-restricted T cells and *i*NKT cell subsets (subset nomenclature (top), **Supplementary Table 16**); results were $\log_2$-transformed and gene-row-centered, and local color scaling was used. Data are combined from at least three independent experiments for each population, with cells pooled from three or more mice in each experiment (**a**,**c**), or are representative of at least two independent experiments (**b**; mean and s.e.m. of three mice).

T cells (CD4+CD8^{int} cells; **Fig. 1d**). Although some of these genes, such as *Zbtb16* (which encodes the transcription factor PLZF) and *Vdr* (which encodes the receptor for vitamin D) encoded molecules known to have important functions in specification of the *i*NKT cell lineage[29–31], many encoded molecules not yet been linked to the development of *i*NKT cells (**Supplementary Tables 7** and **8**). At the final maturation stage of *i*NKT cells, many of the genes with the greatest upregulation encoded molecules that belong to the killer lectin receptor family, which includes activating and inhibitory NKRs (**Fig. 1c**, bottom). By the final stage of ontogeny, *i*NKT cells had expression of transcripts encoding several NKRs similar in amount to that of splenic NK cells (**Supplementary Fig. 2a**). In addition, upregulation of genes encoding NKRs occurred 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 cytometry of thymic *i*NKT cells confirmed the increase in expression 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 followed the same expression kinetics as that of genes encoding NKRs over the course of the DP–to–stage 3 transition (**Fig. 1g**), which suggested that progressive upregulation of the expression of NKRs may be part of a broader gene program. Together these data comprehensively characterized shared and distinct gene-expression changes that occurred in maturing *i*NKT cells in the broader context of αβ T cell development.



Transcriptional signatures of peripheral *i*NKT cells

In both mice and humans, CD4+ and CD4− *i*NKT cells have been reported to secrete different types of cytokines[5,32,33]. In addition, *i*NKT cell subsets

Figure 3 The global transcriptional relationship between NK and *i*NKT cells is similar in magnitude to the relationship between T cells and *i*NKT cells. Euclidean-distance matrix calculated with the 15% most variable probes; numbers in outlined areas indicate average Euclidean distance between intersecting subsets (i–iii indicate areas of interest). Only genes with mean expression values above 120 were included; results were $\log_2$-transformed and mean-centered. Data are combined from at least three independent experiments for each population, with cells pooled from three or more mice in each experiment.

from the liver have 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. Whereas 159 genes were expressed differently in CD4$^+$ versus CD4$^-$ *i*NKT cells in the spleen, and 261 genes were expressed differently in the corresponding subsets in the lung, only 17 transcripts varied in expression in CD4$^+$ versus CD4$^-$ hepatic *i*NKT subsets (**Fig. 2a**, **Supplementary Fig. 3** and **Supplementary Tables 9–14**). That suggested 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 genes encoding NKRs were among those most modulated, with higher expression in CD4$^-$ subsets than in CD4$^+$ subsets (**Fig. 2a** and **Supplementary Fig. 3**). In liver and spleen, this reflected a greater frequency of NKR$^+$ cells among CD4$^-$ *i*NKT cells than among CD4$^+$ *i*NKT cells, as determined by flow cytometry (**Fig. 2b**). The expression of NKR-encoding mRNA in peripheral *i*NKT cells, although lower than that in thymic *i*NKT cells at stage 3, was maintained at higher amounts than that in 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 showed that CD4$^-$ populations were more transcriptionally similar to each other than were CD4$^+$ *i*NKT cells, across tissues. A small number of genes were expressed differently in different tissues regardless of CD4 expression (**Supplementary Fig. 4**), which supported the idea that *i*NKT cells subsets may serve organ-specific functions.

Transcriptional patterns shared by NK cells and *i*NKT cells

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 transcriptional program shared by NK cells and NKT cells than is now appreciated. To assess the extent of transcriptional relatedness between *i*NKT cells and NK cells at a global level, and to compare the NK cell–*i*NKT cell relationship and the relationship of *i*NKT cells to T cells, we calculated Euclidean distances between subsets of steady-state NK cells, T cells and *i*NKT cells. We determined Euclidean distance, a measure of the similarity of the gene-expression patterns of pairwise compared subsets, using the 15% of gene probes with the greatest variability among the subsets analyzed (presented as a matrix in **Fig. 3**). Mature *i*NKT cells and NK cells showed a high degree of similarity in gene expression (**Fig. 3**, area i), in contrast to the greater distances between all MHC-restricted T cells and NK cells (**Fig. 3**, area ii). Although *i*NKT cells had a somewhat closer relationship to memory CD8$^+$ T cells and to certain memory CD4$^+$ T cells than to NK cells, we sorted memory CD4$^+$ T cells through the use of markers with similar expression by *i*NKT cells (CD44$^+$ and CD62L^{lo}) and thus they probably included a substantial proportion of *i*NKT cells. The average Euclidean distance between *i*NKT cells and NK cells was only slightly larger than the distance that separated *i*NKT cells from naive CD4$^+$ T cells and CD8$^+$ T cells (**Fig. 3**, area iii). Those relative relationships were maintained when we analyzed all genes and not just the 15% of genes with the greatest variability (data not shown). To determine whether the NK cell–*i*NKT cell relationship was dependent on the shared expression of NKRs, we recalculated the distance matrix after removing NKRs and related molecules from the data. The outcome of the analysis remained essentially unchanged (**Supplementary Fig. 5**). Thus, the transcriptional relationship between NK cells and *i*NKT cells was not limited to the shared expression of NKRs. Together these data supported the idea of an unexpectedly substantial transcriptional

