



OPEN ACCESS

EDITED BY

Hashem Koohy,
University of Oxford, United Kingdom

REVIEWED BY

Takeshi Nitta,
The University of Tokyo, Japan
Graham Anderson,
University of Birmingham, United Kingdom

*CORRESPONDENCE

Mark M. W. Chong
[✉ mchong@svi.edu.au](mailto:mchong@svi.edu.au)

SPECIALTY SECTION

This article was submitted to
Systems Immunology,
a section of the journal
Frontiers in Immunology

RECEIVED 24 November 2022

ACCEPTED 21 March 2023

PUBLISHED 03 April 2023

CITATION

Oh S, Liu X, Tomei S, Luo M, Skinner JP,
Berzins SP, Naik SH, Gray DHD and
Chong MMW (2023) Distinct
subpopulations of DN1 thymocytes exhibit
preferential $\gamma\delta$ T lineage potential.
Front. Immunol. 14:1106652.
doi: 10.3389/fimmu.2023.1106652

COPYRIGHT

© 2023 Oh, Liu, Tomei, Luo, Skinner,
Berzins, Naik, Gray and Chong. This is an
open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that
the original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution or
reproduction is permitted which does not
comply with these terms.

Distinct subpopulations of DN1 thymocytes exhibit preferential $\gamma\delta$ T lineage potential

Seungyoul Oh^{1,2}, Xin Liu¹, Sara Tomei^{3,4}, Mengxiao Luo^{3,4},
Jarrod P. Skinner¹, Stuart P. Berzins^{5,6}, Shalin H. Naik^{3,4},
Daniel H. D. Gray^{3,4} and Mark M. W. Chong^{1,2*}

¹St Vincent's Institute of Medical Research, Fitzroy, VIC, Australia, ²Department of Medicine (St Vincent's), University of Melbourne, Fitzroy, VIC, Australia, ³The Walter and Eliza Hall Institute of Medical Research, Melbourne, VIC, Australia, ⁴Department of Medical Biology, University of Melbourne, Melbourne, VIC, Australia, ⁵Department of Microbiology and Immunology, University of Melbourne, Melbourne, VIC, Australia, ⁶Institute of Innovation, Science and Sustainability, Federation University Australia, Ballarat, VIC, Australia

The $\alpha\beta$ and $\gamma\delta$ T cell lineages both differentiate in the thymus from common uncommitted progenitors. The earliest stage of T cell development is known as CD4⁻CD8⁻ double negative 1 (DN1), which has previously been shown to be a heterogenous mixture of cells. Of these, only the CD117⁺ fraction has been proposed to be true T cell progenitors that progress to the DN2 and DN3 thymocyte stages, at which point the development of the $\alpha\beta$ and $\gamma\delta$ T cell lineages diverge. However, recently, it has been shown that at least some $\gamma\delta$ T cells may be derived from a subset of CD117⁻ DN thymocytes. Along with other ambiguities, this suggests that T cell development may not be as straightforward as previously thought. To better understand early T cell development, particularly the heterogeneity of DN1 thymocytes, we performed a single cell RNA sequence (scRNAseq) of mouse DN and $\gamma\delta$ thymocytes and show that the various DN stages indeed comprise a transcriptionally diverse subpopulations of cells. We also show that multiple subpopulations of DN1 thymocytes exhibit preferential development towards the $\gamma\delta$ lineage. Furthermore, specific $\gamma\delta$ -primed DN1 subpopulations preferentially develop into IL-17 or IFN γ -producing $\gamma\delta$ T cells. We show that DN1 subpopulations that only give rise to IL-17-producing $\gamma\delta$ T cells already express many of the transcription factors associated with type 17 immune cell responses, while the DN1 subpopulations that can give rise to IFN γ -producing $\gamma\delta$ T cell already express transcription factors associated with type 1 immune cell responses.

