

ARTICLE

DOI: 10.1038/s41467-018-05026-6

OPEN

TCR signal strength controls thymic differentiation of iNKT cell subsets

Kathryn D. Tuttle^{1,2}, S. Harsha Krovi^{1,2}, Jingjing Zhang^{1,2}, Romain Bedel^{1,2,11}, Laura Harmacek^{3,4}, Lisa K. Peterson^{4,12,13}, Leonard L. Dragone^{4,14}, Adam Lefferts^{1,2}, Catherine Halluszczak^{1,2}, Kent Riemondy⁵, Jay R. Hesselberth^{5,6}, Anjana Rao^{7,8,9}, Brian P. O'Connor^{3,4}, Philippa Marrack^{1,2,10}, James Scott-Browne^{7,8} & Laurent Gopin^{1,2}

During development in the thymus, invariant natural killer T (iNKT) cells commit to one of three major functionally different subsets, iNKT1, iNKT2, and iNKT17. Here, we show that T cell antigen receptor (TCR) signal strength governs the development of iNKT cell subsets, with strong signaling promoting iNKT2 and iNKT17 development. Altering TCR diversity or signaling diminishes iNKT2 and iNKT17 cell subset development in a cell-intrinsic manner. Decreased TCR signaling affects the persistence of *Egr2* expression and the upregulation of PLZF. By genome-wide comparison of chromatin accessibility, we identify a subset of iNKT2-specific regulatory elements containing NFAT and Egr binding motifs that is less accessible in iNKT2 cells that develop from reduced TCR signaling. These data suggest that variable TCR signaling modulates regulatory element activity at NFAT and Egr binding sites exerting a determinative influence on the dynamics of gene enhancer accessibility and the developmental fate of iNKT cells.

¹Department of Immunology and Microbiology, University of Colorado Anschutz Medical Campus, 12800 E. 19th Ave, Aurora, CO 80045, USA.

²Department of Biomedical Research, National Jewish Health, 1400 Jackson Street, Denver, CO 80206, USA. ³Center for Genes, Environment, and Health, Department of Biomedical Research, National Jewish Health, 1400 Jackson Street, Denver, CO 80206, USA. ⁴Department of Pediatrics, National Jewish Health, 1400 Jackson Street, Denver 80206 CO, USA. ⁵RNA Bioscience Initiative, University of Colorado School of Medicine, 12800 E. 19th Ave, Aurora 80045 CO, USA. ⁶Department of Biochemistry & Molecular Genetics, University of Colorado School of Medicine, 12800 E. 19th Ave, Aurora, CO 80045, USA. ⁷La Jolla Institute, 9420 Athena Cir, La Jolla 92037 CA, USA. ⁸Sanford Consortium for Regenerative Medicine, 2880 Torrey Pines Scenic Dr, La Jolla, CA 92037, USA. ⁹University of California San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA. ¹⁰Department of Medicine, University of Colorado Anschutz Medical Campus, 12800 E. 19th Ave, Aurora, CO 80045, USA. ¹¹Present address: Department of Oncology, University of Lausanne, Chemin des Boveresses 155, Epalinges 1066, Switzerland. ¹²Present address: ARUP Laboratories, Institute of Clinical and Experimental Pathology, 500 Chipeta Way, Salt Lake City 84108, UT, Switzerland. ¹³Present address: Department of Pathology, University of Utah, 30N 1900E, Salt Lake City 84132, UT, USA. ¹⁴Present address: Merck Research Laboratories, San Francisco, CA, USA. These authors contributed equally: Kathryn D. Tuttle, S. Harsha Krovi. Correspondence and requests for materials should be addressed to L.G. (email: Laurent.Gopin@ucdenver.edu)

• NKT cells recognize lipid antigens presented by CD1d, a non-polymorphic major histocompatibility complex (MHC) class I-like antigen-presenting molecule¹. These cells use a semi-invariant TCR made up mostly of a single invariant TCR α chain (V α 14-J α 18 in mice, V α 24-J α 18 in humans) with certain TCR β chains (V β 8.2, V β 7, or V β 2 in mice, V β 11 in humans) to engage CD1d. In the thymus, iNKT cells develop into three major terminally differentiated and functionally distinct iNKT cell subsets^{2,3}. iNKT1 cells express the transcription factor T-bet and secrete predominantly IFN γ ; iNKT2 cells express high levels of the GATA3 and promyelocytic leukaemia zinc finger (PLZF) transcription factors and secrete IL-4 and IL-13; iNKT17 have intermediate levels of PLZF, are positive for RAR-related orphan receptor gamma (Roryt) expression, and secrete IL-17. Importantly, the relative distribution of the three thymic iNKT subsets varies in different mouse strains and affects the phenotype and activation status of surrounding cells.

Unlike most T cells that leave the thymus to populate the peripheral immune organs, some mature iNKT cells are retained in the thymus and become long-term thymic residents⁴, with potentially important functions. In particular, mature thymic iNKT2 cells produce IL-4 at steady state and thus affect the homeostasis of thymic cell populations². Indeed, IL-4 conditions CD8⁺ T cells to become “memory-like” and to express the transcription factor Eomesodermin⁵. These CD8 “memory-like” T cells have important roles in early defenses, particularly in situations of chronic viral infection^{6,7}. Thymic steady-state IL-4 also drives the acquisition of an activated/memory-like phenotype by Foxp3⁺ regulatory T cells⁸, the production of chemokines by thymic dendritic cells², the thymic exit of mature “conventional” T cells⁹ and also perhaps the commitment of early thymic progenitors to the T cell lineage¹⁰. Additionally, RANKL-expressing CD44⁻ thymic iNKT cells (which are preferentially enriched for iNKT2 and iNKT17 cells) regulate the differentiation of Aire⁺ MHC class II⁺ medullary thymic epithelial cells¹¹ that are involved in clonal deletion of self-reactive T cells¹² and Treg maturation¹³. Altogether, these results suggest that thymic iNKT cells, and particularly the relative subset representation, have fundamental roles in the composition of other thymic cell populations, both by modulating homeostasis and maturation status of these cells, and potentially in shaping the overall size and repertoire diversity of mature “conventional” T cells.

Development of iNKT cells diverges from that of conventional T cells primarily at the double-positive CD4⁺ CD8⁺ (DP) stage and requires TCR recognition of CD1d on DP cells, involving homotypic interactions across a DP–DP synapse where second signals are initiated by the engagement of homophilic receptors of the signaling lymphocytic-activation molecule (SLAM) family, Slamf1 (SLAM) and Slamf6 (Ly108). This signaling recruits the adaptor SLAM-associated protein (SAP) and the Src kinase Fyn, both of which are essential for the development of the iNKT cell lineage³. The TCR signals received by iNKT cell precursors during selection are associated with high expression of the Ras-¹⁴ and Ca²⁺-dependent transcription factors Egr1 and, especially, Egr2^{15,16}. Interestingly, high expression levels of Egr2 in pre-selection DP thymocytes are potentiated by co-stimulation through Ly108^{17,18}. Egr2 directly regulates the expression of several genes involved in the development of iNKT cells, including PLZF and CD122, one of the chains of the IL-15 receptor¹⁵. Egr2 is recruited to the promoter of *Zbtb16* (which encodes PLZF) after TCR engagement and co-stimulation with Ly108^{15,17}. PLZF directs the acquisition of effector properties, such as the upregulation of CD44 and production of effector cytokines^{19,20}. Expression of PLZF directs the acquisition of effector properties by binding and regulating T helper-specific transcription factor genes that in turn control T-helper-specific

programs^{19–21}. Several transcription factors and signaling molecules affect the lineage diversification of iNKT cell subsets. On the whole, however, the mechanisms that control these iNKT cell fate decisions during development remain poorly understood.

Here, we show that TCR signal strength governs the development of iNKT cell subsets in the thymus, with high signal strength being necessary for iNKT2 and iNKT17 development. The avidity of the iNKT TCR–CD1d interaction correlates with iNKT cell subset assignment and the expression of markers reflecting strength of signaling during selection, suggesting that differential TCR signal strength likely has a role in iNKT cell subset differentiation. Limiting the iNKT TCR repertoire diversity through the use of TCR β transgenic mice and/or TCR signaling using mice with Zap70 mutations markedly impacts iNKT2 and iNKT17 cell subset development in a cell-intrinsic manner. Decreased TCR signaling affects the longevity of Egr2 expression and the upregulation of PLZF both in vitro and in vivo. Genome-wide chromatin accessibility analysis reveals subset-specific activity of regulatory elements associated with unique transcription factor binding site signatures. NFAT and Egr binding motifs are found preferentially enriched in chromatin regulatory regions specifically accessible in iNKT2 cells that are not accessible in iNKT2 cells that had developed with reduced TCR signaling. Altogether, these data suggest a model of iNKT cell subset development wherein variable TCR signaling induces changes in chromatin accessibility at NFAT and Egr binding sites that exert a determinative influence on the dynamics of gene enhancer accessibility that in turn determines the developmental fate of iNKT cells.

Results

Characteristics of thymic iNKT cell subsets. We analyzed thymic iNKT cells from BALB/c mice in which the human CD2 gene was knocked-in into the IL-4 locus (KN2), thereby allowing for the identification by flow cytometry of IL-4-secreting iNKT2 cells directly ex vivo. iNKT cells, identified using PBS57–CD1d tetramer and TCR β co-staining, were electronically placed on a grid consisting of 30 gates, allowing for the analysis of iNKT cells expressing different TCR levels and/or different binding to the PBS57–CD1d tetramer for a given level of TCR (Fig. 1a). Thus, each gate examined cells with different avidities for the antigen–CD1d complex. The proportion of each iNKT subset in each of the 30 different gates was calculated (Supplementary Fig. 1) and displayed as a heatmap. Although not all gates contain the same number of cells, displaying the proportion of each iNKT cell subsets in the 30 gates revealed a direct correlation between the overall avidity of the TCR–PBS57–CD1d interaction and the three major iNKT cell subset phenotypes (Fig. 1b). iNKT1 cells express the lowest amount of TCR and are stained by the tetramer with the lowest intensity. iNKT17 cells express intermediate levels of TCR. However, for a given level of TCR expression, iNKT17 cells bind the tetramer more strongly than iNKT1 cells. Finally, iNKT2 cells express the highest TCR levels and stained more brightly with the PBS57–CD1d tetramer. This correlation is not observed with another marker of mature iNKT cells (CD44). These results demonstrate that the different thymic iNKT cell subsets have disparate but overlapping avidities for the PBS57–CD1d complex.

It is conceivable that these observed differences in avidities for the CD1d-foreign antigen PBS57 could also reflect differential avidities for the CD1d-self antigen(s) recognized during positive selection of the different iNKT cell subsets. Consequently, the strength of TCR signaling received by the different iNKT cell subsets during positive selection might be different as well. Surface expression of CD5 and CD6 molecules directly parallels