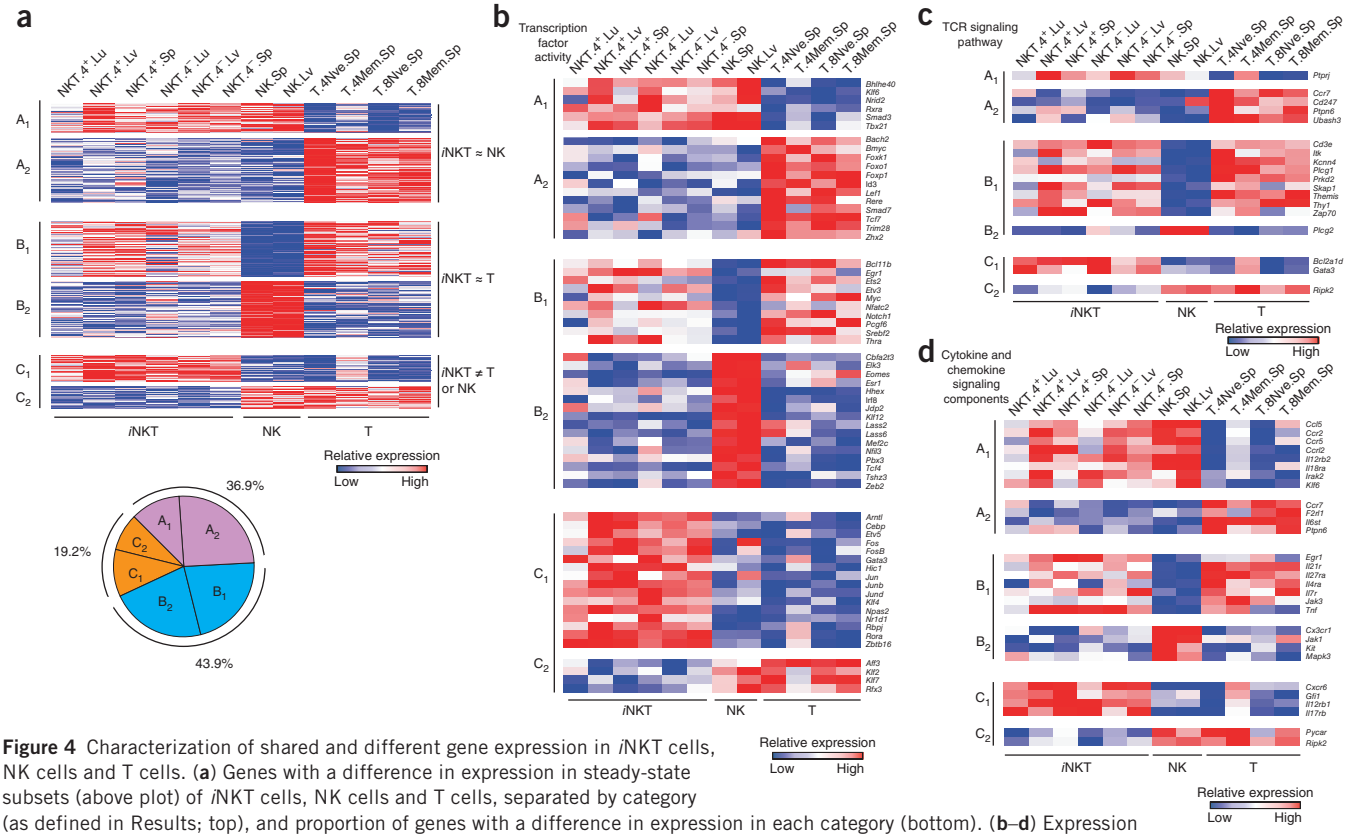


Figure 4 Characterization of shared and different gene expression in *i*NKT cells, NK cells and T cells. (**a**) Genes with a difference in expression in steady-state subsets (above plot) of *i*NKT cells, NK cells and T cells, separated by category (as defined in Results; top), and proportion of genes with a difference in expression in each category (bottom). (**b**–**d**) Expression of genes (categorized as in **a**) in relevant selected functional gene groups. In all heat maps, rows are mean-centered and normalized, and local scaling is used (subset nomenclature, **Supplementary Table 16**). Data are combined from at least three independent experiments for each population, with cells pooled from three or more mice in each experiment.

Table 1 Functional pathway–enrichment analysis of genes in category A_1

Biological process	Genes present	Enrichment (fold)	*P* value
NK cell–mediated immunity (BP00157)	*Klrc2*, *Klrc3*, *Ifng*, *Klrb1f*, *Klrk1*, *Klrb1c*, *Klrc1*	13.70	1.11×10^{-5}
Cytokine- or chemokine-mediated immunity (BP00255)	*Ccrl2*, *Ccr5*, *Ifng*, *Ccr2*, *Xcl1*, *Ccl5*	7.31	1.33×10^{-3}
Cytokine- and chemokine-mediated signaling pathway (BP00107)	*Ccrl2*, *Il12rb2*, *Ccr5*, *Ifng*, *Ccr2*, *Inpp5d*, *Xcl1*, *Ccl5*	4.31	2.49×10^{-3}
Cell motility (BP00287)	*Ccrl2*, *Coro2a*, *Dok2*, *Ccr5*, *Ccr2*, *Anxa1*, *Abi2*, *Dock5*, *Diap1*	3.68	2.98×10^{-3}
Immunity and defense (BP00148:)	*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}
Signal transduction (BP00102)	*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}

Functional biological process–enrichment analysis of genes in category A_1; designations in parentheses (left column) indicate module identifiers (PANTHER classification system). Enrichment and *P* values (from a modified Fisher's exact test) were calculated with DAVID software; enrichment indicates the representation of genes in category A_1 in a given biological process relative to representation that would occur by chance.

relationship between *i*NKT cells and NK cells that was close in magnitude to that between *i*NKT cells and naive T cells.

Shared and distinct *i*NKT cell programs

We next sought to identify the specific genes concordantly regulated by *i*NKT cells, NK cells and T cells. For this purpose, we used one-way analysis of variance (ANOVA) to compare the transcriptomes of peripheral steady-state *i*NKT cells, NK cells and naive and memory $CD4^+$ or $CD8^+$ T cells. We excluded from the analysis genes with low variability over the entire ImmGen Project data set. We found that 20.2% (1,192) of the remaining genes had differences in expression in the three groups ($P < 0.05$ (Bonferroni corrected)). We further classified the modulated genes into the following six categories on the basis of common patterns in two of the three subsets: genes with similar expression in NK cells and *i*NKT cells that was higher (A_1) or lower (A_2) than that in T cells; genes with similar expression in *i*NKT cells and T cells that was higher (B_1) or lower (B_2) than that in NK cells; and genes with expression that was higher (C_1) or lower (C_2) in *i*NKT cells than in both T cells and NK cells (**Fig. 4a**). Genes encoding many transcription factors, TCR signaling components and cytokine- or chemokine-related molecules known to be expressed differently by NK cells, *i*NKT cells and T cells partitioned as expected in these categories (**Fig. 4b–d**). For example, genes encoding the transcription factor PLZF (*Zbtb16*; **Fig. 4b**), the chemokine receptor CXCR6 (*Cxcr6*; **Fig. 4c**), and a component of the high-affinity receptor for interleukin 12 (IL-12 (*Il12rb1*); **Fig. 4d**), all known to be have higher expression in *i*NKT cells than in resting T cells or NK cells[29,30,35,36], partitioned into category C_1 (higher expression in *i*NKT cells than in NK cells or T cells). The genes encoding the transcription factor T-bet (*Tbx21*; **Fig. 4b**), as well as the IL-12RB2 receptor for IL-12 (*Il12rb2*; **Fig. 4d**), known to be upregulated in NK cells and *i*NKT cells[37], partitioned into category A_1 (higher expression in *i*NKT cells and NK cells than in T cells). Genes encoding TCR signaling components such as CD3ε (*Cd3e*), ITK (*Itk*), PLC-γ1 (*Plcg1*) and Zap70 (*Zap70*) segregated as anticipated into category B_1 (higher expression in *i*NKT cells and T cells than in NK cells; **Fig. 4c**,**d**). Functional biological process–enrichment analysis of genes in category A_1 with software from the DAVID bioinformatics database[38] (as described in Online Methods) demonstrated significant enrichment for genes encoding molecules with effector functions, including NK cell–mediated immunity, chemokine or cytokine responses, signal transduction, and cell motility (**Table 1**). The group of transcripts upregulated in *i*NKT cells relative to their expression in NK cells and MHC-restricted T cells (category C_1) showed enrichment for genes encoding molecules involved in distinct biological processes, including proliferation and survival, which probably reflected the uniquely activated phenotype of steady-state *i*NKT cells (**Table 2**).