KEYWORDS

thymocyte, lineage decision, scRNAseq, T cell development, gamma delta ($\gamma\delta$) T cells

Introduction

The $\alpha\beta$ and $\gamma\delta$ T cell lineages both arise from common progenitors that seed the thymus from the bone marrow (1, 2). In mice, the earliest stages of T cell development are termed double-negative (DN) because they lack CD4 and CD8 expression. These are further subdivided into DN1 (CD44⁺CD25⁻), DN2 (CD44⁺CD25⁺), DN3 (CD44⁻CD25⁺), and DN4 (CD44⁻CD25⁻) stages. DN1 thymocytes are known to be a heterogeneous mixture of cells, based on the expression of cell surface markers such as CD24 and CD117 (c-Kit), proliferative capacity, and expression of early T lineage genes (3). Of these, the CD117⁺ fraction, which is referred to as “early thymic progenitors” (ETPs), is able to differentiate into DN2 stage cells (3) and appear to be the most efficient at generating CD4⁺CD8⁺ double-positive (DP) $\alpha\beta$ lineage thymocytes (4). These ETPs are derived from bone marrow progenitors (5), and they still have the capacity to differentiate into NK cells (3, 6, 7) suggesting that they remain uncommitted to the T cell lineage. These cells express transcriptional regulators that are associated with stemness as well early regulators of T cell identity (8). ETPs can be further divided into CD24⁻ and CD24^{lo} subpopulations, termed DN1a and DN1b cells, respectively, which are thought to have a precursor–progeny relationship (3). Interestingly, within these ETPs, there is also a CD63⁺Ly6c⁺ subpopulation that appears to be a granulocyte-committed precursor with no T cell potential (8). Thus, even within ETPs, there is significant heterogeneity.

The $\gamma\delta$ lineage has been proposed to bifurcate from the main developmental pathway at the DN2>DN3 transition, when *Tcrb/g/d* gene rearrangement occurs (9–12), because clonal assays of ETPs and DN2 thymocytes show that a large proportion at these cells can give rise to both lineages, but this biopotency is lost by the DN3 stage (10). As thymocytes do not express cell surface TCR until late DN3, it suggests that the $\alpha\beta$ vs. $\gamma\delta$ commitment occurs independently of a TCR signal. There remains, however, some debate because strong TCR signals can divert TCR β -pT α (pre-TCR)-expressing DN3 thymocytes towards the $\gamma\delta$ lineage (13), suggesting that there remains some plasticity at the DN3 stage. Subsequently, another study showed that when DN3 thymocytes express both a pre-TCR and $\gamma\delta$ TCR, the pre-TCR contributes to generating a strong TCR signal that drives $\gamma\delta$ differentiation (14). Thus, rather than revealing plasticity, strong TCR signaling at the DN3 stage may in fact be re-enforcing the differentiation of $\gamma\delta$ -committed cells.

Thus, the current favored model of early T cell development is that the ETP subpopulation of DN1 thymocytes is the most immature population in the thymus, which progress to the DN2 stage, at which point commitment toward either the $\alpha\beta$ and $\gamma\delta$ lineages is initiated. Completion of lineage commitment then occurs at the DN3 stage. The DN3 stage is also when cells are selected for expression of a functional TCR β in complex with pre-T α , which is known as β -selection, or for expression of a complete $\gamma\delta$ TCR dimer (15).

However, *in vitro* clonal assays have shown that DN3 thymocytes can give rise to a far greater frequency of $\gamma\delta$ cells than when starting

from DN2 thymocytes (10). This is inconsistent with a simple linear progression model, although this discrepancy could reflect differing survival or plating efficiencies. Moreover, a significant oversight of this model is that it ignores the large number of CD117^{lo} DN1 thymocytes that exist. These remaining CD117⁻ DN1 thymocytes can be separated into CD117^{lo}CD24^{hi} (DN1c), CD117⁻CD24^{hi} (DN1d), and CD117⁻CD24⁻ (DN1e) subpopulations. However, they were not previously considered part of the T cell developmental pathway because they do not expand as much as the DN1a and DN1b subpopulations when placed in culture (3). Moreover, they appear to differentiate faster than DN2 thymocytes when cultured on OP9-DL1 monolayers, whereas DN1a and DN2b thymocytes progress with kinetics consistent of being developmentally earlier than DN2 (3).

Potentially, these CD117^{lo} DN1 thymocytes may primarily be the progenitors of non-T lineages as it has been shown that dendritic cells can differentiate from these cells as well as from ETPs (16, 17). However, it was recently shown that a subset of IL-17-producing $\gamma\delta$ T cells are in fact derived from Sox13-expressing DN1d thymocytes, not from ETPs (18), and therefore not *via* the canonical ETP>DN2>DN3 pathway.