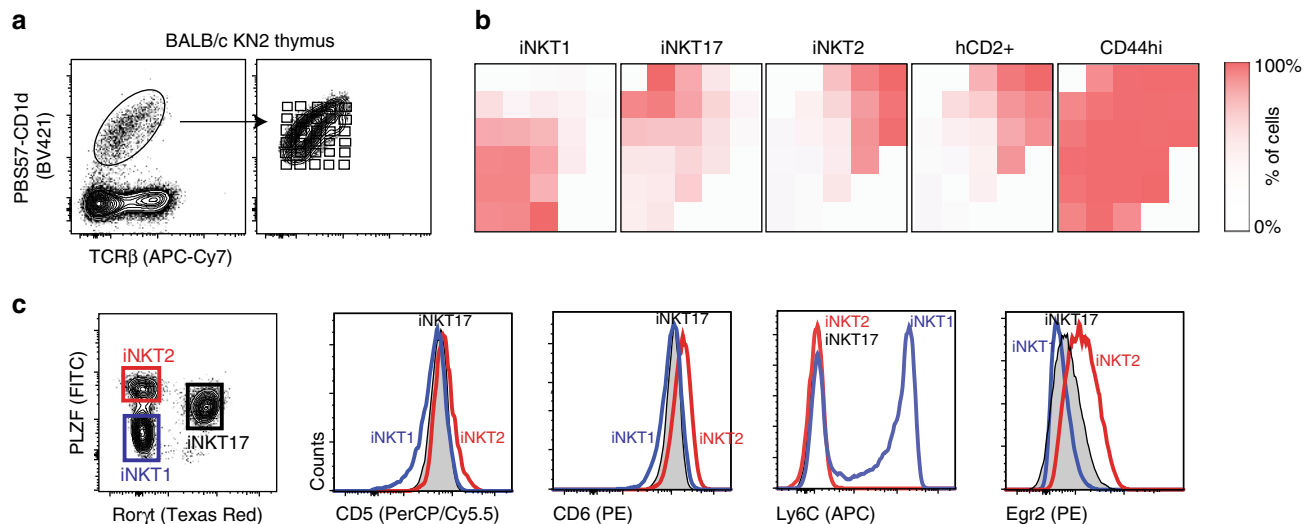


Fig. 1 Thymic iNKT subsets have different avidities for the PBS57–CD1d tetramer that correlate with expression of surrogate markers reflective of strength of signaling during selection. **a** Cells from the thymus of BALB/c IL-4 reporter KN2 mice were stained with anti-TCR β mAbs and PBS57–CD1d tetramers. iNKT cells were then electronically placed on a grid consisting of 30 gates, allowing for the analysis of iNKT cells expressing different TCR levels and/or different binding to the PBS57–CD1d tetramer for a given level of TCR. **b** The proportion of iNKT1 (PLZF^{lo}, Roryt⁻, Tbet⁺), iNKT17 (PLZF^{int}, Roryt⁺, Tbet⁻), iNKT2 (PLZF^{hi}, Roryt⁻, Tbet⁻, or hCD2⁺) as well as the proportion of CD44^{hi} cells in each gate was recorded and displayed as a heatmap. **c** Representative histograms for the expression of CD5, CD6, Ly6C, and Egr2 in each iNKT cell subset (as defined by the gating strategy shown on the left) is shown. Data are representative of $n > 3$ mice for each staining

the signaling intensity of the positively selecting TCR–MHC–ligand interaction^{22–26}, while expression of Ly6C by CD4 T cells is inversely correlated with self-reactivity of their TCRs²⁷. As seen in Fig. 1c, iNKT2 cells express the highest level of CD5 and CD6 proteins while they are negative for Ly6C. iNKT17 cells have lower levels of CD5 and CD6 expression than iNKT2 cells and are also Ly6C negative. Finally, iNKT1 cells express the lowest levels of CD5 and CD6 proteins of the three subsets, while >60% of the cells are also Ly6C positive. Thus, on the basis of expression of these 3 markers, it appears that the three iNKT cell subsets have received different strengths of signaling during selection in the thymus, with the following hierarchy iNKT2 > iNKT17 > iNKT1. We also examined Egr2 expression in the three iNKT cell subsets (Fig. 1c). The results show that iNKT2 cells express the highest level of Egr2 followed by iNKT17 and iNKT1 cells. These results are in agreement with the expression of GFP driven from the immediate early gene *Nr4a1* (*Nur77*) locus, which directly correlates with the strength of TCR stimulus²⁸. Altogether, these results demonstrate that the different iNKT cell subsets receive signals of different strength during their development in the thymus and that the avidity of the TCR repertoires expressed by each subset positively correlates with this signaling strength.

Development of iNKT subsets in mice with a fixed TCR β chain.

We examined whether the development of iNKT cell subsets would be disturbed if the specificity of the repertoire was fixed. We used transgenic mice that express the TCR β chain derived from the OVA-IA^d-specific clone DO-11.10 (DO). The composition of iNKT cell subsets in the thymus of 8-week-old B6 and BALB/c mice is different with BALB/c mice having large populations of iNKT2 and iNKT17 cells compared to B6 mice (Fig. 2)². Interestingly, the B6 \times BALB/c F1 strain recapitulates the BALB/c phenotype (Fig. 2), allowing us to analyze the effect of TCR β DO expression on iNKT cell subset development in B6 mice and F1 (B6 \times BALB/c) mice. Enforced expression of a single TCR β chain strongly affects the development of iNKT cell subsets on both

genetic backgrounds (Fig. 2). The proportions of iNKT2 and iNKT17 cells in DO11.10 TCR β transgenic mice are largely reduced compared to those of control mice with polyclonal repertoires, demonstrating that the specificity of the iNKT TCR has a key role in the development of iNKT cell subsets.

V β repertoire and Egr2 levels in thymic iNKT cell subsets.

In addition to its surface expression level, the nature of the TCR itself can affect the duration and the strength of TCR signaling²⁹. Therefore, we assessed the percentage of V β usage for the three iNKT cell subsets expressing different avidities for the PBS57–CD1d tetramer. BALB/c thymic iNKT cells were electronically gated based on three discrete TCR levels (Fig. 3a). Although the correlation between avidity and iNKT cell subsets held generally true (Fig. 1a), each iNKT subset could nevertheless be found, albeit in very low frequencies in some instances, for each level of TCR expression (Fig. 3a). We examined the expression of V β 2, V β 7, V β 8.1/8.2, and V β 8.3 by the three iNKT cell subsets found for each level of TCR (Fig. 3b). In agreement with previous results², we found that the proportion of V β usage is different between iNKT cells subsets. Unexpectedly, however, the proportion of V β usage within a given subset also varies between cells having different avidities for the PBS57–CD1d complex. For example, 52% of iNKT17 cells expressing low TCR levels use the gene segment V β 8.3, while only 5.1% of iNKT17 cells that are TCR^{hi} are V β 8.3⁺ (Fig. 3b). Likewise, 19% of TCR^{lo} iNKT2 cells use V β 8.1/8.2 but this proportion increased to 55% for iNKT2 cells expressing high levels of TCRs (Fig. 3b). These results demonstrate that the repertoire of each iNKT cell subset is not only unique to a given subset but also varies as a function of the level of TCR surface expression. We also analyzed the levels of Egr2 expression in each iNKT subset expressing different TCR levels. Each iNKT cell subset expresses a limited range of Egr2 protein (Fig. 3c, d) that appears delimited by the levels of TCR surface expression for a given subset. For example, the few iNKT1 cells that are TCR^{hi} express Egr2 levels that remain lower than TCR^{lo} iNKT2 cells (Fig. 3d). Thus, it appears that each iNKT

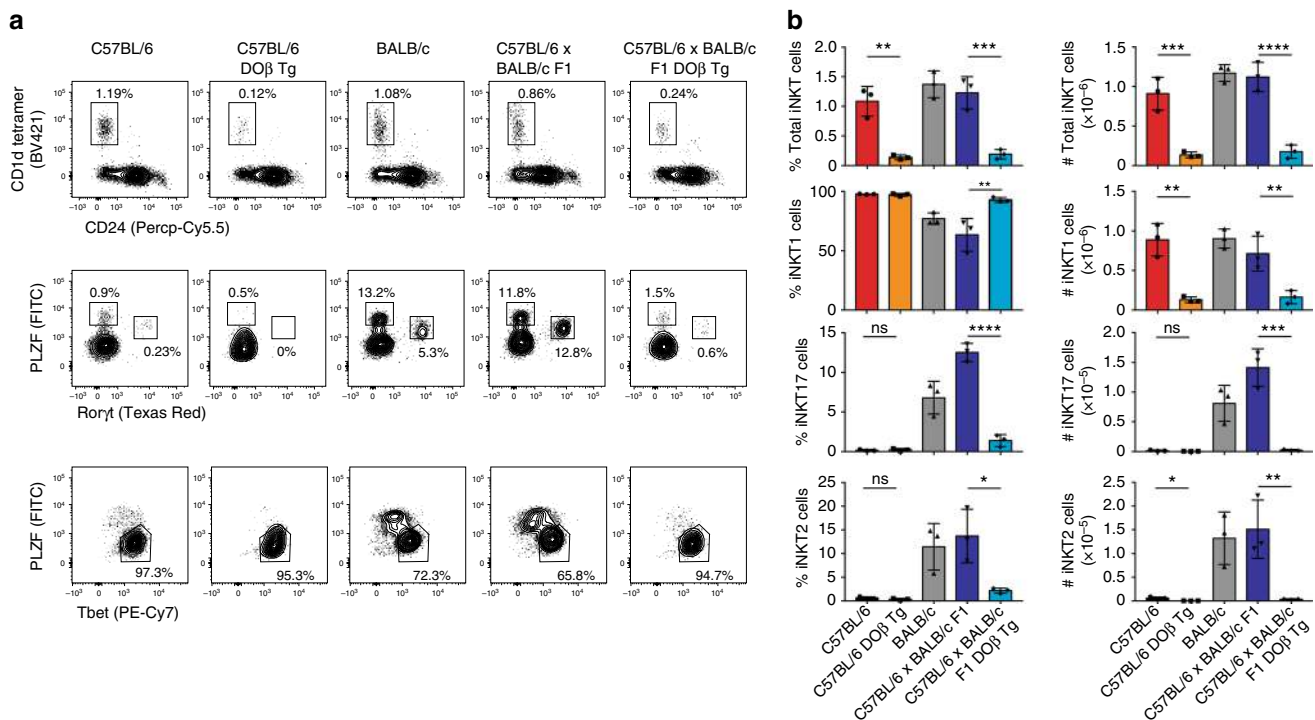


Fig. 2 Development of iNKT cell subsets in DO11.10 TCR β transgenic mice on the C57BL/6 or C57BL/6 \times BALB/c F1 background. **a** The thymi of 8-week-old C57BL/6, C57BL/6 DO11.10 TCR β transgenic, BALB/c, C57BL/6 \times BALB/c F1, and BALB/c \times C57BL/6 DO11.10 TCR β transgenic mice were stained with anti-CD24 mAbs, PBS57-CD1d tetramers and mAbs specific for the transcription factors PLZF, Roryt, and Tbet. Representative flow cytometry plots for the frequency of total mature iNKT cells and iNKT subsets in each genotype is shown. **b** Summary of the data shown in **a** with three mice per group. Data are mean \pm SD. Significance was assessed using multiple comparison one-way analysis of variance (ANOVA) * p < 0.05; ** p < 0.01; *** p < 0.001; **** p < 0.0001; ns, not significant. For clarity purposes, the only statistics shown are between mice with or without expression of the DO11.10 TCR transgene for a given genetic background

cell subset expresses a unique TCR β repertoire modulated by the overall level of TCR expression on the cells but that the combination results in discrete signal strength.

Thymic iNKT cell subsets in Zap70-mutant mice. To further test the idea that strength of TCR signaling affects iNKT cell fate choice, we examined iNKT cell development in mouse models with hypo-responsiveness to TCR stimulation, the YYAA and SKG mice^{30,31}. Both mutations diminish TCR signal strength and thus provide opportunities to analyze the effects of signal strength on iNKT cell differentiation on both B6 and BALB/c genetic backgrounds. Both strains show skewed development of iNKT subsets (Fig. 4). In the thymus of YYAA mice, iNKT2 cells are largely decreased proportionally and numerically compared to the B6 controls (Fig. 4a, b), while in SKG mice, which have a more profound impairment of TCR signaling than the YYAA mice³⁰, most of the iNKT cells are iNKT1 cells in striking contrast to those in age-matched control BALB/c mice (Fig. 4a, c). Because differences in subset ratios are stronger between BALB/c and SKG mice, subsequent experiments are only carried out with these two strains.

Cell-intrinsic Zap70 affects iNKT subset development. To determine whether the SKG mutation effects on iNKT cell subset development are cell intrinsic or extrinsic, we generated competitive bone marrow chimeras. BALB/c-derived bone marrow cells produce all three iNKT cell subsets with large proportions of iNKT2 (40–50%) and iNKT17 (5–6%) cells (Fig. 4d). By contrast, SKG-derived iNKT cells are composed mostly of iNKT1 cells, with poor representation of iNKT2 (4%) and iNKT17 (0.3%)

cells. These results demonstrate that iNKT cell precursors need to receive adequate TCR signaling mediated through Zap70 in order to develop efficiently into the iNKT2 and iNKT17 lineages and that cell extrinsic signals in the thymus cannot compensate for this deficiency.