We found that the number of genes with similar expression patterns in the *i*NKT and NK lineages (440 genes (36.91%); categories A_1 and A_2) was similar to the number of genes with similar expression patterns in *i*NKT cells and T cells (523 genes (43.88%); categories B_1 and B_2). In addition, about one fifth of the genes that 'passed' the ANOVA were regulated uniquely in *i*NKT cells (229 genes (19.21%); categories C_1 and C_2; **Fig. 4a** and **Supplementary Table 15**). Thus, in addition

Table 2 Functional pathway–enrichment analysis of genes in category C_1

Biological process	Genes present	Enrichment (fold)	*P* value
Jnk cascade (BP00116)	*Vav3*, *Dusp1*, *Jun*, *Jund*, *Junb*, *Cxcl10*	14.49	5.45×10^{-5}
Oncogene (BP00265)	*Fos*, *Vav3*, *Lmo4*, *Jun*, *Jund*, *Etv5*, *Junb*	10.35	5.53×10^{-5}
Inhibition of apoptosis (BP00263)	*Il4*, *Bcl2a1d*, *Bcl2a1c*, *Cebpb*, *Socs2*, *Bcl2a1b*, *Bcl2a1a*	8.11	2.13×10^{-4}
Apoptosis (BP00179)	*Il4*, *Bcl2a1d*, *Bcl2a1c*, *Cebpb*, *Socs2*, *Bcl2a1b*, *Bcl2a1a*, *Tnfsf14*, *Nfkbia*, *Gadd45b*, *Emp1*	3.25	1.88×10^{-3}
Oncogenesis (BP00281)	*Fos*, *Vav3*, *Lmo4*, *Jun*, *Lund*, *Fam129a*, *Etv5*, *Junb*, *Emp1*	3.13	7.59×10^{-3}
Intracellular signaling cascade (BP00111:)	*Il4*, *Cap2*, *Vav3*, *Socs2*, *Rab4a*, *Nfkbia*, *junb*, *Cxcl10*, *Plk3*, *Dusp1*, *Jun*, *Jund*, *Gadd45b*, *Tbkbp1*, *Dusp6*, *Rab27a*	2.78	5.12×10^{-4}
mRNA transcription (BP00040)	*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}

Functional biological process–enrichment analysis of genes in category C_1 (presented as in **Table 1**). Jnk, kinase.

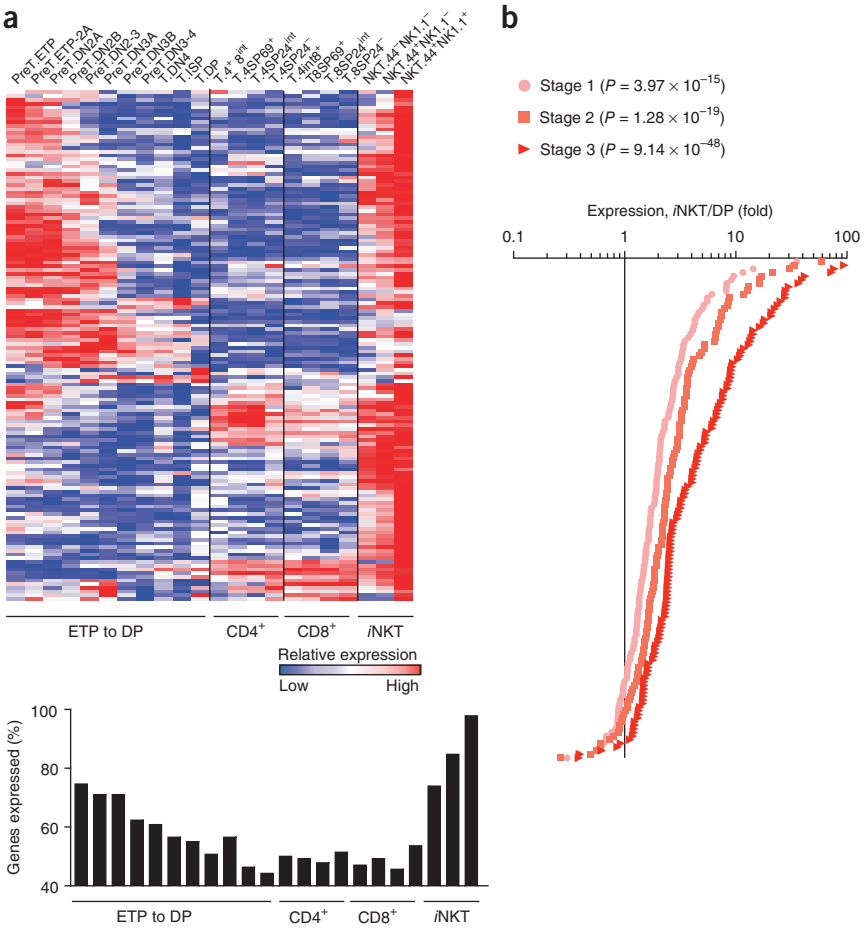


Figure 5 Transcriptional programs shared by NK cells and *i*NKT cells are acquired during thymic maturation of *i*NKT cells. (**a**) Expression of genes significantly upregulated in peripheral NK cells and *i*NKT cells (category A_1) over the course of the thymic maturation of T cells and *i*NKT cells, ordered by hierarchical clustering with Pearson correlation (top; presented as in **Fig. 4b–d**); and frequency of genes exceeding the threshold for expression over the course of thymic maturation (below). (**b**) Distribution of genes in category A_1 in *i*NKT cells at stages 1, 2 and 3 relative to that of DP thymocytes (according to change in expression); genes are in rank order along the vertical axis. *P* values, category A_1 versus all expressed genes (Kolmogorov-Smirnov test). Data are combined from at least three independent experiments for each population, with cells pooled from three or more mice in each experiment.

to having a distinct gene-expression program, *i*NKT cells shared a transcriptional program with steady-state NK cells that was extensive and similar in magnitude to that shared by *i*NKT cells and other αβ T cells, consistent with the Euclidean-distance analysis.