This finding that IL-17-producing $\gamma\delta$ T cells can differentiate from DN1d thymocytes also points to another feature of $\gamma\delta$ T development that differs from the development of most $\alpha\beta$ T cells, that effector outcomes are determined in the thymus rather than in the periphery. First, the expression of IL-17A or IFN γ by mature $\gamma\delta$ T cells correlates with V γ chain usage. V γ 2⁺ $\gamma\delta$ T cells tend to produce IL-17A while V γ 1.1⁺ $\gamma\delta$ T cells tend to produce IFN γ (19, 20). Additionally, weak TCR signaling may promote an IL-17A phenotype in $\gamma\delta$ T cells, whereas strong signaling promotes an IFN γ phenotype (21).

While there is little controversy to $\alpha\beta$ T cell development progressing *via* the DN2 and DN3 stages and that at least some $\gamma\delta$ cells bifurcate at the DN3 stage, the ambiguities described suggest that the $\alpha\beta$ vs. $\gamma\delta$ decision lineage decision is not as strict as the prevailing model and that there is still much to be clarified at these early stages of T cell development. Notably, CD117^{lo} DN1 thymocytes cannot be simply omitted from a model of T cell development as highlighted by the finding that at least some IL-17-producing $\gamma\delta$ T cells develop from DN1d thymocytes. To better characterize the earliest stage of T cell development, particularly the composition of the DN1 population, we performed single-cell RNA sequencing (scRNAseq) of mouse DN and $\gamma\delta$ thymocytes to determine the transcriptional heterogeneity at single-cell resolution. By delineating transcriptionally distinct subpopulations and assessing their lineage potential, we better clarify the composition of the DN populations, particularly of DN1 thymocytes.

Materials and methods

Mice and thymocyte preparations

Thymuses were harvested from C57BL/6 mice at 6–7 weeks of age. All experiments were approved by the St Vincent’s Hospital Animal Ethics Committee and performed under the Australian code for the care and use of animals for scientific purposes.

Total thymocytes were obtained by crushing the thymus through a metal sieve to generate a single cell suspension. The cells were washed with PBS and filtered through a 70 μm sieve to remove any clumps. CD4⁺- and CD8⁺-expressing thymocytes were depleted using anti-CD4 and CD8 magnetic-activated cell-sorting (MACS) beads (Miltenyi Biotec), according to the manufacturer's instructions. The depleted thymocyte preparation was stained with surface antibodies for sorting on a FACSAria (BD Biosciences).

Flow cytometry

For the analysis of cell surface phenotype, cells were simply stained with antibodies. For the analysis of intracellular cytokine expression, cells were first restimulated *in vitro* with 50 ng/mL PMA + 2 $\mu\text{g}/\text{mL}$ ionomycin (both Sigma-Aldrich) in the presence of Monensin (BD Biosciences) at 37°C for 2.5 h before staining with cell surface antibodies. The cells were then fixed with the Intracellular Fixation and Permeabilization buffer set (eBioscience) and stained with antibodies against cytokines. All antibodies were purchased from eBiosciences except for the anti-CD53 and TCR V γ 1.1 antibodies, which were purchased from BD Biosciences. The full list of antibodies can be found in [Supplementary Table 1](#). Flow cytometry data were acquired on LSR Fortessa III (BD Biosciences) and analyzed with the FlowJo v10.7.0 (TreeStar) software. When only analyzing cell surface phenotype, dead cells were excluded using DAPI. For intercellular cytokine analyses, the cells were not stained with DAPI but were gated on live cells determined by size.

Single-cell RNA sequencing

Sorted thymocytes were counted and checked for viability, then loaded onto the Chromium platform (10x Genomics) for scRNAseq library construction using the Single Cell V2 or V3.1 Reagent Kit according to the manufacturer's instructions. Libraries were sequenced using 150-cycle/150-bp-reads NextSeq500 (Illumina) or 300-cycle/150-bp-reads Novaseq (Illumina). Sequencing files were demultiplexed and aligned to the *Mus musculus* transcriptome reference (mm10), and count matrix was extracted using the Cell Ranger Single Cell software v2.1.1 or v4.0.0 (10x Genomics) (22).

The Illumina sequencing output was pre-processed with Seurat (v2.3 or v3.2.2) on R (v3.6.3 and 4.1.0). Cells with <500 genes, >5000 genes, or >7% mitochondria gene expression were filtered out as low-quality cells. Following normalization and removal of confounders, highly variable genes were identified and selected using the VST selection method (23). Unsupervised linear principal component analysis (PCA) was performed on these highly variable genes to group them into 20 principal components. Cell clustering was implemented using the number of components that retain >90% of variance of gene expression in the data.