Zap70 activity-dependent gene expression in stage 0 cells. To determine whether changes in strength of TCR signaling might have affected the developmental program of iNKT cells, we used RNA-Seq to compare gene expression in the earliest iNKT cell progenitors, just after positive selection (stage 0 cells), between BALB/c and SKG mice. iNKT cells were enriched using PBS57-CD1d tetramer and magnetic beads (Fig. 5a). Stage 0 iNKT cells were defined as CD44^{lo}, CD24⁺, CD4^{low/-}, CD8⁻, CD69⁺, and sorted using these markers. This sorting strategy eliminates DP thymocytes that invariably contaminate the CD24⁺ fraction and that do not represent genuine stage 0 iNKT cells that are actively signaling (Fig. 5a and Supplementary Fig 2). Confirmation that these cells are genuine iNKT precursors is provided by the analysis of the sequencing reads that are mapped to the TCRA locus (Supplementary Fig 3). We detected statistically significant differential expression (FDR < 0.1) of 238 genes between the two genotypes (Fig. 5b), with a significant enrichment for genes implicated in cell activation and immune cell development (Supplementary Fig 4). Amongst the differentially regulated genes, several are known regulators of iNKT cell development, including *Egr1* and *Egr2*, *Myc*, *Tcf7*, *Klf2*, *Runx3*, and *RORA*. Interestingly, several factors involved in chromatin remodeling and transcriptional regulation, which are induced during positive selection of T lymphocytes (*Patz1* and *Satb1*), are also differentially regulated. Altogether, these results demonstrate

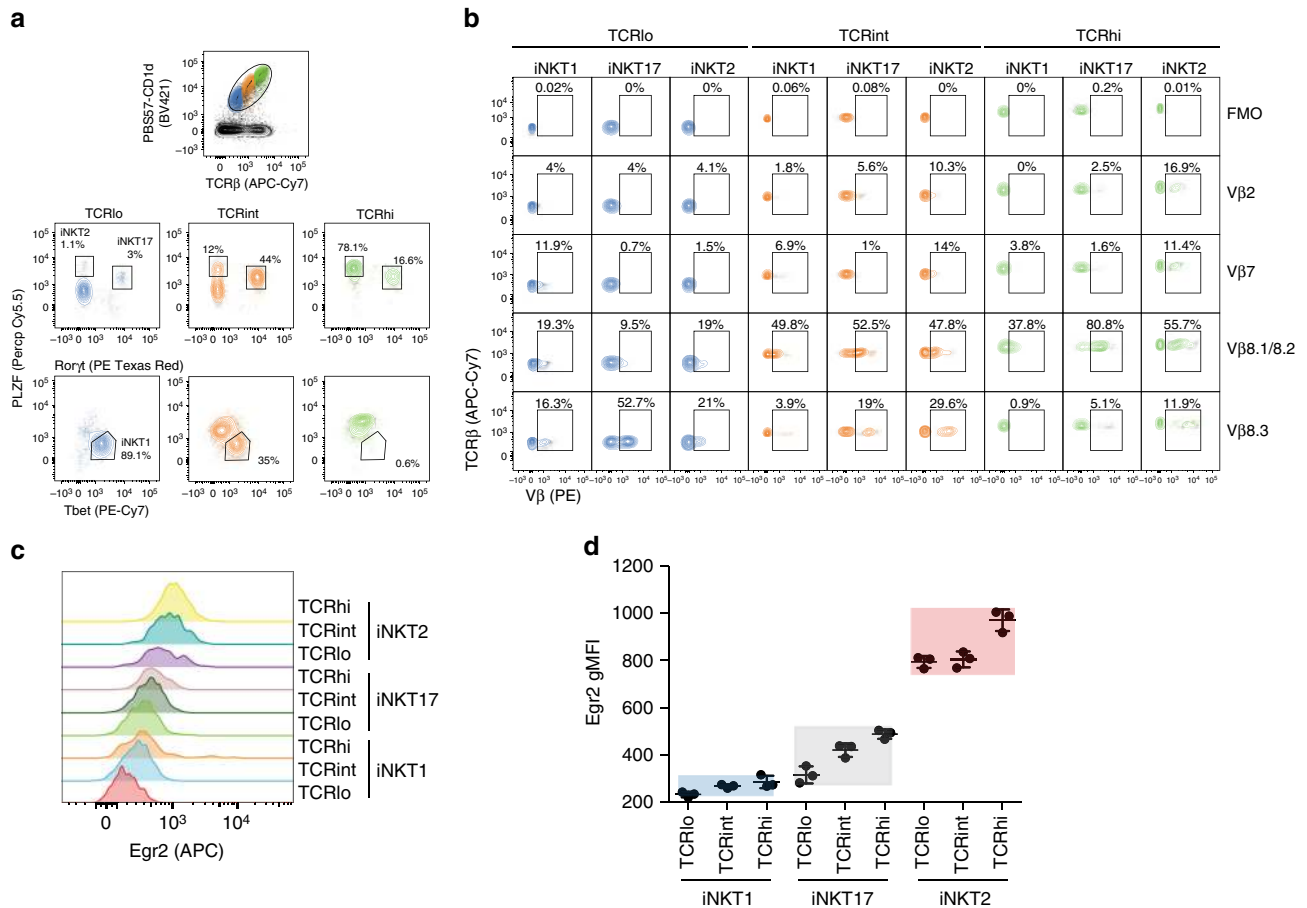


Fig. 3 Thymic iNKT subsets with different TCR levels express different proportions of Vβ usage but discrete Egr2 levels. Cells from the thymus of 8-week-old BALB/c mice were stained with anti-TCRβ mAbs, PBS57-CD1d tetramers, mAbs specific for the transcription factors PLZF, Rorγt, and Tbet and the indicated anti-Vβ mAbs. **a** iNKT cells were subdivided based on Vβ expression of different TCR levels. The proportion of iNKT cell subsets for each level of TCR expression (TCR^{low} (blue), TCR^{int} (orange) and TCR^{high} (green)) is shown. **b** The percentage of Vβ usage found in each iNKT subset expressing three distinct TCR levels is shown with the fluorescence minus one control (FMO). Data are representative of three independent mice analyzed. **c** Example of Egr2 expression in each iNKT cell subsets expressing different levels of TCR. **d** The gMFI of Egr2 expression was recorded from three independent staining of iNKT cell subsets expressing different TCR levels from one individual mouse as gated in **a**. Three independent mice were analyzed in this way with similar results

that TCR hypo-responsiveness affects iNKT cell development even at the earliest stage of commitment.

Reduced Egr2 expression in iNKT cells from SKG mice. In agreement with the RNA-seq data, the levels of Egr2 in stage 0 iNKT cells of SKG mice (gMFI 4866 ± 421) are slightly reduced (10%) compared to those in BALB/c animals (gMFI 5432 ± 293), although the results do not reach statistical significance (Fig. 5d). According to the conventional “staging” system of iNKT cell development³, immediately following stage 0, precursor cells (iNKTp) are CD44^{lo} CD24⁻ CD69⁻ and PLZF^{hi}. The cells subsequently reach stage 2 (CD44^{hi}), which contains functionally mature iNKT cell subsets³. We observed that the differences in the levels of Egr2 expression between BALB/c and SKG mice are more pronounced as the cells progress in their maturation (Fig. 5c, d). iNKTp cells from SKG mice express 1/3 less Egr2 than their counterpart cells in wild-type mice (Fig. 5d). While iNKT2 cells express the highest levels of Egr2 of the three iNKT subsets in BALB/c mice (Figs. 1 and 5d), its level of expression in the few iNKT2 cells of SKG mice (gMFI 469 ± 35) is comparable to the levels found in BALB/c iNKT1 cells (gMFI 459 ± 15). Expression of PLZF and Rorγt in iNKT2 and iNKT17 cells respectively, are also lower in SKG mice (Fig. 5d). Altogether,

these results demonstrate that during development of iNKT cells in SKG mice, Egr2 is induced at near normal levels in stage 0 iNKT cells but its expression is not sustained effectively as the cells mature.

Modulation of Egr2 and PLZF induction in vitro. While induction of Egr2 expression in developing T cells is known to occur downstream of TCR signaling, whether PLZF is induced in an Egr2-dependent manner in response to TCR signaling has been controversial^{15,32}. Pre-selection DP thymocytes from the thymus of BALB/c and SKG mice were stimulated with anti-CD3, anti-Ly108 or a combination of both, and followed over time for the expression of Egr2 and PLZF. As seen in Fig. 6, TCR stimulation alone of BALB/c DP thymocytes triggers a few cells to produce Egr2 but its level of expression, which peaks around 40 h post stimulation, remains low. In these conditions, we do not detect PLZF expression at any time point post stimulation. As reported previously^{17,18}, co-stimulation with the SLAM family member Ly108 increases the proportion of cells expressing Egr2 as well as the level of expression per cell. Interestingly, cells expressing the highest levels of Egr2 begin to express the PLZF protein around 32 h post stimulation with the proportion of PLZF⁺ cells increasing as a function of time. These results

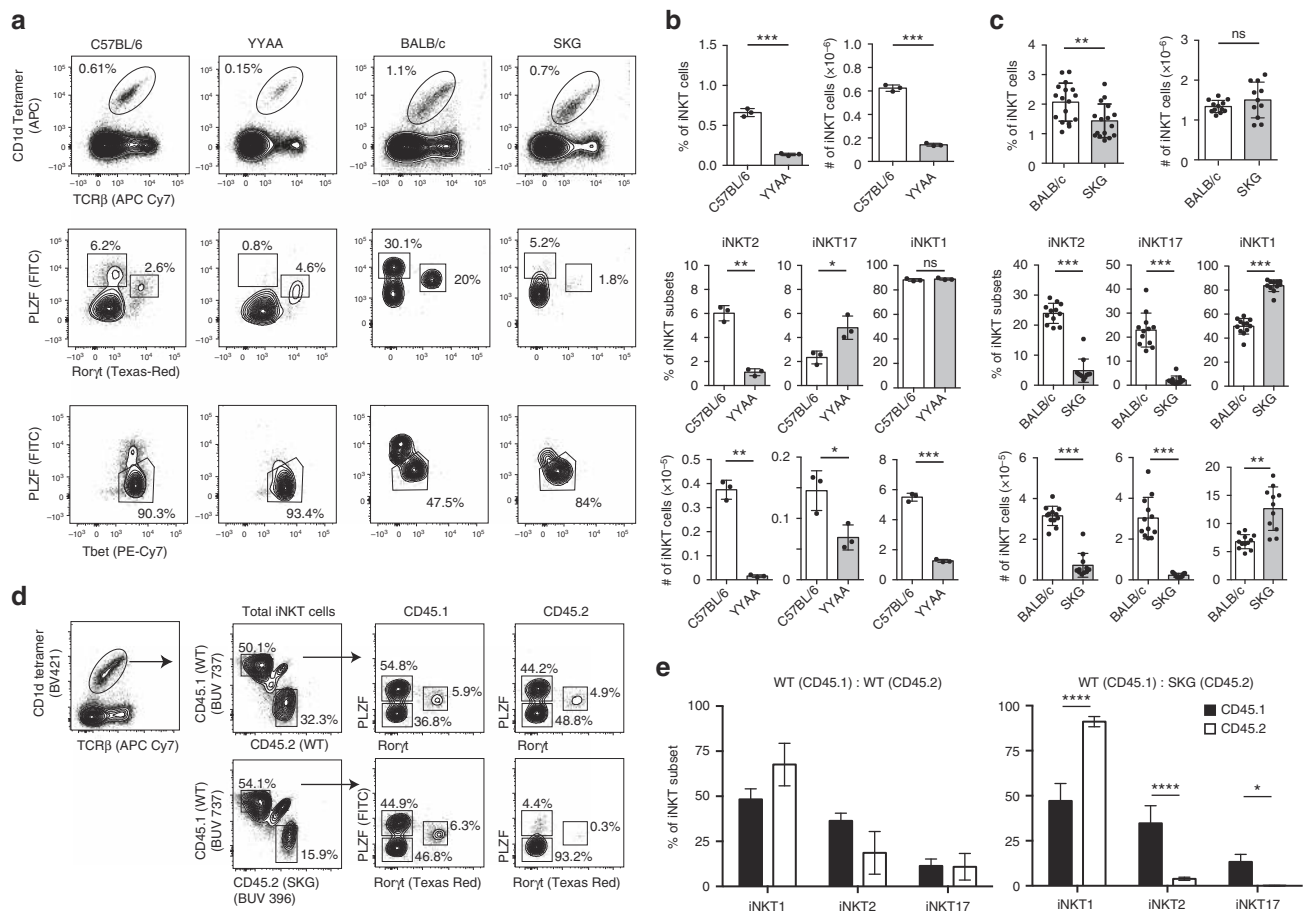


Fig. 4 Cell-intrinsic deficiency in iNKT2 and iNKT17 subsets development in mice with Zap70 hypomorphic mutations. **a** Representative flow cytometry plots showing the frequency of total iNKT cells and iNKT cell subsets in the thymi of 8-week-old C57BL/6, C57BL/6 YYAA, BALB/c, and BALB/c SKG mice. Each genotype was analyzed independently and compared to wild-type control mice of the appropriate background. **b, c** Summary of the data shown in **a** with percentages and cell numbers of total iNKT cells and iNKT cell subsets in the various genotypes analyzed. Data are mean \pm SD. Significance was assessed using the unpaired *t*-test. **d** Analysis of thymic iNKT cell subsets in competitive bone marrow chimera. Congenically marked bone marrow cells from BALB/c wild-type mice (CD45.1) were mixed in a 1:1 ratio with (CD45.2) BALB/c or SKG bone marrow cells and injected into lethally irradiated F1 (CD45.2 \times CD45.1) BALB/c mice. The thymi of the chimera were analyzed for iNKT cell subset composition 8–10 weeks post reconstitution and representative flow cytometry plots are shown. After gating on total iNKT cells, the proportion of iNKT cell subsets derived from either CD45.2⁺ or CD45.1⁺ bone marrow cells is shown. **e** Summary of the data shown in **d** with ≥ 5 mice per group. Data are represented as mean \pm SD. Significance was assessed using one-way analysis of variance (ANOVA) **p* < 0.05; ***p* < 0.01; ****p* < 0.001; *****p* < 0.0001; ns, not significant