Thymic induction of programs shared by NK cells and *i*NKT cells

To determine at what point during thymic development the transcriptional program shared by mature *i*NKT cells and NK cells was induced, we analyzed by hierarchical clustering the expression patterns of genes in category A_1 over the course of the differentiation of MHC-restricted T cells and *i*NKT cells. Although approximately 75% of the shared genes upregulated in NK cells and *i*NKT cells were expressed in early thymic precursors, these genes were then largely shut down or downregulated by the DP thymocyte stage and were not reactivated in differentiating $CD4^+$ or $CD8^+$ T cells. In contrast, differentiating *i*NKT cells upregulated those same genes, so that thymic *i*NKT cells at stage 3 expressed more than 90% of the genes (**Fig. 5a**). Consistent with that observation, the distribution of genes in category A_1, calculated by comparison of the expression of each gene in developing *i*NKT cells with that in their DP progenitors, showed an increasing shift toward higher values with *i*NKT cell maturation. All three distributions were significantly different from the baseline distribution for all expressed genes, as determined by the Kolmogorov-Smirnov test (**Fig. 5b**). In contrast, the distribution calculated by comparison of naive splenic $CD4^+$ or $CD8^+$ T cells with DP cells showed no significant difference between the gene set of category A_1 and all expressed genes. In contrast to genes in category A_1, the genes significantly downregulated in mature NK cells and *i*NKT cells relative to their expression in T cells (category A_2) were not elicited as much in differentiating thymic *i*NKT cells as in $CD4^+$ or $CD8^+$ T cells (**Supplementary Fig. 6**). These data extended our earlier observation that *i*NKT cells acquired NKR expression at the end of development and indicated that a large part of the transcriptional program shared by *i*NKT cells and NK cells was acquired during thymic maturation.

The *i*NKT cell programs in γδ T cells

We next sought to determine if other lymphocytes with innate features might also have the transcriptional program noted above. Innate lymphocyte subsets include γδ T cells that can be categorized on the basis of their use of the TCR γ-chain variable region (V_γ) and δ-chain variable region (V_δ). The subset bearing $V_\gamma2$ 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) secrete mainly T_H1 or T_H2 cytokines[39]. Distinct intraepithelial lymphocyte (IEL) $CD8\alpha\alpha^+$ γδ T cell populations have an effector phenotype similar to that of *i*NKT cells[40,41].

The genes shared by NK cells and *i*NKT cells (category A_1) had relatively low expression in splenic $CD44^-$ γδ T cells, similar to that observed in resting T cells. A subset of those genes was upregulated in splenic $CD44^+$ $V_\gamma2^+$ and $V_\gamma2^-$ cells. Notably, we found that the expression of a large cluster of genes in category A_1 was higher in all tissue-resident IEL γδ T cells than in the reference NK and *i*NKT populations (**Fig. 6a**). Distributions made by comparison of the expression of genes in category A_1 by splenic or IEL γδ T cells and T cells showed significant enrichment for overexpressed genes (**Fig. 6b**). IEL γδ T cells expressed an even smaller proportion of genes in category A_2 (downregulated in NK cells and *i*NKT cells relative to their expression in T cells) than did *i*NKT cells, although most of those genes were expressed in splenic γδ T cells (data not shown). Although it was less prominent than for the shared NK cell–*i*NKT cell program (category A_1), a portion of the genes upregulated in *i*NKT cells but not in T cell or NK cell subsets (category C_1) also had relatively high expression in γδ T cell subsets (**Supplementary Fig. 7a**). Together these data suggested that most of the gene program shared by NK cells and *i*NKT cells, as well as part of the program of genes upregulated differently in *i*NKT cells, was also used by populations of γδ T cells at steady state.

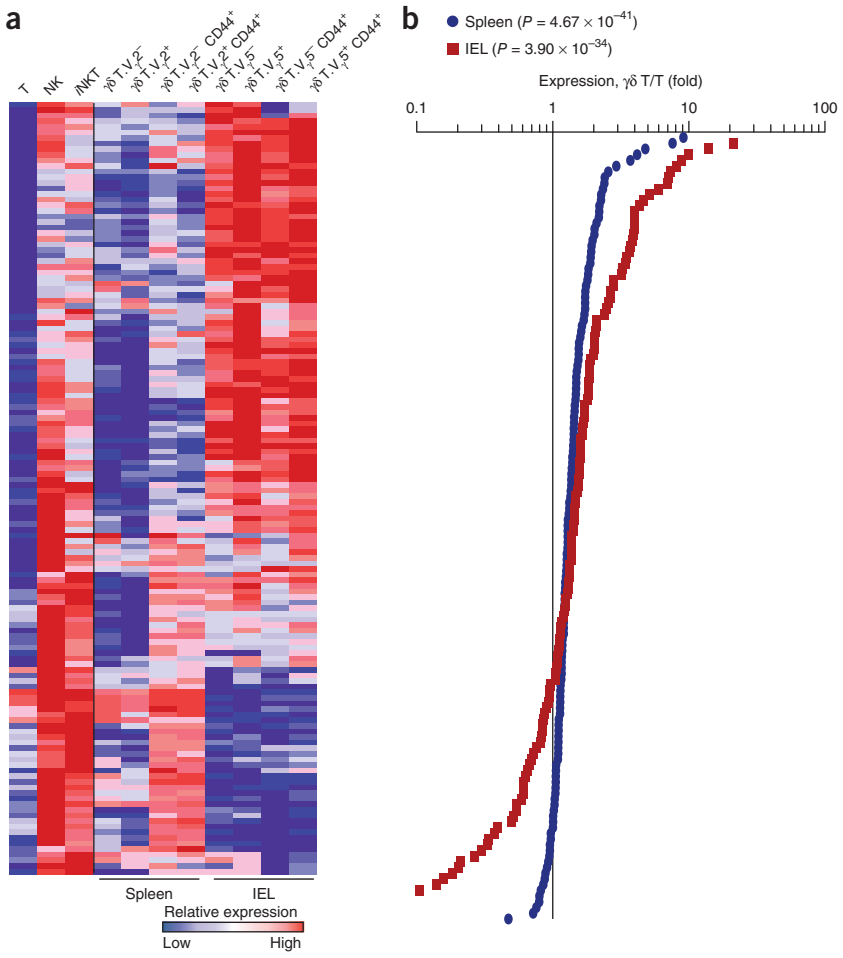


Figure 6 Activated splenic γδ T cells and IEL γδ T cells express genes in the program shared by NK cells and *i*NKT cells. (**a**) Expression of genes significantly upregulated in peripheral NK cells and *i*NKT cells relative to their expression in T cells (category A_1) in various splenic and IEL γδ T cell subsets (top; presented as in **Fig. 4b–d**). Left three columns, averaged expression values for the subsets used to define the ANOVA categories (*i*NKT cells, NKT cells and T cells). (**b**) Distribution of genes in category A_1 for averaged expression values for splenic and IEL γδ T cell subsets compared with the averaged values for the T cell subset (presented as in **Fig. 5b**). Data are combined from at least three independent experiments for each population, with cells pooled from three or more mice in each experiment.

The NK cell–*i*NKT cell program in CD8$^+$ T cells

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

DISCUSSION

Although they express a functional TCR, *i*NKT cells do not fit the paradigm of adaptive T cell immunity[3,14]. How the innate features of *i*NKT cells are acquired and how the transcriptional programs involved relate to those active in lymphocytes with similar functional attributes remains incompletely defined.