DoubletFinder (v2.0.3) was applied to remove likely sequencing doublets before downstream analyses (24). The expected number of doublets was calculated as 0.75% of recovered cells. The remaining

cells were re-clustered and visualized with t-distributed stochastic neighbor embedding (t-SNE) or uniform manifold approximation and projection (UMAP) dimensional reduction. Differential expression between subclusters was carried out using the FindAllMarkers function, with default parameters; differentially expressed genes with adjusted p-value <0.05 and fold change >0.5 or <-0.5 (log2FC) were considered unless otherwise stated.

Cell cycle genes specific to the G1, S, or G2/M stages were used to perform cell cycle scoring and assign cells to their respective stage of the cell cycle (25). Cell cycle genes were regressed out using Seurat's built-in regression model.

Merging of multiple scRNAseq datasets

To compare cell types and proportions across three independent sequencing runs, the datasets were integrated as described at <https://satijalab.org/seurat/archive/v3.0/integration.html> (26). The Seurat package (v3.2.2) was used to assemble multiple distinct scRNAseq datasets into an integrated dataset and cell cycling scores were calculated. To remove technical variability, the datasets were pre-processed and normalized using SCTransform (27). To correct for experimental batch effect, integration anchors were identified between the experiments then merged using canonical correlation analysis. Linear dimensional reduction was applied and principal components that retain >90% of variance of gene expression in the data were included for downstream analysis. Unsupervised clustering was implemented on the integrated data.

Pseudotime trajectory construction

Filtered 10x data was imported into Monocle 2 by generating a cell dataset from the raw counts slot of the Seurat object. Cells were ordered into a branch pseudotime trajectory according to the procedure recommended in the Monocle 2 documentation (28). The highly variable genes identified by Seurat were chosen as ordering genes to recover pseudospacial trajectories using the setOrderingFilter, reduceDimension, and orderCells functions in Monocle 2 with default parameters. Differential expression between pseudotime states was determined using the Seurat function FindAllMarkers.

Slingshot (29) was also employed to infer developmental trajectories using the umap, clusterLabels = seurat_clusters, and start.clus = X functions, with DN1 cells fixed as the starting point.

OP9-DL1 co-cultures

Thymocyte subpopulations of interest were purified by MACS depletion followed by cell sorting and then plated onto OP9-DL1 monolayers (30). The OP9-DL1 cells were inactivated with Mitomycin C (Stem Cell) immediately prior to use. 10³ sorted thymocytes were seeded per well in 96-well plates in α MEM (Thermo Fisher) supplemented with 20% FCS (Bovogen

Biologicals), penicillin/streptomycin/gentamycin (Sigma), 2 ng/mL murine IL-7 (Peprotech), and 5 ng/mL human FLT3L (Peprotech). The media were refreshed every 2 days and freshly inactivated OP9-DL1 cells were added every 4 days.

Cellular barcoding

Sorted total DN1 thymocytes were pre-cultured on OP9-DL1 for 24 h. The cells were then transduced with barcode lentivirus library (31) in StemSpan medium (Stem Cell Technologies) by centrifugation at 900×g for 1.5 h at room temperature. A viral titer pre-determined to give 5–10% transduction efficiency was used to ensure that the cells were not transduced with multiple barcodes; 2 ng/mL murine IL-7 and 5 ng/mL human FLT3L were then added to each well and the cells were returned to the incubator. The following day, fresh α MEM with supplements was added. $\alpha\beta$ (CD90.2⁺ CD8 α ⁺ TCR $\gamma\delta$ ⁻) and $\gamma\delta$ (CD90.2⁺ CD8 α ⁻ TCR $\gamma\delta$ ⁺) lineage cells were sorted after 14 d and 20 d of the OP9-DL1 co-culture.