establish that the transcription factor PLZF can be induced in DP thymocytes following TCR/Ly108 stimulation and that only the cells that sustain the highest levels of Egr2 can also express PLZF. Stimulation of SKG DP cells in the same conditions induces a slightly lower percentage of Egr2⁺ cells than the BALB/c cells (Fig. 6b, c) but the levels of Egr2 induced never reach the levels found in the BALB/c cells at any time point post stimulation, and few PLZF⁺ cells are detected in these conditions. These results recapitulate *ex vivo* observations (Fig. 5) and further support the hypothesis that modulation of TCR signal strength can affect iNKT cell subset differentiation, in part by modulating expression of Egr2 and PLZF.

Zap70 activity affects chromatin landscape of iNKT subsets. To assess how TCR signal strength might influence lineage differentiation via regulatory elements, we compared chromatin accessibility between BALB/c and SKG iNKT1 and iNKT2 cells with ATAC sequencing (ATAC-seq). iNKT1 and iNKT2 cell subsets show a unique chromatin landscape, with regions of accessibility specific to each subset. In agreement with their distinct functionalities, the accessibility of subset signature genes

(*Ifny*, *Tbx21*, *Cxcr3*, *Ccr5*, *Il2rb*, *Fasl*, *Bcl2*, *Il2*, *Zbtb16*, *Slamf6*, *Il17rb*, *Il6ra*, *Ccr7*) is different between subsets (Supplementary Figs. 5 and 6). For example, the chromatin landscapes of the *Ifny* and *Tbx21* genes, encompassing not only their promoters but also intragenic and intergenic regions extending as far as several kilobases away, are clearly accessible in iNKT1 but not in iNKT2 cells. By contrast, the chromatin landscape surrounding the *Zbtb16* (encoding PLZF), *Slamf6* (encoding Ly108), *Il17rb*, *Il6ra*, and *Ccr7* genes is clearly more accessible in iNKT2 than iNKT1 cells, in agreement with the differential expression of these genes between these two iNKT subsets^{33–35}. Globally, 17,464 regions (out of a total of 90,051 accessible regulatory regions identified) are differentially accessible between BALB/c iNKT1 and iNKT2 cells (Fig. 7a). Similarly, 11,065 regions are differentially accessible between SKG iNKT1 and iNKT2 cells (Fig. 7c). Regions that are more accessible in BALB/c iNKT1 than in BALB/c iNKT2 cells also tend to be more accessible in SKG iNKT1 than in SKG iNKT2 cells (Fig. 7b, d). Conversely, the same holds true for regions that are more accessible in iNKT2 over iNKT1 cells (Fig. 7b, d). These results suggest that the mechanisms involved in setting up the chromatin accessibility

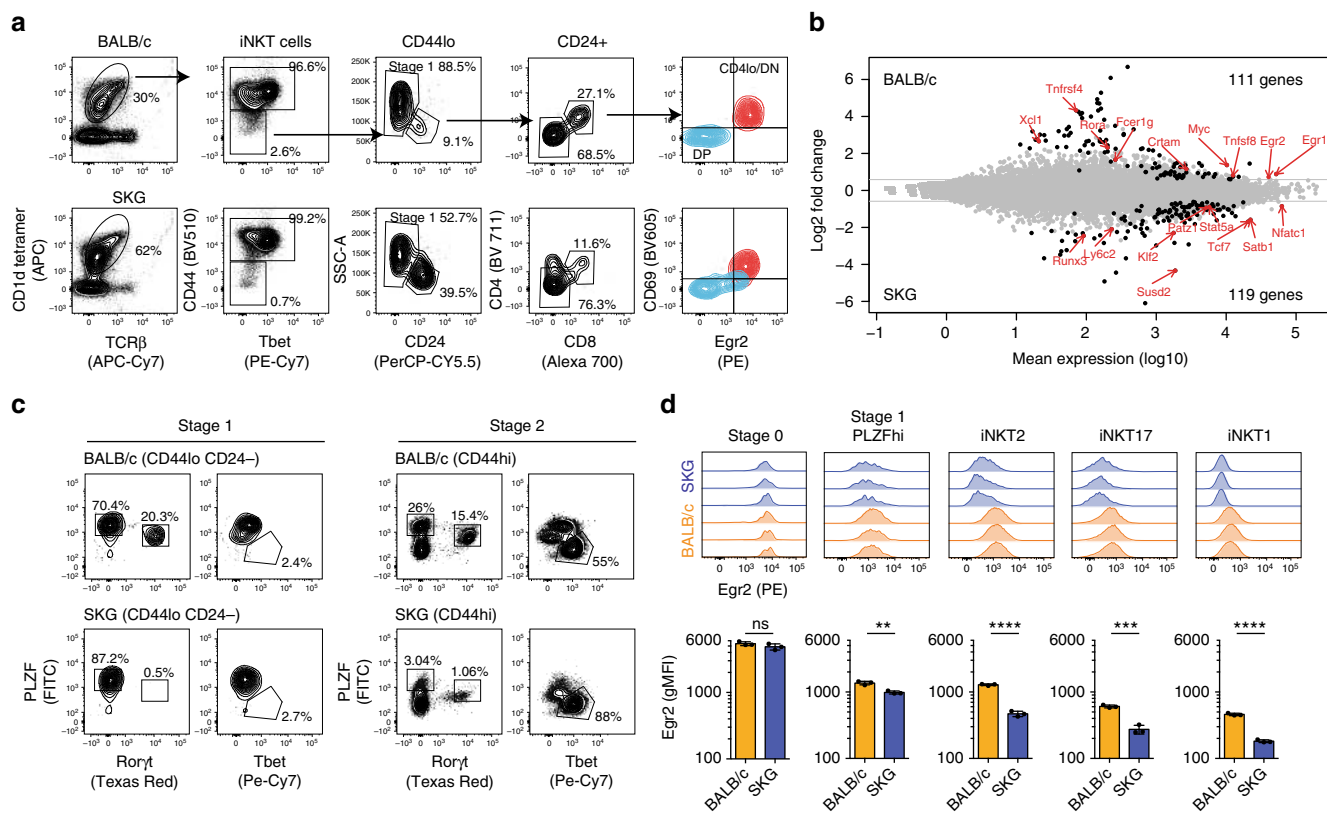


Fig. 5 Differential gene expression in iNKT cells of BALB/c and SKG mice. **a** Gating strategy used to identify and sort stage 0 iNKT cells in BALB/c (top) and SKG (bottom) mice. Enrichment of iNKT cells from the thymi of BALB/c or SKG mice was achieved after staining with APC-labeled PBS57-CD1d tetramer and enrichment using anti-APC magnetic beads. Stage 0 iNKT cells were defined as TCRβ⁺ PBS57-CD1d tetramer⁺ CD44⁻ CD24⁺ CD4^{lo} CD8⁻ CD69⁺. As seen on the figure, these cells are also Egr2⁺ (red) in contrast with the CD4⁺CD8⁺ (blue) population that invariably contaminates the cell preparation. **b** Mean average plot of genes expressed differentially in stage 0 iNKT cells isolated from SKG mice relative to their expression in such cells from BALB/c control mice. Data are representative of one experiment with three biological replicates per genotype. 238 genes (displayed in black) were found to be differentially expressed >1.5-fold with a false discovery rate <0.1. **c** Analysis of transcription factor (PLZF, Roryt, and Tbet) expression in stage 1 (CD44⁻ CD24⁻) and stage 2 (CD44⁺) iNKT cells from the thymi of BALB/c and SKG mice. Data are representative of >5 mice analyzed for each genotype. **d** Analysis of Egr2 expression level in BALB/c (orange) and SKG (purple) iNKT cells at different developmental stages. Results from three mice of each genotypes analyzed independently are shown. The mean ± SD of Egr2 gMFI in each iNKT cell developmental stage in BALB/c and SKG mice is shown. Significance was assessed using the unpaired *t*-test. ***p* < 0.01; ****p* < 0.001; *****p* < 0.0001; ns, not significant

landscape specific to each iNKT cell subset are conserved between BALB/c and SKG mice. To ask whether there are specific changes in chromatin accessibility in iNKT cells from SKG mice, we compared the ATAC-seq data between BALB/c and SKG backgrounds for each subsets. In iNKT1 cells, we identified 267 regions that are differentially accessible between BALB/c and SKG, suggesting that reduced TCR signaling only minimally impacts the differentiation of the iNKT1 lineage (Fig. 7e). By comparison, 4254 regions are differentially accessible between BALB/c and SKG iNKT2 cells (Fig. 7e), with >75% of these regions only accessible in BALB/c cells with reduced accessibility in SKGs. To group regulatory elements with similar profiles on their relative signal in each subsets between BALB/c and SKG mice, we partitioned all differentially accessible regions into 6 clusters using *k*-means. Regions in clusters 1, 4, and 5 define the iNKT1 chromatin landscape signature and are significantly enriched for T-box, Runt, and ETS binding sequences, while clusters 6 and 3 are associated with iNKT2 cells and are enriched for Tcf, Roryt, NFκB, Egr and NFAT-binding sites (Fig. 7f). Interestingly, cluster 3, which defines peaks found in BALB/c iNKT2 cells but absent from SKG iNKT2 cells, is strongly enriched for NFAT and Egr binding sites (Fig. 7b), suggesting that this unique signature is affected by the SKG mutation (Fig. 7b). Conversely, cluster 2, which defines peaks found in SKG

iNKT cells but absent from BALB/c iNKT cells, shows an enrichment for NFκB-binding sites (Fig. 7f). These results suggest that in addition to losing the NFAT/Egr signature, potentially due to a quantitative decrease in TCR signaling, the SKG mutation also alters the signal received by the cells in a qualitative manner.

PD1 expression in activated CD8 T cells is controlled by regulatory regions in the *Pcd1* locus. Accessibility to these regions has been directly linked to the binding of NFAT^{36,37}. The same regulatory regions are found accessible in BALB/c iNKT cells (Fig. 7g), with small differences between BALB/c iNKT1 and iNKT2 cells, suggesting that the locus might be differentially regulated in these subsets (Fig. 7g). Importantly however, accessibility to these sites is largely diminished in both iNKT subsets derived from SKG mice (Fig. 7g). Consequently, PD1 is poorly expressed on SKG iNKT cells compared to the wild-type cells (Fig. 7h). These results illustrate the loss of the NFAT signature in iNKT cells from the SKG mice (Fig. 7f).