The transcriptional programs we found to be operating during the maturation of *i*NKT cells were consistent with published reports and included genes encoding PLZF, the vitamin D receptor, T-bet, components of the transcription factor NF-κB and GTPase Ras–mitogen-activated protein kinase (MAPK) pathways, and NKRs[4,37,44,45]. By comparing developing *i*NKT cells with differentiating MHC-restricted T cells, we have defined transcriptional programs of genes expressed specifically by *i*NKT cells at stage 1 shortly after the DP branch point. We have thus identified a large number of genes encoding molecules not previously known to affect the biology of *i*NKT cells but that probably specifically modulate the thymic maturation of *i*NKT cells.

Many genes reported to be involved in the development of *i*NKT cell were not expressed in an *i*NKT cell specific–manner. For example, genes encoding members of the NF-κB and Ras-MAPK pathway, such as the p50 and RelB subunits of NF-kB, Ras-MAPK and Egr2, were upregulated in both *i*NKT and non-*i*NKT thymocyte populations[4] (data not shown). Other genes encoding molecules known to affect the ontogeny of *i*NKT cells, such as the transcription factor Bcl-11b[46] and the chromatin modifier Med1 (ref. 47), showed little or no transcriptional variation over the course of *i*NKT cell differentiation. The functional regulation of those 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, which are developmentally distant relatives. We hypothesized that NK cells and *i*NKT cells share a broader transcriptional program than is now appreciated. Our analyses showed that the transcriptional patterns that *i*NKT cells shared with NK cells made up nearly

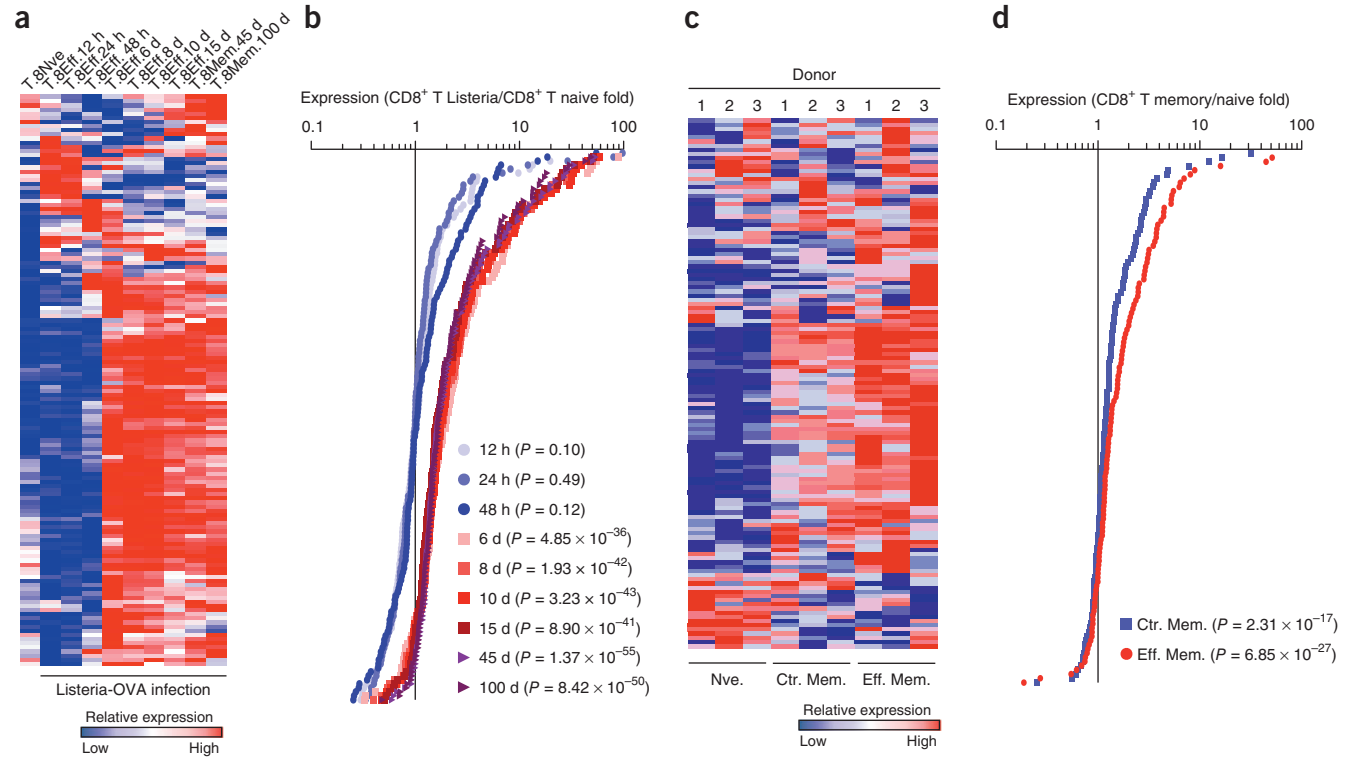


Figure 7 Programs shared by NK cells and *i*NKT cells are induced in activated CD8⁺ T cells. (**a**) Expression of genes significantly upregulated in peripheral NK cells and *i*NKT cells (category A_1) relative to their expression in antigen-specific CD8⁺ T cells over the course of infection with OVA-expressing *L. monocytogenes* (*Listeria*-OVA; presented as in **Fig. 4b–d**). (**b**) Distribution of genes in category A_1 for activated CD8⁺ T cells versus that for naive control cells (presented as in **Fig. 5b**). (**c**) Expression of the human homologs of genes in category A_1 in human peripheral blood CD8⁺ T cell subsets from various donors (1–3)[43] (presented as in **Fig. 4b–d**). (**d**) Distribution of the gene set of category A_1 for human memory CD8⁺ T cells versus that for the naive control cells, with human homologs of the genes in category A_1 (presented as in **Fig. 5b**). Data are combined from at least three independent experiments for each population, with cells pooled from three or more mice in each experiment.

as large a part of the *i*NKT transcriptome as those shared with MHC-restricted T cells. Furthermore, we found that the program shared by NK cells and *i*NKT cells was also active in steady-state γδ T cells subsets similarly poised for rapid, innate-like responsiveness. The NK cell–*i*NKT cell program was also induced in adaptive αβ T cells, but only days after antigen-specific stimulation. These data suggested that the program shared by NK cells and *i*NKT cells represents a core effector program operational in lymphocytes of distinct lineages, consistent with the many functional abilities that *i*NKT cells, NK cells and activated MHC-restricted T cells have in common.