Barcode library construction was performed as described previously (31). The cells were lysed in 0.5 mg/ml Proteinase K (Invitrogen) in Direct PCR Lysis Buffer (Viagen) at 55°C for 2 h. The Proteinase K was then inactivated at 85°C for 30 min and 95°C for 5 min. The lysate was split into two wells for technical replicate PCRs. A first round of PCR was performed using 1× Standard-Taq magnesium-free reaction buffer pack (NEB) with 2 mM MgCl₂ (NEB), 0.2 mM dNTPs (made in house), 0.5 μ M TopLiB forward primer (TGCTGCCGTCAACTAGAACA), and 0.5 μ M of BotLiB reverse primer (GATCTCGAATCAGGCGCTTA) for 32 cycles (1 cycle at 95°C for 5 min, 30 cycles at 95°C for 15 s, 57.2°C for 15 s, and 72°C for 15 s followed by 1 cycle at 72°C for 10 min). A second round of PCR was then performed to add different Illumina index to each sample by amplifying the first-round PCR product with a sample specific Illumina forward index primer and a common Illumina reverse index primer for 32 cycles (1 cycle at 95°C for 5 min, 30 cycles at 95°C for 15 s, 57.2°C for 15 s, and 72°C for 15 s followed by 1 cycle at 72°C for 10 min). An aliquot of the PCR product was run on 2% agarose gel to check for barcode amplification, then the samples were pooled, and the DNA was cleaned using NucleoMag SPRI beads (Machery-Nagel).

The 75-cycle sequencing runs were performed on a NextSeq instrument (Illumina). The data was demultiplexed and aligned to the reference barcode library using the *processAmplicons* function from the edgeR package (32). The barcode counts were then processed in the following steps: 1) barcodes with less than two read counts in all sample/cell types were excluded from the analysis; 2) barcodes that were detected in the water control were also removed from the analysis; 3) the read count between technical replicates of the same sample was averaged and the total read count in each sample was then normalized to 100%. The normalized barcode profiles were checked for any biases by t-SNE clustering. DBSCAN (version 1.1-8) was then used to classify barcodes based on their corresponding t-SNE coordinates.

Fetal thymic organ culture (FTOC)

Fetal thymic lobes were isolated from embryos at gestational Day 14.5 following timed pregnancies of C57BL/6 female mice. They were cultured for 5 days on 0.8-mm isopore membranes (Millipore) atop surgical gelfoam sponge (Ferrosan Medical Devices) soaked in RPMI-1640 (Sigma) supplemented with 10% FCS (Bovogen Biologicals), 20 mM HEPES (Sigma-Aldrich), 50 mM 2-mercaptoethanol (Sigma), and 1.35 mM 2'-deoxyguanosine (dGuo, Sigma) to deplete endogenous thymocytes. The depleted thymic lobes were then transferred onto new sponges soaked in fresh media with supplements but without dGuo for 2 days before repopulation with thymocyte progenitors. To repopulate thymic lobes, they were placed in 20 mL hanging drop cultures on Terasaki plates (Sigma-Aldrich) containing 5×10² to 2×10³ sorted thymocytes for 24 h before returning to fresh sponges. The media were refreshed every 3–4 days. After 14 days, single cell suspensions of the thymic lobes were generated by passing through a 70- μ m sieve for analysis by flow cytometry.

Statistical analysis

Statistical testing was performed with one-way analyses of variance (ANOVA) using Prism v9 (GraphPad). p-Values are shown as * < 0.05, ** < 0.01, *** < 0.001, and **** < 0.0001 where each statistical significance was found, and all data are represented as means \pm S.E.M.

Data availability

All scRNAseq datasets have been deposited in the NCBI Gene Expression Omnibus repository under accession number GSE188913.

Results

A comprehensive transcriptional map of early T cell development at single-cell resolution

To better characterize the heterogeneity of the early stages of T cell development, we analyzed the transcriptional landscape of DN and $\gamma\delta$ thymocytes at single-cell resolution. Cells were sorted from the thymus of C57BL/6 mice for analysis by 10x scRNAseq over three independent runs (Figures 1A, B; Table 1; Supplementary Figure 1A). The first run consisted of total DN (defined as CD4⁻ CD8⁻ B220⁻ CD11b⁻ CD11c⁻ NK1.1⁻ TCR β ⁻) and TCR $\gamma\delta$ ⁺ thymocytes, the second consisted only of DN1 and DN2 (CD4⁻ CD8⁻ CD44⁺ B220⁻ CD11b⁻ CD11c⁻ NK1.1⁻ TCR β ⁻ TCR $\gamma\delta$ ⁻) cells, and the third involved sorting DN1+DN2, DN3 and TCR $\gamma\delta$ ⁺ cells separately and mixing back together post-sort at a ratio of 55% to 30% to 15%, respectively. The latter two runs were performed to