While the chromatin accessibility profile of the *Zbtb16* locus (encoding PLZF) is not disturbed by the SKG mutation in any iNKT subset (Supplementary Fig 6), we noticed that the loci surrounding two of the transcription factors known to regulate the expression of PLZF in iNKT cells (Egr2 and Runx1^{15,38}) are affected (Fig. 7i, j). Within 50 Kb of the *Egr2* transcription start, we identified five putative regulatory regions, some of which are

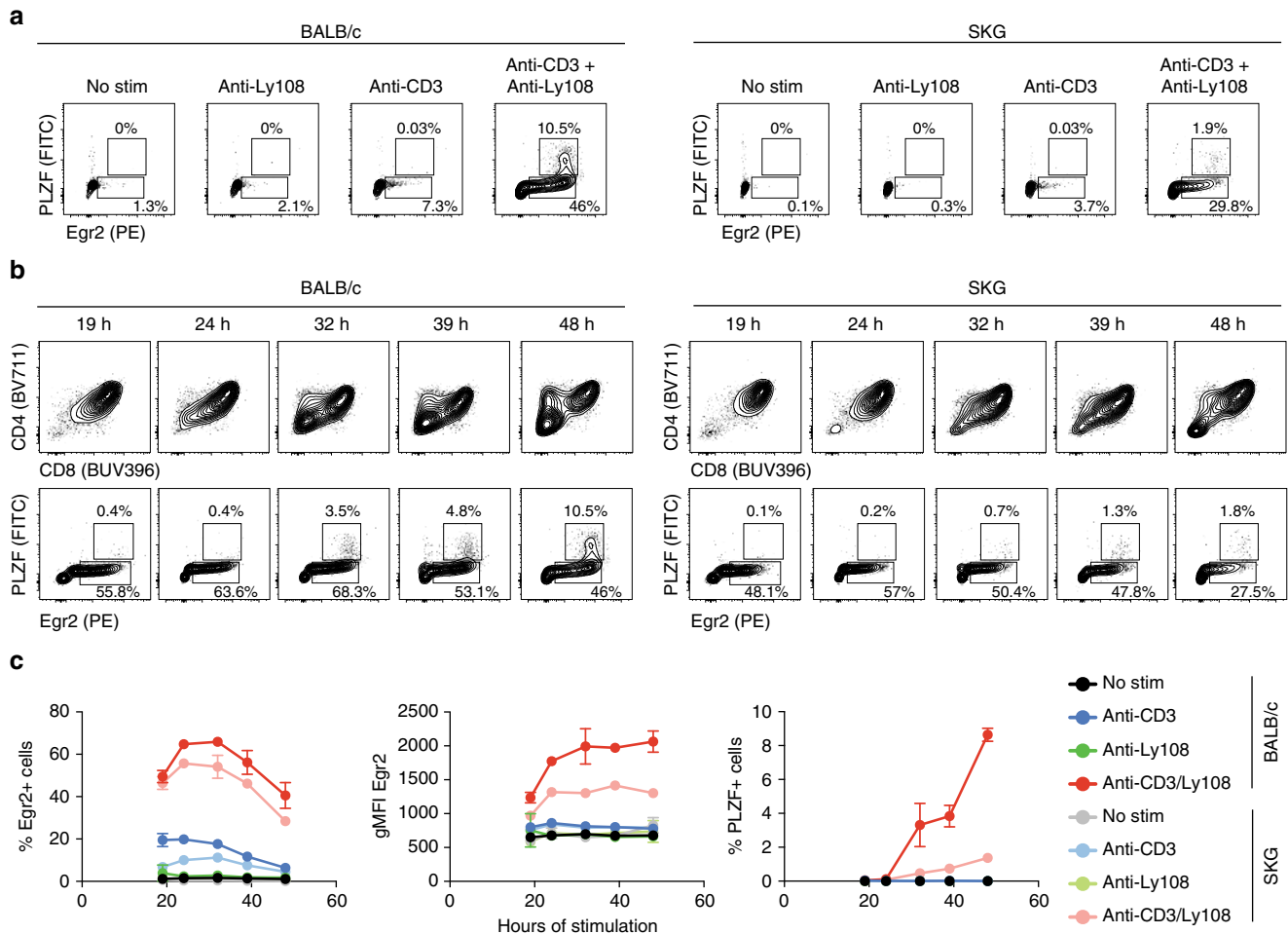


Fig. 6 Induction of Egr2 and PLZF in pre-selection DP thymocytes stimulated with plate-bound anti-CD3 and anti-Ly108 mAbs is affected by the SKG mutation. **a** Expression of Egr2 and PLZF proteins in pre-selection BALB/c or SKG DP thymocytes stimulated with media alone (no stim), plate-bound anti-CD3, anti-Ly108, or anti-CD3 + anti-Ly108 for 48 h. At the end of the stimulation, cells were stained using anti-CD4, -CD8, -Egr2, -PLZF mAbs, and a viability dye. The percentages of live cells that are Egr2⁺ and PLZF⁺ are indicated for each stimulatory condition. **b** Time course of CD4, CD8, Egr2, and PLZF expression in pre-selection BALB/c or SKG DP thymocytes stimulated with plate-bound anti-CD3 + anti-Ly108 mAbs. Data are shown for live cells in each conditions. **c** Summary of the data shown in **b** for each stimulatory condition. The percentage of Egr2⁺ cells, the gMFI of Egr2 expression, and the percentage of PLZF⁺ cells are shown. Each experimental condition was tested in duplicate wells and the data are representative of greater than three independent experiments with similar results

shown to directly bind NFAT in CD8 T cells (Fig. 7i). Importantly, accessibility to these five regions is different between BALB/c iNKT1 and iNKT2 cells, suggesting that control of Egr2 expression between iNKT cell subsets involves different regulatory mechanisms. Nevertheless, for both iNKT subsets, the SKG mutation affects the pattern of chromatin accessibility of putative Egr2 regulatory regions (Fig. 7i). These results are consistent with the diminished levels of Egr2 expression within these cells (Fig. 5d). Similar results are found for the *Runx1* locus, where subset-specific chromatin accessibility peaks can clearly be delineated (Fig. 7j). Two of these regulatory regions (one of them directly binds NFAT in CD8 T cells) are affected by the SKG mutation (Fig. 7j), suggesting that like Egr2, Runx1 expression might also have been disturbed by the SKG mutation. Finally, analysis of the transcription starting site (TSS) of the *Klf2* gene demonstrates more accessibility in BALB/c iNKT1 over iNKT2 cells, in agreement with the known differential expression of this gene between iNKT1 and iNKT2 cells^{2,35}. Accessibility at the TSS is largely increased in both iNKT subsets derived from the SKG mice (Fig. 7k). Such increase in *Klf2* expression is already apparent in stage 0 SKG iNKT cells (Fig. 5b).

To identify potential gene targets of the regulatory elements that are differentially accessible between BALB/c and SKG iNKT cells, we performed gene ontology analysis using GREAT (Genomic Regions Enrichment of Annotation Tool)³⁹ (Supplementary Fig 7). For cluster 3, we observe a significant enrichment of gene ontology (GO) terms for the regulation of kinase activity and Notch signaling (Supplementary Fig 7 and Supplementary Table 1). Several genes encoding proteins involved in the regulation of mitogen-activated protein kinases (MAPK) activity, including the dual-specificity protein phosphatases Dusp4, Dusp5, and Dusp6⁴⁰, as well as the MAPK/Erk pathway regulators sprouty proteins 1 and 2⁴¹ show differential chromatin accessibility profiles between iNKT cell subsets that are affected by the SKG mutation (Supplementary Fig. 8). Similar observations are made for the Notch pathway-associated genes, *Hey1*, *Jag1*, *Ovol2*, and *Rbpj* (Supplementary Fig 9). Interestingly, several of these peaks (but not all) bind NFAT in CD8 T cells (Supplementary Figs. 8 and 9). Altogether, these results suggest a potential link between the reduced TCR signaling in SKG mice and alterations of MAPK/Erk activity and Notch signaling in having a role in the development of iNKT cell subsets.

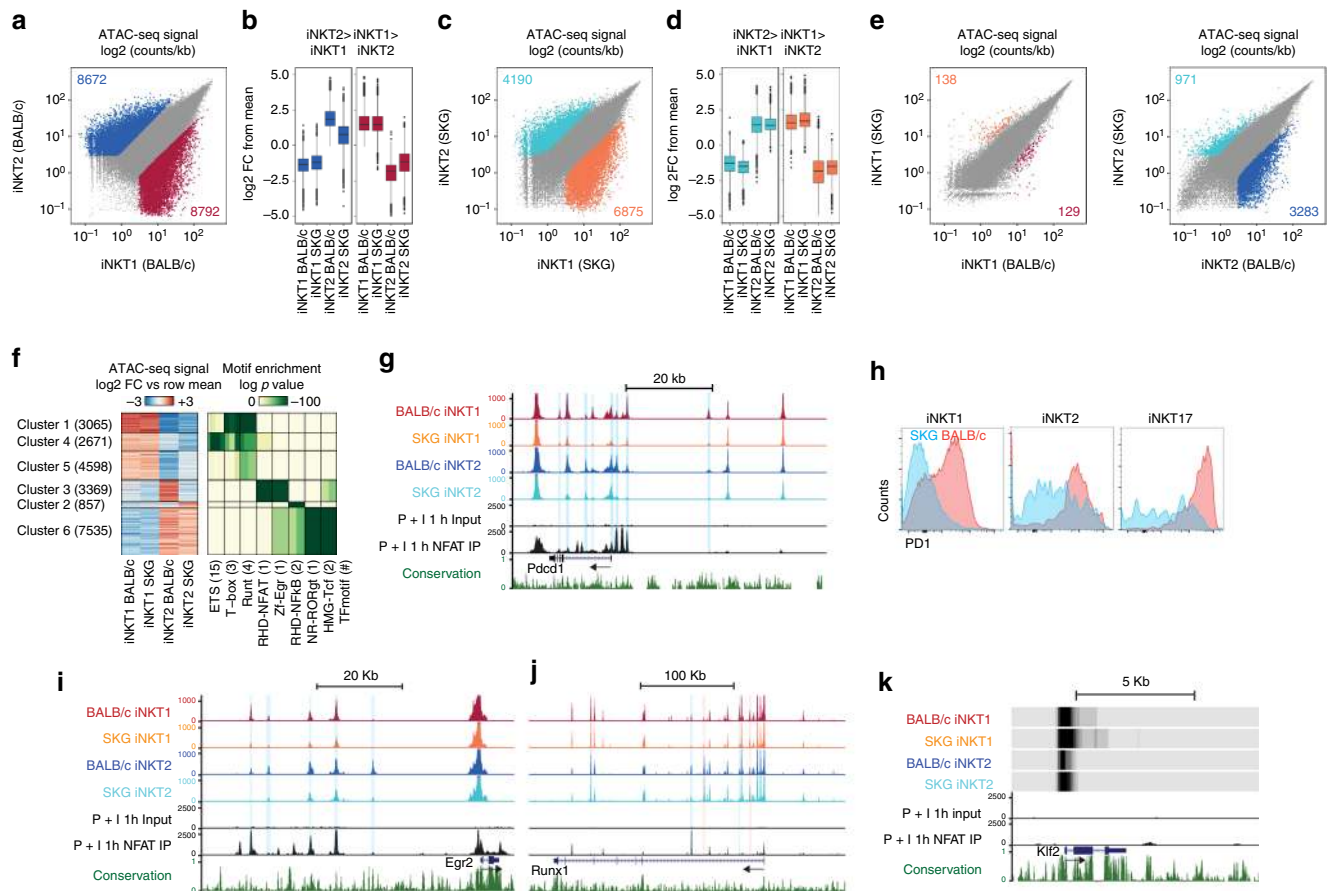


Fig. 7 Genome-wide chromatin landscapes of iNKT1 and iNKT2 subsets are affected by the SKG mutation. **a, c** Scatterplots of mean ATAC-seq counts per peak comparing iNKT1 and iNKT2 subsets from BALB/c (**a**) or SKG (**c**) mice. **b, d** Boxplots of ATAC-seq counts per plots from the indicated samples (labeled at the bottom) at common or differentially accessible regions from the comparison labeled above. Box indicates interquartile range with whiskers ± 1.5 times this range and outlier points. **e** Scatterplots of mean ATAC-seq counts per peak comparing iNKT1 from BALB/c vs SKG mice and iNKT2 subsets from BALB/c vs SKG mice. **f** k-means clustered log₂ fold-change from mean ATAC-seq signal for all differentially accessible regions compared to all differentially accessible regions in iNKT1 and iNKT2 cells from BALB/c and SKG mice. The enrichment of known motifs within each cluster of differentially accessible regions compared to all differentially accessible regions in iNKT1 and iNKT2 cells from BALB/c and SKG mice. All motifs with an enrichment log *p*-value < -35 and found in 10% or more regions in at least one cluster are shown. **g** Mean ATAC-seq coverage at the *Pdcd1* locus with a scale 0–1000 for ATAC-Seq tracks. NFAT ChIP-seq coverage from CD8 T cells stimulated or not with PMA + ionomycin for 1 h³⁷ are also shown. **h** Flow cytometry analysis of PD1 expression on thymic iNKT cell subsets from BALB/c (red) and SKG (blue) mice. **i** Mean ATAC-seq coverage at the *Egr2* locus with a scale 0–1000 for ATAC-Seq tracks. **j** Mean ATAC-seq coverage at the *Runx1* locus with a scale 0–1000 for ATAC-Seq tracks. For **g, i, and j**, ATAC peaks with differential signals where BALB/c > SKG mice are highlighted in blue, while peaks with signals SKG > BALB/c mice in at least one subset are in red. **k** ATAC-Seq of chromatin accessibility at the transcription start site of *Klf2* in thymic iNKT1 and iNKT2 subsets from BALB/c and SKG mice

Discussion

In the thymus, T cell precursors are selected to undergo different fates and functions such as becoming $\gamma\delta$ or $\alpha\beta$ T cells⁴², CD4 or CD8 T cells⁴³, positively selected or clonally deleted⁴⁴, and “conventional” or “agonist-selected” lymphocytes⁴⁵. In these processes, TCR signals cooperate with co-stimulatory molecules, cytokines, chemokines, integrins, and metabolic signals to regulate the generation of T cells with very different phenotypes and functions.