The factors that regulate the expression of genes in this shared effector program remain to be fully defined. Nevertheless, our data have provided several clues. Acquired largely at the end of thymic *i*NKT cell maturation, the shared program mirrored the cells' acquisition of NKR expression, which is driven by the transcription factor T-bet both in NK cells and *i*NKT cells[37]. Thus, it is likely that T-bet, whose transcript was among those significantly upregulated in both NK cells and *i*NKT cells relative to its expression in resting T cells, has an important role in eliciting and maintaining at least part of the shared effector lymphocyte program. T-bet-deficient *i*NKT cells have lower mRNA and protein expression of IFN-γ, granzyme B, Fas ligand, CCR5 and CD38, all molecules that are part of the shared program[44]. IL-15, reported to act upstream of T-bet in maturing *i*NKT cells[45], is also probably involved in the acquisition of this program by *i*NKT cells. Furthermore, IL-15 and T-bet are important for the regulation of 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 have an important role in inducing the expression of genes in the shared effector program in several cell types.

We have also identified genes encoding many previously unappreciated transcriptional regulators with similar expression patterns in NK cells and *i*NKT cells that may help regulate the core effector program in innate lymphocytes. For example, the expression of genes encoding the transcription factor Bhlhe40, a regulator of the circadian rhythm[50] with immunomodulatory functions in CD4⁺ T cells[51,52], as well as Smad3, a transcription factor important for the 'tuning' of lymphocyte activation mediated by transforming growth factor-β, was upregulated in both NK cells and *i*NKT cells relative to their expression in resting T cells. Also, the expression of genes encoding many other factors was downregulated in NK cells and *i*NKT cells. Among those were *Lef1* and *Tcf7*, which encode transcription factors downstream of the Wnt signaling pathway that are important for the establishment and maintenance of T cell identity[53]; Wnt signaling is typically repressed after the activation of T cells[54,55]. The relatively low expression of *Tcf7* and *Lef1* in NK cells and *i*NKT cells during development is consistent with acquisition of the terminally differentiated effector phenotype of these cells at steady state.

Approximately 20% of the genes with significantly different expression in NK cells, *i*NKT cells and T cells had transcriptional patterns specific to *i*NKT cells. As expected, the transcription factors PLZF and GATA-3 were among the factors encoded by these genes. The group of genes 'preferentially' upregulated in *i*NKT cells (category C_1) showed enrichment for those encoding molecules involved in cellular activation and survival programs. For example, genes encoding several

members of the AP-1 family of transcription factors (*Jun*, *Junb*, *Jund*, *Fos*, *Fosb* and *Cebpb*), which are normally induced in cells only after activation, had relatively high expression in *i*NKT cells. Constitutive expression of AP-1 transcription factors by *i*NKT cells is consistent with their poised effector phenotype.

Together our data offer a new view of *i*NKT cells and their relationships with other lymphocyte lineages. Using the Immunological Genome Project database, we have identified gene-expression programs that both link to *i*NKT cells and differentiate *i*NKT cells from other lymphocytes of the innate and adaptive immune systems. Our data have identified extensive genetic modules shared with NK cells and other innate-like T cells that were also elicited several days after activation in adaptive T cells. By defining both distinct and shared transcriptional programs in *i*NKT cells, our data open avenues for future research and bring into clearer focus how lymphocyte populations that differ substantially in their ontogeny can ultimately serve similar effector functions through modular expression of genes in similar transcriptional programs.

METHODS

Methods and any associated references are available in the online version of the paper.

Accession codes. GEO: microarray data, GSE15907.

Note: Supplementary information is available in the online version of the paper.

ACKNOWLEDGMENTS

We thank the tetramer facility of the US National Institutes of Health for ongoing support; and S. Raychaudhuri, X. Hu and H. Li for advice, discussions and technical assistance. Supported by the US National Institutes of Health (R01AI063428 to M.B.B., T32AI007306 to P.J.B, and R24AI072073 to the ImmGen Project consortium).

AUTHOR CONTRIBUTIONS

N.R.C. primarily wrote the manuscript, conceived of and did experiments and analyzed the data; P.J.B. contributed substantially to the manuscript, conceived of and did experiments and analyzed the data; T.S. conceived of and did experiments and analyzed the data; G.F.W. did experiments; M.B. and J.K. assisted with experimental design and interpretation of the data; M.B.B. substantially contributed to the manuscript and supervised all experimental design, performance and data analysis; and The ImmGen Project Consortium contributed to experimental design and data collection.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at http://www.nature.com/doifinder/10.1038/ni.2490.
Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