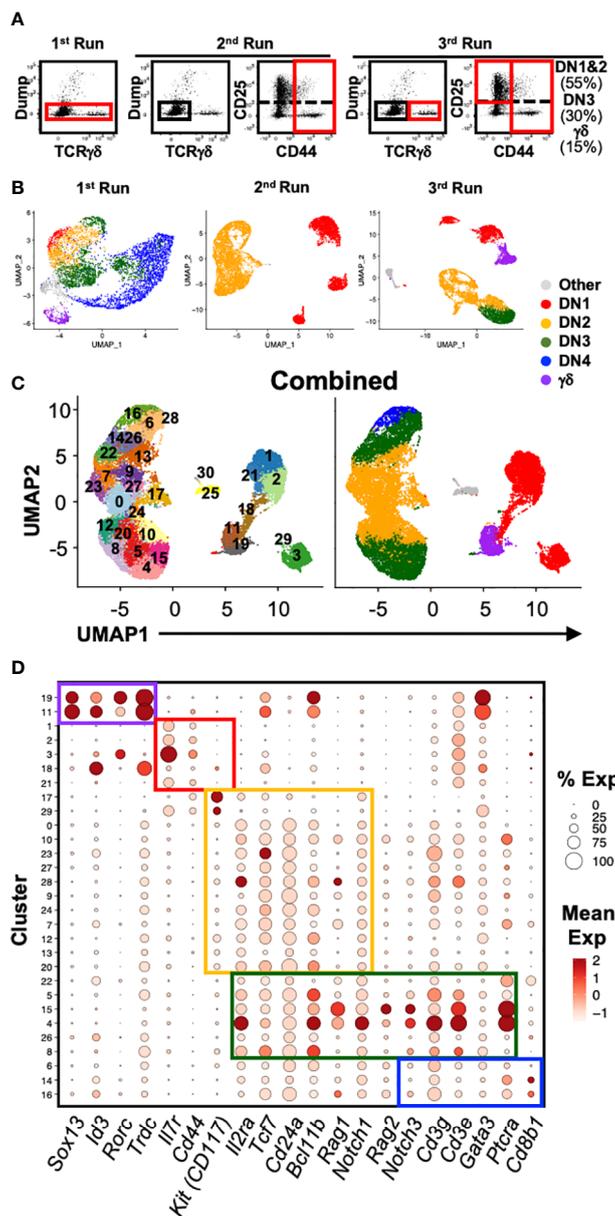


FIGURE 1

Single-cell RNA sequencing analysis of early T cell development. (A) Shown are the gating strategies used to sort DN and $\gamma\delta$ thymocytes from C57BL/6 mice for 10x scRNAseq. Three separate runs were completed: 1st = total DN and TCR $\gamma\delta^+$ cells; 2nd = Only DN1 and DN2 cells; 3rd = DN1 plus DN2, DN3 and TCR $\gamma\delta^+$ cells were sorted separately, and then mixed back together post-sort at a ratio of 55% to 30% to 15% respectively. Dump = CD4, CD8, B220, CD11b, CD11c, NK1.1 and TCR β . (B) Following processing of the 10x data on Cell Ranger, each dataset was analyzed for clustering based on the first 12 principal components in Seurat. Shown is the UMAP visualization of each run color-coded to DN developmental stage, TCR $\gamma\delta^+$ thymocytes or other (non-thymocytes). (C) The three datasets, totaling 22,094 cells, were integrated with SCTransform, then clustered with Seurat. A minimal resolution of 2.0 was selected such that no cluster contained a mixture of pre- and post-T lineage commitment cells (DN2a versus DN2b) or pre- and post- β -selection cells (DN3a versus DN3b). The resulting clusters (left plot) were then annotated to DN developmental stage, TCR $\gamma\delta^+$ thymocytes or other (non-thymocytes) (right plot). (D) Dot plots showing expression of selected markers used to assign individual clusters to DN development stage. The markers are grouped (colored boxes) based on contribution to assigning to each of the 5 broad populations.

ensure that a sufficient number of DN1 and DN2 cells were captured for high-resolution analysis that is not possible by analyzing total DN cells.