Here, we demonstrate that TCR signal strength is an essential variable in shaping the development of iNKT cell subsets. The three major thymic iNKT subsets generally have different avidities (a combination of TCR expression levels and TCR affinities) for the PBS57–CD1d complex, which is correlated with different levels of proteins whose surface expression levels reflect the TCR signal strength initially perceived during selection^{22,26–28}. Thus, as reported previously for conventional T cells^{23,25}, the strength of self-reactivity appears directly related to the strength of TCR binding to the α -galactosylceramide analog (PBS57) presented by

CD1d. Interestingly, α GC was recently identified in the mouse thymus where it was proposed to have a role in the positive selection of iNKT cells⁴⁶. iNKT2 cells stain the brightest with the PBS57–CD1d tetramer, followed by iNKT17 cells and finally iNKT1 cells. While the quantity of surface TCR available for ligand interaction reflects this differential staining, it is possible that it also mirrors the signaling capacity of the iNKT cell subsets, with high TCR expressing cells capable of sustaining TCR engagement for longer period of time that in turn affect the kinase activity of Zap70²⁹. A more direct measure of TCR signal strength received by each subset, through direct staining of the TCR-induced transcription factor Egr2, further reveals that iNKT2 cells indeed receive stronger and/or more persistent TCR signaling than iNKT17 and iNKT1 cells in vivo. This differential Egr2 expression appears to be regulated at the transcriptional level³³.

As reported previously², the frequency of V β usage varies among iNKT cell subsets. Interestingly, V β usage within a given subset varies as a function of TCR level as well. Notwithstanding

these variations, discrete levels of *Egr2* expression are observed in each iNKT cell subset, irrespective of TCR expression levels. These results support the idea that the actual V β -Va usage (and perhaps CDR3 β sequence) expressed by each iNKT cell precursor is responsible for the differential signaling observed, in good agreement with recent findings showing that iNKT cell subset development is affected by the biophysical properties of the TCR–antigen–CD1d interaction^{47,48}. That said, we do not observe disparate TCR expression levels on the surface of stage 0 precursors. It is possible that small differences in TCR expression levels on iNKT cell precursor cells exist, but cannot be resolved with the current sensitivity of flow cytometry. Additionally, it is also possible that, for equivalent levels of TCR expression, the exact nature of each TCR dictates the signal strength received by each precursor and that as a consequence of this signaling, each precursor cell modulates its TCR surface level as the cell progresses through development⁴⁹. In support of this hypothesis, iNKT2 cells express the highest levels of TCR and also express the highest levels of GATA3⁵⁰. While GATA3 is an autoactivator of its own expression⁵¹, it also directly binds to several gene loci, including *Tcr α* , *Tcr β* , *Cd3d*, and *Cd3g*⁵², and controls surface TCR expression level^{52,53}. This might provide a positive regulatory loop whereby cells that signal strongly would initially induce high levels of GATA3, leading to increased surface expression of the TCRs and increased avidity for their cognate ligands.

To test whether differential TCR signaling is required for the proper development of iNKT cell subsets, we either reduced the TCR repertoire diversity of iNKT cells through the use of TCR β transgenic mice or we reduced TCR signaling capacity of the polyclonal repertoire through the use of mice containing hypomorph mutations in the proximal tyrosine kinase Zap70. In both cases, we observe that limiting the strength of TCR signaling or the diversity of the TCR repertoire, preferentially affects, in a cell-intrinsic and genetic background independent manner, the development of iNKT2 and iNKT17 cells, while iNKT1 cells are less disturbed. These results are in agreement with the idea that the stronger TCR signaling observed in the two former subsets is indispensable to their proper development and cannot be replaced by external cues. Transcriptomic analysis of the stage 0 iNKT population from wild-type and Zap70-mutant mice demonstrates changes in expression of several transcripts coding for genes that have previously been involved in iNKT cells development, including *Egr*, *Myc*, *Tcf7*, and *Klf2*. For example, stage 0 iNKT cells from SKG mice are enriched 4.5-fold ($p_{\text{adj}} < 0.07$) in *Klf2*-encoding transcripts compared to their wild-type counterpart. Interestingly, *Klf2* deficiency results in an increase in the proportion of iNKT2 cells². Conversely, *Myc*-encoding mRNA is increased in cells derived from the BALB/c mice compared to SKG mice (2.1-fold, $p_{\text{adj}} < 10^{-6}$). In the absence of *Myc*, development of iNKT1 cells is blocked and the remaining thymic iNKT cells have a phenotype reminiscent of iNKT2 cells instead⁵⁴, suggesting a role for *Myc* in the subset differentiation of iNKT cells. These experiments argue that reduced TCR signal strength during the positive selection of iNKT cells changes the expression levels of several genes that affect subsequent subset differentiation.

Surprisingly, we observe only small reductions in expression of *Egr2* (1.6 fold, $p_{\text{adj}} < 0.5$) and *Egr1* mRNA (1.8 fold, $p_{\text{adj}} < 0.0002$) in stage 0 iNKT cells from SKG mice compared to the BALB/c mice. Analysis of *Egr2* protein levels in stage 0 cells confirms these results. This is unexpected because both genes are well known downstream targets of TCR signaling during thymic selection⁵⁵ and have been previously implicated in iNKT cell development^{14–16}. While the precise role of *Egr1* remains unclear and appears subtle in wild-type iNKT cells^{14,16}, lack of *Egr2* expression by contrast leads to severe deficiency early in iNKT

cell development¹⁵. Nevertheless, reduction of TCR signaling only minimally affects the induction of *Egr2* in early iNKT progenitors. However, analysis of the subsequent stages of development expose a severe loss of *Egr2* expression in iNKT cell from SKG mice compared to the levels observed in the same cells from BALB/c mice. Thus, in the presence of reduced Zap70 activity, the acute induction of *Egr2* appears conserved but its prolonged expression is not. This phenomenon can be mimicked in vitro upon co-stimulation of pre-selection DP thymocytes with anti-Ly108 and anti-CD3 mAbs. While very high levels of *Egr2* expression can be achieved acutely in these conditions, more than 24 h of sustained stimulation is required for the highest *Egr2*-expressing cells to express PLZF. Stimulation of SKG cells in the same conditions does not sustain the levels of *Egr2* expression over time and few SKG DP thymocytes induce PLZF. Several conclusions can be reached from these experiments. First, expression of the PLZF protein can be induced in pre-selection DP thymocytes in vitro and requires prolonged co-signaling through the TCR and Ly108 molecules. These results support the hypothesis that TCR signaling can instruct the development of “innate” (as defined by PLZF expression) lymphocytes^{15,56}, including iNKT cells³, in contrast with other data³². Second, only the highest *Egr2*-expressing cells are able to express the PLZF protein, expanding upon previous results where gene expression was examined by qPCR in entire cell populations¹⁷. Third, high *Egr2* expression levels are necessary but not sufficient to induce PLZF, as it needs to be sustained for more than 24 h. We propose that reduced TCR signaling in SKG mice affects the longevity of *Egr2* expression potentially through epigenetic mechanisms involving downstream TCR signaling molecules. Because the persistence of TCR signaling sustains cross talk with other pathways that can further modulate T cell fate^{57–60}, other events capable of modulating the strength and/or duration of signaling in iNKT cell precursors could also affect subset specification.

Global epigenetic information is more stable than gene expression profiles and propagates information over time during development and differentiation⁶¹. Changes in chromatin accessibility between iNKT1 and iNKT2 cells suggest lineage-specific activity of regulatory elements. Several regions surrounding iNKT1-expressed genes are only accessible in iNKT1 but not in iNKT2 cells. Globally, these iNKT1-specific chromatin accessibility regions are enriched for Runt, ETS and T-box binding motifs, suggesting that transcription factors belonging to these families might contribute to iNKT1 development by binding to these differentially accessible regions. Importantly, few differences in chromatin accessibility are observed between iNKT1 cells of BALB/c or SKG mice, implying that reduced TCR signaling does not appreciably affect the chromatin accessibility necessary to establish the iNKT1 program. Conversely, we detected accessible regulatory elements that are specific to iNKT2 cells. Generally, iNKT2-specific regions are enriched for Tcf, Ror γ t, *Egr*, and NFAT-binding sites, indicating that commitment to the iNKT2 program might instead require binding of these transcription factors to these regions. Remarkably, part of this iNKT2 signature is affected by the SKG mutation with 30% of BALB/c iNKT2-specific regions being inaccessible in SKG iNKT2 cells. These regions are enriched for *Egr* and NFAT-binding sites, two transcription factors with activity directly linked to TCR signaling. These results suggest that efficient development of iNKT2 cells requires higher activity of these two transcription factors for binding to these regions. Altogether, our data point towards a model of iNKT cell subset development where variable TCR signaling might induce epigenetic changes at NFAT and *Egr* binding sites that may have corresponding effects on gene expression and differentiation. Whether these effects are due to direct binding of these transcription factors and/or binding in

collaboration with other transcription factors at these sites remains to be determined.

Interestingly, the chromatin accessibility surrounding the *Egr2* locus (and *Runx1*) is different between iNKT cell subsets, with potential regulatory enhancer regions only accessible within specific subsets. Furthermore, accessibility at several of these enhancer regions is diminished and/or absent in SKG-derived cells. Thus, modulation of accessibility to these enhancer regions might have affected gene expression between subsets and between BALB/c and SKG mice. Specific deletion of these enhancer regions in vivo will be necessary to formally test this hypothesis. However, because transcriptional regulation of PLZF expression in iNKT cells has been linked to *Egr2*¹⁵ and *Runx1*³⁸, reduced expression of these two transcription factors in SKG cells would be expected to affect the expression of PLZF and thereby development of iNKT2/17 cells as well.

Our ATAC-seq analysis of iNKT2 cells in BALB/c and SKG mice also reveals that expression of several genes coding for regulators of the MAPK/Erk pathway might be affected by the SKG mutation. These might be implicated in feedback circuitry mechanisms essential in modulating the persistence of the signal⁶², as *Egr2* induction downstream of TCR also requires the Ras/MAPK/Erk pathway¹⁴. A stronger resurgence of Erk phosphorylation is observed in the stimulation of pre-selection DP thymocytes in the presence of Ly108 co-stimulation¹⁷, suggesting a potential role in regulating Erk oscillations over time and as a consequence, possibly, the dynamics of epigenetic changes and gene transcription in the cells⁶³.

Finally, the Notch signaling pathway has an important role in the differentiation of T cell effector subsets⁶⁴. The chromatin accessibility surrounding several genes involved in this pathway is also affected by the reduced TCR signaling. While it is currently unclear whether any of these potential enhancer regions modulate gene activity, the results nevertheless suggest a potential cross talk between TCR and Notch in further modulating signaling and the differentiation of iNKT subsets in the thymus⁶⁵. Future experiments will explore these possibilities.