1. Heng, T.S.P. *et al.* The Immunological Genome Project: networks of gene expression in immune cells. *Nat. Immunol.* **9**, 1091–1094 (2008).
2. Cohen, N.R., Garg, S. & Brenner, M.B. Antigen presentation by CD1 lipids, T cells, and NKT cells in microbial immunity. *Adv. Immunol.* **102**, 1–94 (2009).
3. Bendelac, A., Savage, P.B. & Teyton, L. The biology of NKT cells. *Annu. Rev. Immunol.* **25**, 297–336 (2007).
4. Godfrey, D.I., Stankovic, S. & Baxter, A.G. Raising the NKT cell family. *Nat. Immunol.* **11**, 197–206 (2010).
5. Coquet, J.M. *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. USA* **105**, 11287–11292 (2008).
6. Cerundolo, V., Silk, J.D., Masri, S.H. & Salio, M. Harnessing invariant NKT cells in vaccination strategies. *Nat. Rev. Immunol.* **9**, 28–38 (2009).
7. Makino, Y., Kanno, R., Ito, T., Higashino, K. & Taniguchi, M. Predominant expression of invariant Vα14$^+$ TCR α chain in NK1.1$^+$ T cell populations. *Int. Immunol.* **7**, 1157–1161 (1995).
8. Bendelac, A. Mouse NK1$^+$ T cells. *Curr. Opin. Immunol.* **7**, 367–374 (1995).
9. Godfrey, D.I., MacDonald, H.R., Kronenberg, M., Smyth, M.J. & Van Kaer, L. NKT cells: what's in a name? *Nat. Rev. Immunol.* **4**, 231–237 (2004).
10. Grégoire, C. *et al.* The trafficking of natural killer cells. *Immunol. Rev.* **220**, 169–182 (2007).
11. Matsuda, J.L. *et al.* Tracking the response of natural killer T cells to a glycolipid antigen using CD1d tetramers. *J. Exp. Med.* **192**, 741–754 (2000).
12. Cooper, M.A. *et al. In vivo* evidence for a dependence on interleukin 15 for survival of natural killer cells. *Blood* **100**, 3633–3638 (2002).
13. Matsuda, J.L. *et al.* Homeostasis of V alpha 14i NKT cells. *Nat. Immunol.* **3**, 966–974 (2002).
14. Brigl, M. & Brenner, M.B. How invariant natural killer T cells respond to infection by recognizing microbial or endogenous lipid antigens. *Semin. Immunol.* **22**, 79–86 (2010).
15. Lanier, L.L. Evolutionary struggles between NK cells and viruses. *Nat. Rev. Immunol.* **8**, 259–268 (2008).
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.* **41**, 1913–1923 (2011).
17. Kawamura, T. *et al.* NKG2A inhibits invariant NKT cell activation in hepatic injury. *J. Immunol.* **182**, 250–258 (2009).
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.* **167**, 4180–4186 (2001).
19. Ota, T. *et al.* IFN-γ-mediated negative feedback regulation of NKT-cell function by CD94/NKG2. *Blood* **106**, 184–192 (2005).
20. Brennan, P.J. *et al.* Invariant natural killer T cells recognize lipid self antigen induced by microbial danger signals. *Nat. Immunol.* **12**, 1202–1211 (2011).
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* **27**, 597–609 (2007).
22. Salio, M. *et al.* Modulation of human natural killer T cell ligands on TLR-mediated antigen-presenting cell activation. *Proc. Natl. Acad. Sci. USA* **104**, 20490–20495 (2007).
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.* **152**, 219–226 (2008).
24. Andrews, D.M., Scalzo, A.A., Yokoyama, W.M., Smyth, M.J. & Degli-Esposti, M.A. Functional interactions between dendritic cells and NK cells during viral infection. *Nat. Immunol.* **4**, 175–181 (2003).
25. Brigl, M., Bry, L., Kent, S.C., Gumperz, J.E. & Brenner, M.B. Mechanism of CD1d-restricted natural killer T cell activation during microbial infection. *Nat. Immunol.* **4**, 1230–1237 (2003).
26. Fernandez, N.C. *et al.* Dendritic cells directly trigger NK cell functions: cross-talk relevant in innate anti-tumor immune responses *in vivo*. *Nat. Med.* **5**, 405–411 (1999).
27. Vincent, M.S. *et al.* CD1-dependent dendritic cell instruction. *Nat. Immunol.* **3**, 1163–1168 (2002).
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.* **5** Suppl 1, S49–S59 (2005).
29. Savage, A.K. *et al.* The transcription factor PLZF directs the effector program of the NKT cell lineage. *Immunity* **29**, 391–403 (2008).
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.* **9**, 1055–1064 (2008).
31. Yu, S. & Cantorna, M.T. The vitamin D receptor is required for iNKT cell development. *Proc. Natl. Acad. Sci. USA* **105**, 5207–5212 (2008).
32. Gumperz, J.E., Miyake, S., Yamamura, T. & Brenner, M.B. Functionally distinct subsets of CD1d-restricted natural killer T cells revealed by CD1d tetramer staining. *J. Exp. Med.* **195**, 625–636 (2002).
33. Watarai, H. *et al.* Development and function of invariant natural killer T cells producing T_h2- and T_h17-cytokines. *PLoS Biol.* **10**, e1001255 (2012).
34. Crowe, N.Y. *et al.* Differential antitumor immunity mediated by NKT cell subsets *in vivo*. *J. Exp. Med.* **202**, 1279–1288 (2005).
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.* **208**, 1163–1177 (2011).
36. Johnston, B., Kim, C.H., Soler, D., Emoto, M. & Butcher, E.C. Differential chemokine responses and homing patterns of murine TCRαβ NKT cell subsets. *J. Immunol.* **171**, 2960–2969 (2003).
37. Townsend, M.J. *et al.* T-bet regulates the terminal maturation and homeostasis of NK and Vα14i NKT cells. *Immunity* **20**, 477–494 (2004).
38. Thomas, P.D. *et al.* PANTHER: a browsable database of gene products organized by biological function, using curated protein family and subfamily classification. *Nucleic Acids Res.* **31**, 334–341 (2003).
39. O'Brien, R.L. & Born, W.K. γδ T cell subsets: a link between TCR and function? *Semin. Immunol.* **22**, 193–198 (2010).
40. Fahrer, A.M. *et al.* Attributes of gammadelta intraepithelial lymphocytes as suggested by their transcriptional profile. *Proc. Natl. Acad. Sci. USA* **98**, 10261–10266 (2001).
41. Shires, J., Theodoridis, E. & Hayday, A.C. Biological insights into TCRγδ$^+$ and TCRαβ$^+$ intraepithelial lymphocytes provided by serial analysis of gene expression (SAGE). *Immunity* **15**, 419–434 (2001).
42. Vivier, E. & Anfossi, N. Inhibitory NK-cell receptors on T cells: witness of the past, actors of the future. *Nat. Rev. Immunol.* **4**, 190–198 (2004).
43. Gattinoni, L. *et al.* A human memory T cell subset with stem cell-like properties. *Nat. Med.* **17**, 1290–1297 (2011).

44. Matsuda, J.L. *et al.* T-bet concomitantly controls migration, survival, and effector functions during the development of Valpha14i NKT cells. *Blood* **107**, 2797–2805 (2006).
45. Gordy, L.E. *et al.* IL-15 regulates homeostasis and terminal maturation of NKT cells. *J. Immunol.* **187**, 6335–6345 (2011).
46. Kastner, P. *et al.* Bcl11b represses a mature T-cell gene expression program in immature CD4+CD8+ thymocytes. *Eur. J. Immunol.* **40**, 2143–2154 (2010).
47. Yue, X., Izcue, A. & Borggrefe, T. Essential role of mediator subunit Med1 in invariant natural killer T-cell development. *Proc. Natl. Acad. Sci. USA* **108**, 17105–17110 (2011).
48. Sullivan, B.M., Juedes, A., Szabo, S.J., von Herrath, M. & Glimcher, L.H. Antigen-driven effector CD8 T cell function regulated by T-bet. *Proc. Natl. Acad. Sci. USA* **100**, 15818–15823 (2003).
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.* **27**, 2885–2891 (1997).
50. Honma, S. *et al.* Dec1 and Dec2 are regulators of the mammalian molecular clock. *Nature* **419**, 841–844 (2002).
51. Miyazaki, K. *et al.* The role of the basic helix-loop-helix transcription factor Dec1 in the regulatory T cells. *J. Immunol.* **185**, 7330–7339 (2010).
52. Sun, H., Lu, B., Li, R.Q., Flavell, R.A. & Taneja, R. Defective T cell activation and autoimmune disorder in Stra13-deficient mice. *Nat. Immunol.* **2**, 1040–1047 (2001).
53. Weber, B.N. *et al.* A critical role for TCF-1 in T-lineage specification and differentiation. *Nature* **476**, 63–68 (2011).
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.* **176**, 1439–1446 (2006).
55. Gattinoni, L. *et al.* Wnt signaling arrests effector T cell differentiation and generates CD8+ memory stem cells. *Nat. Med.* **15**, 808–813 (2009).