Methods

Mice. The SKG mice on the BALB/c background³¹, and YYAA mice on the C57BL/6 background³⁰ have been described previously. IL-4 reporter mice KN266 backcrossed to the BALB/c background and the YYAA mice were graciously provided by Dr. Lee Reinhardt and Dr. Arthur Weiss, respectively. TCR transgenic mice on the C57BL/6 background expressing the TCR β chain from the DO11.10 hybridoma have been described previously⁶⁷. These mice were crossed with BALB/c mice to generate F1 expressing the transgene or not. Wild-type F1 littermates were used as controls. C57BL/6, BALB/c, and CD45.1 congenic BALB/c mice (CByJ.SJL(B6)-Ptpcr^c/J) were purchased from Jackson Laboratories. All mice were used between 8 and 10 weeks and were age-matched for each experiment. All mice were raised in a specific pathogen-free environment at the Biology Resource Center in National Jewish Health and the Office of Laboratory Animal Research at UCD. All animal procedures were approved by the NJH (AS2780-10-16) and UCD (B-64314(05)1E) Institutional Animal Care and Use Committees and were carried out in accordance with the approved guidelines.

Thymocyte isolation and flow cytometry. Single-cell suspensions were prepared from the thymus by manual disruption using syringe plunger. PBS57-CD1d1 tetramer was obtained from the National Institutes of Health Tetramer Core Facility. The complete list of surface antibodies used is as follows: Anti-CD5 (53-7.3), anti-CD6 (BX222), anti-Ly6c (AL-21), anti-V β 2 (B20.6), anti-V β 7 (TR310), anti-V β 8.1-8.2 (MR5-2), anti-V β 8.3 (1B3.3), anti-*Egr2* (erongr2), anti-CD24 (M1/69), anti-CD69 (H1.2F3). From BD Biosciences: anti-TCR β (H57-597), anti-CD8 α (53-6.7); from BioLegend: anti-CD44 (IM7); anti-CD4 (RM4-5). Surface antibody staining was done then cells were fixed and permeabilized using the Foxp3 buffer set (eBioscience). Fixed and permeabilized cells were incubated with intracellular antibodies including anti-PLZF (Mags.21F7; eBioscience), anti-Tbet (4B10; BioLegend), and anti-Roryt (Q31-378; BD Biosciences). Cells were analyzed on a BD LSRFortessa (BD Biosciences) and data were processed with FlowJo software (TreeStar).

Enrichment of CD1d-reactive thymocytes. Thymocytes were enriched for PBS57-CD1d-reactive cells by incubating thymocyte cell suspensions with PE or APC conjugated PBS57-CD1d tetramers for 45 min at 4 °C, then incubated with anti-PE or anti-APC magnetic microbeads (Miltenyi Biotec) for 15 min at 4 °C, followed by separation by using an autoMACS Pro Separator (Miltenyi Biotec) according to manufacturer's instructions. Cells were first stained for surface markers and then intracellularly stained for flow cytometric analysis.

Bone marrow chimeras. Bone marrow cells from wild-type congenic BALB/c mice (CD45.1), BALB/c mice (CD45.2), and SKG (CD45.2) mice were harvested and depleted of CD90-expressing cells using magnetic beads. Bone marrow cells from wild-type CD45.1 mice were mixed at 1:1 ratio with either cells derived from wild-type CD45.2 mice or SKG mice. Overall, 5×10^6 cells were injected intravenously into lethally irradiated wild-type BALB/c (CD45.1 \times CD45.2) recipient mice (1000 rads). Eight to ten weeks after reconstitution, chimeric mice were euthanized and thymocytes were analyzed by flow cytometry.

In vitro stimulation of pre-selection DP thymocytes. Total thymocytes from BALB/c and SKG mice were incubated with PE-labeled anti-CD127 (A7R34), anti-CD3 (145-201), anti-CD25 (PC61.5), anti-CD69 (H1.2F3), and PBS57/CD1d tetramers for 20 min at room temperature, washed and depleted using anti-PE magnetic beads and autoMACS pro separator (Miltenyi Biotec). Overall, 5×10^5 pre-selection DP thymocytes (purity > 95%, verified by flow cytometry analysis for each experiment) were then placed in wells of 96-well plates some of which had been coated overnight with 1 μ g/ml anti-CD3 (2C11) \pm 5 μ g/ml anti-Ly108 (13G3-19D). At different time points post stimulation, cells were stained with surface markers CD4 and CD8, fixed with the Foxp3 fixation/permeabilization kit and stained intracellularly for *Egr2* and PLZF. Each condition was tested in duplicate for each independent experiment.

RNA-Seq. Stage 0 iNKT cells (PBS57-CD1d tetramer⁺, CD44⁻, CD24⁺, CD69⁺, CD4^{lo/-}, CD8⁻) were sorted from the pooled thymi of three BALB/c or SKG mice ($\sim 1 \times 10^3$ cells were recovered) directly into Trizol buffer. Three independent samples were generated for each genotype. RNA quality was evaluated with the Bioanalyzer RNA pico kit (Agilent Technologies). Libraries were prepared from polyA⁺ RNA using the Whole Transcriptome Amplification Sequencing Technology SEQR kit (Sigma-Aldrich). Briefly, purified RNA was reverse transcribed using SuperScript II, Oligo dT30 VN primers, and template switching primers. A preamplification step of eight PCR cycles was performed using the Kapa HiFi Hotstart kit (Kapa Biosystems). The PCR product was purified using AMPureXP beads (Beckman Coulter) and 1 ng was further used for library preparation using the Nextera XT LibraryPrep kit (Illumina). Tagmented DNA was amplified with a 12-cycle PCR and again purified with AMPureXP beads. Library size distribution and yield were evaluated using the Bioanalyzer high-sensitivity DNA kit. Libraries were pooled at equimolar ratios and sequenced with the rapid run protocol on a HiSeq. 2500 (Illumina) with 50-nt single end cycling.

RNA-Seq analysis. Reads were aligned using Hisat2 (v.2.1.0) with unstranded default parameters and the primary GRCm38 genome assembly (downloaded 08/26/2016). Reads overlapping exons were enumerated using featureCounts (subread v.1.4.4) requiring a minimum mapping quality of 10 using GRCm38.85 ensemble annotations (downloaded 08/26/2016). Counts were analyzed using the R package DESeq2 and libraries were normalized using the default median ratio method. Differentially expressed genes were required to pass a linear fold-change threshold of 1.5 and a FDR threshold of 0.1.

ATAC-Seq. ATAC-seq was performed according to Buenrostro et al.⁶⁸ with 50,000 sorted iNKT1 (CD27⁺, CD122⁺) or iNKT2 (CD27⁺, CD122⁻) cells. iNKT17 cells were not analyzed due to the low number of cells recovered from the SKG mice. Cells were lysed in 50 μ l lysis buffer (10 mM TrisHCl, pH 7.4, 10 mM NaCl, 3 mM MgCl₂, 0.1% (v/v) Molecular biology-grade IGEPAL CA-630) to extract nuclei. Nuclei were pelleted at 500 \times g. Nuclei were resuspended in 50 μ l \times TD buffer containing 2.5 μ l transposase (Nextera, Illumina). The transposase reaction was conducted for 40 min at 37 °C. DNA was purified using MinElute PCR Purification columns and kit (Qiagen) and barcoding were performed using Illumina compatible index primers (IDTdna.com). PCR was conducted for 11–12 cycles. Library purification was performed with the MinElute PCR Purification Kit (Qiagen) and library size selection was carried out using AMPureXP (Beckman Coulter). Libraries were quantified and size distribution was assessed using the Bioanalyzer High Sensitivity DNA Kit (Agilent). Paired-end sequencing was performed on a HiSeq 2500 (Illumina) with 50 cycles for each read. Raw data from the sequencer was demultiplexed and FASTQ files were generated using bcl2fastq conversion software (Illumina).

ATAC-Seq analysis. Sequencing reads in FASTQ format were mapped to mouse genome (mm10) using bowtie (version 1.0.0) with parameters “-p 8 -m 1 --best --strata -X 2000 -S --fr --chunkmbs 1024.” Unmapped reads were processed with trim_galore using parameters “--paired --nextera --length 37 --stringency 3

–three_prime_clip_R1 1 --three_prime_clip_R2 1” before attempting to map again using the above parameters. These two bam files were merged and processed to remove reads mapping to the mitochondrial genome and duplicate reads (with picard MarkDuplicates). For visualization, genomic coverage for individual replicates were computed on 10 bp windows with MEDIPS using full fragments captured by ATAC-seq and used to generate average coverage with the Java Genomics Toolkit for each group. To identify peaks, the bam files containing unique, non-chrM reads were processed with samtools and awk using “{if(sqrt(\$9/\$9) < 100) print(0),}” to identify nucleosome free DNA fragments <100 nt in length. These subnucleosomal fragments were used to call peak summits for each replicate with MACS2 using parameters “--nomodel -q 0.00001 --keep-dup all --call-summits.” The summits for each peak from all replicates were expanded to regions with a uniform size of 500 bp. These regions from all replicates were merged into one global set of peaks and were filtered to remove peaks on the Y chromosome or those that overlapped ENCODE blacklisted regions. The number of transposase insertions within each region was computed for each replicate and these raw ATAC-seq counts per peak for all replicates were normalized using voom. Pairwise contrasts were performed with limma and differentially accessible regions were filtered based on an FDR adjusted *p*-value of < 0.01 and an estimated fold-change of at least 4. We computed the ATAC-seq density (number of transposase insertion sites per kilobase) and accessible regions were defined as those with a mean of 3 normalized insertions per kilobase. We used HOMER to identify motifs for transcription factor binding sites enriched in different groups of peaks.

Statistical analysis. Groups were compared with two-tailed Student’s *t*-tests or one-way analysis of variance (ANOVA) with Prism 7 (Graphpad Software, Inc.). **p* < 0.05, ***p* < 0.01, ****p* < 0.001, and *****p* < 0.0001.

Data availability. Data that support the findings of this study have been deposited in NCBI GEO with the accession code [GSE110327](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE110327) (RNA-Seq) and [GSE111427](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE111427) (ATAC-Seq).