ImmGen Project Consortium: Paul Monach[7], Susan A Shinton[8], Richard R Hardy[8], Radu Jianu[9], David Koller[9], Jim Collins[10], Roi Gazit[11], Brian S Garrison[11], Derrick J Rossi[11], Kavitha Narayan[4], Katelyn Sylvia[4], Joonsoo Kang[4], Anne Fletcher[12], Kutlu Elpek[12], Angelique Bellemare-Pelletier[12], Deepali Malhotra[12], Shannon Turley[12], Adam J Best[13], Jamie Knell[13], Ananda Goldrath[13], Vladimir Jojic[14], Daphne Koller[14], Tal Shay[15], Aviv Regev[15], Nadia Cohen[16], Patrick Brennan[16], Michael Brenner[16], Taras Kreslavsky[12], Natalie A Bezman[17], Joseph C Sun[17], Charlie C Kim[17], Lewis L Lanier[17], Jennifer Miller[18], Brian Brown[18], Miriam Merad[18], Emmanuel L Gautier[18,19], Claudia Jakubzick[18], Gwendalyn J Randolph[18,19], Francis Kim[20], Tata Nageswara Rao[20], Amy Wagers[20], Tracy Heng[21], Michio Painter[21], Jeffrey Ericson[21], Scott Davis[21], Ayla Ergun[21], Michael Mingueneau[21], Diane Mathis[21] & Christophe Benoist[21]

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

ONLINE METHODS

Mice. Male C57BL/6 mice 6 weeks of age were obtained from Jackson Laboratories 1 week before organ collection. 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. Antibody to CD19 (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 (α-galactosylceramide analog) were provided by the tetramer facility of the US National Institutes of Health.

Cell isolation, microarray analysis and subset nomenclature. All purification of cells of the immune system was done in strict adherence to standard operating procedure guidelines of the ImmGen Project. Additional information about the thymic, NK cell, γδ T cell and $CD8^+$ T cell populations used for comparison with *i*NKT cells is available at the ImmGen Project website. For *i*NKT cells, thymocytes, splenocytes, liver and lung mononuclear cells were isolated from five to ten mice per sample. Thymocytes were disaggregated and then blocked with anti-CD16-CD32 (2.4G2; eBioscience), then were stained with fluorophore-labeled antibodies for depletion of non-*i*NKT cell populations and were separated with anti-fluorophore magnetic beads (Miltenyi). Spleens were disaggregated, treated with ammonium chloride–potassium bicarbonate lysis buffer (Lonza) to remove red blood cells, and then depleted of non-*i*NKT cells as described above. Lungs and livers were collected after perfusion with cold PBS and were mechanically homogenized. Lungs were digested for 15 min at 37 °C in 7 U/ml of Liberase III enzyme (Roche) in DMEM, then were filtered and washed. Livers were homogenized and liver mononuclear cells were isolated by Ficoll density gradient centrifugation. *i*NKT cell-enriched thymocytes, splenocytes, liver mononuclear cells and lung cells were stained for cell surface markers (information about the staining procedure, including gating strategy, is available at the ImmGen Project website) and were double sorted directly into TRIzol (Invitrogen) at a purity of >99%. By staining of thymocytes from CD1d-deficient mice that lack *i*NKT cells, a false-positive rate of 2.7% ± 0.9% was estimated for $CD44^-$ $NK1.1^-$ (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 with a FACSAria, with the exception of the $CD4^-$ *i*NKT subset from the lungs, for which only a single replicate passed quality control (disussed below). RNA extraction, microarray hybridization (Affymetrix MoGene 1.0 ST array) and data were processing were done at the ImmGen Project processing center. Additional details are available in **Supplementary Table 16** (subset nomenclature key), Data Generation and Quality Control pipeline documentation, or the ImmGen Project website. Human and mouse data sets not from the ImmGen Project were from data sets from the GEO (Gene Expression Omnibus) database repository of the National Center for Biotechnology Information[43,56]. Human homologs of mouse genes were identified with the HomoloGene automated system for detecting homologs (National Center for Biotechnology Information).

Data analysis and visualization. Data from the March 2011 ImmGen Project release (802 arrays with 22,268 probe sets) were used. Probe sets associated with the same gene symbol were consolidated by selection of the probe set with the highest mean expression overall. For heat maps, data were $\log_2$-transformed, and a relative color scale with row centering (subtraction of the mean) and normalization was used. Heat maps were produced with the HeatmapViewer module of GenePattern software (Broad Institute). Where indicated, Pearson correlation with pairwise complete linkage was applied to rows for clustering analysis. Volcano plots were produced by with the Multiplot module of GenePattern.

For *K*-means clustering analysis, genes were prefiltered for a mean expression value of ≥120 (cutoff above which genes have a 95% chance of expression) and for a change in expression of over twofold between any two subsets analyzed. The ExpressCluster application of GenePattern (http://www.broadinstitute.org/cancer/software/genepattern/) was used for clustering, with $K = 10$.

For Euclidean-distance matrices, the 15% most variable genes were identified with the PopulationDistances application of GenePattern, which filters probes on the basis of a variation of ANOVA with the geometric s.d. of populations to 'weight' genes that vary in multiple populations. The selected genes were $\log_2$-transformed, filtered for probes with a mean expression value of ≥120, and centered on the mean before visualization. Genes encoding 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**) and were manually removed for the second analysis (**Supplementary Fig. 5**). *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, Euclidean-distance analysis.

For ANOVA, data were $\log_2$-transformed, low-variability genes with a s.d. of ≤0.5 across all ImmGen Project samples were removed and only genes with expression values of ≥120 in two or more arrays were considered, which left 5,900 genes. ANOVA was done with MATLAB software function 'anova1' (MathWorks) for comparison of *i*NKT cell, NK cell and T cell populations, and the Bonferroni correction was applied to the resulting list of *P* values, and 1,192 (of 5,900) genes passed the ANOVA test ($P < 0.05$). 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. On the basis of the results of this analysis, the 1,192 genes were then separated into 14 gene categories. By comparison of the average expression of each gene in NK cells, *i*NKT cells and T cells, these categories were then sorted into six groups of genes: genes with the most similar expression in NK cells and *i*NKT cells that were upregulated (A_1) or downregulated (A_2) relative to the expression in T cells; genes with the most similar expression in T cells and *i*NKT cells that were upregulated (B_1) or downregulated (B_2) relative to the expression in NK cells; and genes with upregulated (C_1) or downregulated (C_2) expression in *i*NKT cells relative to the expression in T cells or NK cells. Functional gene-set enrichment analysis was done with software from the DAVID bioinformatics database (Database for Annotation, Visualization and Integrated Discovery; Version 6.7; National Institute for Allergy and Infectious Diseases of the US National Institutes of Health)[57]. Biological processes of the PANTHER database of gene products organized by biological function (Protein Analysis Through Evolutionary Relationships)[38] are presented in **Tables 1** and **2**. *P* values calculated with DAVID represent a modified Fisher's exact test. Biological processes with *P* values less than 0.01 are presented.

For estimation of the significance of enrichment for the distributions associated with the heat maps in **Figures 5–7**, Kolmogorov-Smirnov *P* values were calculated with JMP software (SAS Institute) with comparison of the selected gene set with all genes meeting the criteria for expression (>120) in the samples tested.

56. Marshall, H.D. *et al.* Differential expression of Ly6C and T-bet distinguish effector and memory Th1 $CD4^+$ cell properties during viral infection. *Immunity* **35**, 633–646 (2011).
57. Huang da, W., Sherman, B.T. & Lempicki, R.A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44–57 (2009).