Received: 25 April 2018 Accepted: 7 June 2018

Published online: 09 July 2018

References

- Bendelac, A., Savage, P. B. & Teyton, L. The biology of NKT cells. *Annu. Rev. Immunol.* **25**, 297–336 (2007).
- Lee, Y. J., Holzapfel, K. L., Zhu, J., Jameson, S. C. & Hogquist, K. A. Steady-state production of IL-4 modulates immunity in mouse strains and is determined by lineage diversity of iNKT cells. *Nat. Immunol.* **14**, 1146–1154 (2013).
- Gapin, L. Development of invariant natural killer T cells. *Curr. Opin. Immunol.* **39**, 68–74 (2016).
- Berzins, S. P., McNab, F. W., Jones, C. M., Smyth, M. J. & Godfrey, D. I. Long-term retention of mature NK1.1+ NKT cells in the thymus. *J. Immunol.* **176**, 4059–4065 (2006).
- Carty, S. A., Koretzky, G. A. & Jordan, M. S. Interleukin-4 regulates eomesodermin in CD8+ T cell development and differentiation. *PLoS ONE* **9**, e106659 (2014).
- Lee, A. et al. IL-4 induced innate CD8+ T cells control persistent viral infection. *PLoS Pathog.* **11**, e1005193 (2015).
- Renkema, K. R. et al. IL-4 sensitivity shapes the peripheral CD8+ T cell pool and response to infection. *J. Exp. Med.* **213**, 1319–1329 (2016).
- Kang, B. H. et al. PLZF+ innate T cells support the TGF- β -dependent generation of activated/memory-like regulatory T cells. *Mol. Cells* **39**, 468–476 (2016).
- White, A. J. et al. A type 2 cytokine axis for thymus emigration. *J. Exp. Med.* **214**, 2205–2216 (2017).
- Barik, S. et al. IL-4/IL-13 signaling inhibits the potential of early thymic progenitors to commit to the T cell lineage. *J. Immunol.* **199**, 2767–2776 (2017).
- White, A. J. et al. An essential role for medullary thymic epithelial cells during the intrathymic development of invariant NKT cells. *J. Immunol.* **192**, 2659–2666 (2014).
- Anderson, M. S. & Su, M. A. Aire and T cell development. *Curr. Opin. Immunol.* **23**, 198–206 (2011).
- Aschenbrenner, K. et al. Selection of Foxp3+ regulatory T cells specific for self antigen expressed and presented by Aire+ medullary thymic epithelial cells. *Nat. Immunol.* **8**, 351–358 (2007).
- Hu, T., Gimferrer, I., Simmons, A., Wiest, D. & Alberola-Ila, J. The Ras/MAPK pathway is required for generation of iNKT cells. *PLoS ONE* **6**, e19890 (2011).
- Seiler, M. P. et al. Elevated and sustained expression of the transcription factors Egr1 and Egr2 controls NKT lineage differentiation in response to TCR signaling. *Nat. Immunol.* **13**, 264–271 (2012).
- Lazarevic, V. et al. The gene encoding early growth response 2, a target of the transcription factor NFAT, is required for the development and maturation of natural killer T cells. *Nat. Immunol.* **10**, 306–313 (2009).
- Dutta, M. et al. A role for Ly108 in the induction of promyelocytic zinc finger transcription factor in developing thymocytes. *J. Immunol.* **190**, 2121–2128 (2013).
- Baglaenko, Y. et al. Invariant NKT cell activation is potentiated by homotypic trans-Ly108 interactions. *J. Immunol.* **198**, 3949–3962 (2017).
- Savage, A. K. et al. The transcription factor PLZF directs the effector program of the NKT cell lineage. *Immunity* **29**, 391–403 (2008).
- 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).
- Mao, A. P. et al. Multiple layers of transcriptional regulation by PLZF in NKT-cell development. *Proc. Natl Acad. Sci. USA* **113**, 7602–7607 (2016).
- Azzam, H. S. et al. CD5 expression is developmentally regulated by T cell receptor (TCR) signals and TCR avidity. *J. Exp. Med.* **188**, 2301–2311 (1998).
- Mandl, J. N., Monteiro, J. P., Vrsekooop, N. & Germain, R. N. T cell-positive selection uses self-ligand binding strength to optimize repertoire recognition of foreign antigens. *Immunity* **38**, 263–274 (2013).
- Fulton, R. B. et al. The TCR’s sensitivity to self peptide-MHC dictates the ability of naive CD8 T cells to respond to foreign antigens. *Nat. Immunol.* **15**, 266–274 (2014).
- Persaud, S. P., Parker, C. R., Lo, W.-L., Weber, K. S. & Allen, P. M. Intrinsic CD4 + T cell sensitivity and response to a pathogen are set and sustained by avidity for thymic and peripheral complexes of self peptide and MHC. *Nat. Immunol.* **15**, 266–274 (2014).
- Orta-Mascaró, M. et al. CD6 modulates thymocyte selection and peripheral T cell homeostasis. *J. Exp. Med.* **213**, 1387–1397 (2016).
- Martin, B. et al. Highly self-reactive naive CD4 T cells are prone to differentiate into regulatory T cells. *Nat. Commun.* **4**, 2209 (2013).
- Moran, A. E. et al. T cell receptor signal strength in Treg and iNKT cell development demonstrated by a novel fluorescent reporter mouse. *J. Exp. Med.* **208**, 1279–1289 (2011).
- Klammt, C. et al. T cell receptor dwell times control the kinase activity of Zap70. *Nat. Immunol.* **16**, 961–969 (2015).
- Hsu, L. Y., Tan, Y. X., Xiao, Z., Malissen, M. & Weiss, A. A hypomorphic allele of ZAP-70 reveals a distinct thymic threshold for autoimmune disease versus autoimmune reactivity. *J. Exp. Med.* **206**, 2527–2541 (2009).
- Sakaguchi, N. et al. Altered thymic T-cell selection due to a mutation of the ZAP-70 gene causes autoimmune arthritis in mice. *Nature* **426**, 454–460 (2003).
- Zhang, S., Laouar, A., Denzin, L. K. & Sant’Angelo, D. B. Zbtb16 (PLZF) is stably suppressed and not inducible in non-innate T cells via T cell receptor-mediated signaling. *Sci. Rep.* **5**, 12113 (2015).
- Georgiev, H., Ravens, L., Benarafa, C., Förster, R. & Bernhardt, G. Distinct gene expression patterns correlate with developmental and functional traits of iNKT subsets. *Nat. Commun.* **7**, 13116 (2016).
- Engel, I. et al. Innate-like functions of natural killer T cell subsets result from highly divergent gene programs. *Nat. Immunol.* **17**, 728–739 (2016).
- Lee, Y. J. et al. Lineage-specific effector signatures of invariant NKT cells are shared amongst $\gamma\delta$ T, innate lymphoid, and Th cells. *J. Immunol.* **197**, 1460–1470 (2016).
- Scott-Brownie, J. P. et al. Dynamic changes in chromatin accessibility occur in CD8(+) T cells responding to viral infection. *Immunity* **45**, 1327–1340 (2016).
- Martinez, G. J. et al. The transcription factor NFAT promotes exhaustion of activated CD8+ T cells. *Immunity* **42**, 265–278 (2015).
- Mao, A. P., Ishizuka, I. E., Kasal, D. N., Mandal, M. & Bendelac, A. A shared Runx1-bound Zbtb16 enhancer directs innate and innate-like lymphoid lineage development. *Nat. Commun.* **8**, 863 (2017).
- McLean, C. Y. et al. GREAT improves functional interpretation of cis-regulatory regions. *Nat. Biotechnol.* **28**, 495–501 (2010).
- Owens, D. M. & Keyse, S. M. Differential regulation of MAP kinase signalling by dual-specificity protein phosphatases. *Oncogene* **26**, 3203–3213 (2007).
- Cabrita, M. A. & Christofori, G. Sprouty proteins, masterminds of receptor tyrosine kinase signaling. *Angiogenesis* **11**, 53–62 (2008).
- Hayes, S. M., Li, L. & Love, P. E. TCR signal strength influences alphabeta/gammadelta lineage fate. *Immunity* **22**, 583–593 (2005).
- Singer, A., Adoro, S. & Park, J. H. Lineage fate and intense debate: myths, models and mechanisms of CD4- versus CD8-lineage choice. *Nat. Rev. Immunol.* **8**, 788–801 (2008).
- Hogquist, K. A., Baldwin, T. A. & Jameson, S. C. Central tolerance: learning self-control in the thymus. *Nat. Rev. Immunol.* **5**, 772–782 (2005).

45. Hogquist, K. A. & Jameson, S. C. The self-obsession of T cells: how TCR signaling thresholds affect fate 'decisions' and effector function. *Nat. Immunol.* **15**, 815–823 (2014).
46. Kain, L. et al. The identification of the endogenous ligands of natural killer T cells reveals the presence of mammalian alpha-linked glycosylceramides. *Immunity* **41**, 543–554 (2014).
47. Cruz Tleugabulova, M. et al. Discrete TCR binding kinetics control invariant NKT cell selection and central priming. *J. Immunol.* **197**, 3959–3969 (2016).
48. Clancy-Thompson, E. et al. Monoclonal invariant NKT (iNKT) cell mice reveal a role for both tissue of origin and the TCR in development of iNKT functional subsets. *J. Immunol.* **199**, 159–171 (2017).
49. Schrum, A. G., Turka, L. A. & Palmer, E. Surface T-cell antigen receptor expression and availability for long-term antigenic signaling. *Immunol. Rev.* **196**, 7–24 (2003).
50. Carr, T. et al. The transcription factor lymphoid enhancer factor 1 controls invariant natural killer T cell expansion and Th2-type effector differentiation. *J. Exp. Med.* **212**, 793–807 (2015).
51. Ouyang, W. et al. Stat6-independent GATA-3 autoactivation directs IL-4-independent Th2 development and commitment. *Immunity* **12**, 27–37 (2000).
52. Wei, G. et al. Genome-wide analyses of transcription factor GATA3-mediated gene regulation in distinct T cell types. *Immunity* **35**, 299–311 (2011).
53. Ling, K. W. et al. GATA3 controls the expression of CD5 and the T cell receptor during CD4 T cell lineage development. *Eur. J. Immunol.* **37**, 1043–1052 (2007).
54. Mycko, M. P. et al. Selective requirement for c-Myc at an early stage of V(alpha)14i NKT cell development. *J. Immunol.* **182**, 4641–4648 (2009).
55. Shao, H., Kono, D. H., Chen, L. Y., Rubin, E. M. & Kaye, J. Induction of the early growth response (Egr) family of transcription factors during thymic selection. *J. Exp. Med.* **185**, 731–744 (1997).
56. Kreslavsky, T. et al. TCR-inducible PLZF transcription factor required for innate phenotype of a subset of gammadelta T cells with restricted TCR diversity. *Proc. Natl Acad. Sci. USA* **106**, 12453–12458 (2009).
57. Lee, I. H., Li, W. P., Hisert, K. B. & Ivashkiv, L. B. Inhibition of interleukin 2 signaling and signal transducer and activator of transcription (STAT)5 activation during T cell receptor-mediated feedback inhibition of T cell expansion. *J. Exp. Med.* **190**, 1263–1274 (1999).
58. Gomez-Rodriguez, J. et al. Itk-mediated integration of T cell receptor and cytokine signaling regulates the balance between Th17 and regulatory T cells. *J. Exp. Med.* **211**, 529–543 (2014).
59. Gomez-Rodriguez, J. et al. Itk is required for Th9 differentiation via TCR-mediated induction of IL-2 and IRF4. *Nat. Commun.* **7**, 10857 (2016).
60. Yamane, H. & Paul, W. E. Early signaling events that underlie fate decisions of naive CD4(+) T cells toward distinct T-helper cell subsets. *Immunol. Rev.* **252**, 12–23 (2013).
61. Lara-Astiaso, D. et al. Immunogenetics. Chromatin state dynamics during blood formation. *Science* **345**, 943–949 (2014).
62. Alam, R. & Gorska, M. M. Mitogen-activated protein kinase signalling and ERK1/2 bistability in asthma. *Clin. Exp. Allergy* **41**, 149–159 (2011).
63. Waters, K. M., Cummings, B. S., Shankaran, H., Scholpa, N. E. & Weber, T. J. ERK oscillation-dependent gene expression patterns and deregulation by stress response. *Chem. Res. Toxicol.* **27**, 1496–1503 (2014).
64. Amsen, D., Helbig, C. & Backer, R. A. Notch in T cell differentiation: all things considered. *Trends Immunol.* **36**, 802–814 (2015).
65. Oh, S. J. et al. Notch 1 and Notch 2 synergistically regulate the differentiation and function of invariant NKT cells. *J. Leukoc. Biol.* **98**, 781–789 (2015).
66. Mohrs, K., Wakil, A. E., Killeen, N., Locksley, R. M. & Mohrs, M. A two-step process for cytokine production revealed by IL-4 dual-reporter mice. *Immunity* **23**, 419–429 (2005).
67. Marrack, P. et al. The somatically generated portion of T cell receptor CDR3a contributes to the MHC allele specificity of the T cell receptor. *eLife* **6**, e30918 (2017).
68. Buenostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y. & Greenleaf, W. J. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods* **10**, 1213–1218 (2013).

Acknowledgements

We thank Jennifer Matsuda, John Cambier, and Lee Reinhardt for their comments and support; the National Jewish Health flow cytometry facility, the Mucosal and Vaccine Research Colorado Flow Core, and the University of Colorado flow cytometry shared resource facility for assistance with cell sorting; the Center for Genes, Environment, and Health at National Jewish Health for ATAC sequencing; the Genomics and Microarray core and the RNA Bioscience Initiative at the University of Colorado Denver for RNA-Seq and the National Institutes of Health core facility for CD1d tetramers. This work was supported by National Institutes of Health Grants AI121761 and AI124076 (to L.G.) and AI18785 (to P.M.); the National Institutes of Health NIAID training grant (T32-AI07405) award (to K.T.); The American Association for Immunologists Careers in Immunology Fellowship (to L.G.); and The Cancer Center Support Grant P30CA046934.

Author contributions

L.G. wrote the manuscript and designed figures. L.G., K.T., S.H.K., and J.S.-B. designed experiments. K.T., S.H.K., J.Z., R.B., C.H., A.L., and L.H. conducted experiments and acquired data. L.G., K.T., S.H.K., J.-S.B., and K.R. analyzed and interpreted data. L.K.P., L.L.D., A.R., J.R.H., P.M., and B.O. provided invaluable reagents and insights. L.G., S.H.K., J.S.-B., and P.M. edited the manuscript. All authors approved the final manuscript.


Additional information

Supplementary Information accompanies this paper at <https://doi.org/10.1038/s41467-018-05026-6>.

Competing interests: The authors declare no competing interests.

Reprints and permission information is available online at <http://npg.nature.com/reprintsandpermissions/>

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018