A total of 22,094 high-quality cells passed quality control checks across the three datasets. The datasets were first integrated by anchoring common cells (26) in order to assemble a global view of early T cell development. To recover biological distinction from

the different replicates and minimize batch-associated variability, the pooled data was normalized using SCTransform. Following dimensional reduction, unsupervised clustering was performed using the first 12 PCs. Next, we performed clustering of the cells at different resolutions starting from 0.5 up to 3.0. We selected a minimum value that was sufficient to separate pre- and post-T lineage commitment cells (DN2a vs. DN2b) and pre- and post- β -

TABLE 1 Summary of single-cell RNA sequencing runs.

Parameter	1 st 10x run	2 nd 10x run	3 rd 10x run
Cell sorting	DN + $\gamma\delta$	DN1 + DN2	DN1 & DN2 (55%) $\gamma\delta$ (15%) DN3 (30%)
Number of cells loaded	17,500	18,000	17,500
Targeted number of cells (recovery)	10,000	10,000	10,000
Estimated number of cells (From Cell Ranger)	5,527	8,887	8,869
Mean Reads per cell	82,999	58,813	45,036
Median Genes per cell	2,149	1,814	2,735
Median UMI Counts	5,982	4,830	8,343
Filtering (no. cells removed)	5,253 (274)	8,851 (36)	7,990 (879)
Detecting highly variable genes	2,036 HVGs (Feature Selection)	2000 HVGs (VST Selection)	2000 HVGs (VST Selection)
Dimensionality reduction	15PCs (90% variance)	11PCs (91% variance)	11PCs (90% variance)
DoubletFinder (PC distance matrix)		pK = 0.005	pK = 0.24
Number of doublets	700	750	750
Doublets removal (Dimensionality reduction)	2,066 HVGs 15PCs (90% variance)		10PCs (90% variance)
Resolution (no. clusters)	2.0 (19)	2.0 (26) (25- doublets removed)	2.0 (27)
Cell cycle regression (re-clustering)	14PCs (>90% variance)	14PCs (>90% variance)	
Cell cycle regression (re-clustering)	2.0 (19)	2.0 (25)	

selection cells (DN3a vs. DN3b) into different clusters; 2.0 was determined to be the minimum resolution. This identified 30 distinct clusters (Figure 1C; Supplementary Figure 1B), which were assigned to a canonical DN stage or $\gamma\delta$ thymocytes based on the expression of key marker genes (Figure 1D; Supplementary Figure 2). High *Cd44* and *Il7r* expression but low T lineage gene expression, including *Il2ra*, *Tcf7*, *Cd24a*, *Notch1*, and *Bcl11b*, identified DN1 cells. DN2 cells were identified by upregulation of T lineage genes and downregulation of *Il7r*. DN3 cells were identified by *Ptcra* and further upregulation of T lineage genes. Low *Cd8b1* and loss of *Il2ra* distinguished DN4 from DN3 cells. High *Trdc*, *Id3*, *Sox13*, and *Rorc* identified $\gamma\delta$ thymocytes. This indicates that multiple subpopulations correspond to each of the canonical DN stages.

Because there is a massive expansion of cell number between the ETP and DP stages, we also tested the effect of cell cycle gene expression on the clustering. Cell cycling scores were calculated from the integrated data and regressed using Seurat's built-in regression model (Supplementary Figure 3). There was a slight variation in the number of output clusters, but we did not observe any biological variability that could be explained by cell cycle status. The genetic profiles were nearly identical between the outputs with

cell cycle genes left in and regressed out. Thus, the transcriptional heterogeneity of DN thymocytes observed is not simply a result of being in a different phase of the cell cycle. This also meant that exclusion of cell cycle genes was not necessary for downstream analyses.

Trajectory analysis implies that $\gamma\delta$ T cell development branches from DN1 thymocytes

We next employed Monocle 2 (28) to infer a potential developmental pathway from the scRNAseq data of DN and $\gamma\delta$ thymocytes by ordering the cells based on tracking gene expression in pseudotime analysis (Figure 2A). This assembled the cells along an asymmetric trajectory that divided into six states (Figure 2B). Each state was then analyzed for the expression of signature genes (Figure 2C) to assign to a stage in development. State 2, comprising DN1 and some DN2a cells, was identified as pre-T lineage committed cells, and therefore the starting point. State 3 contained T lineage committed cells and cells that had passed β -selection, and therefore corresponded to the main $\alpha\beta$ pathway, with

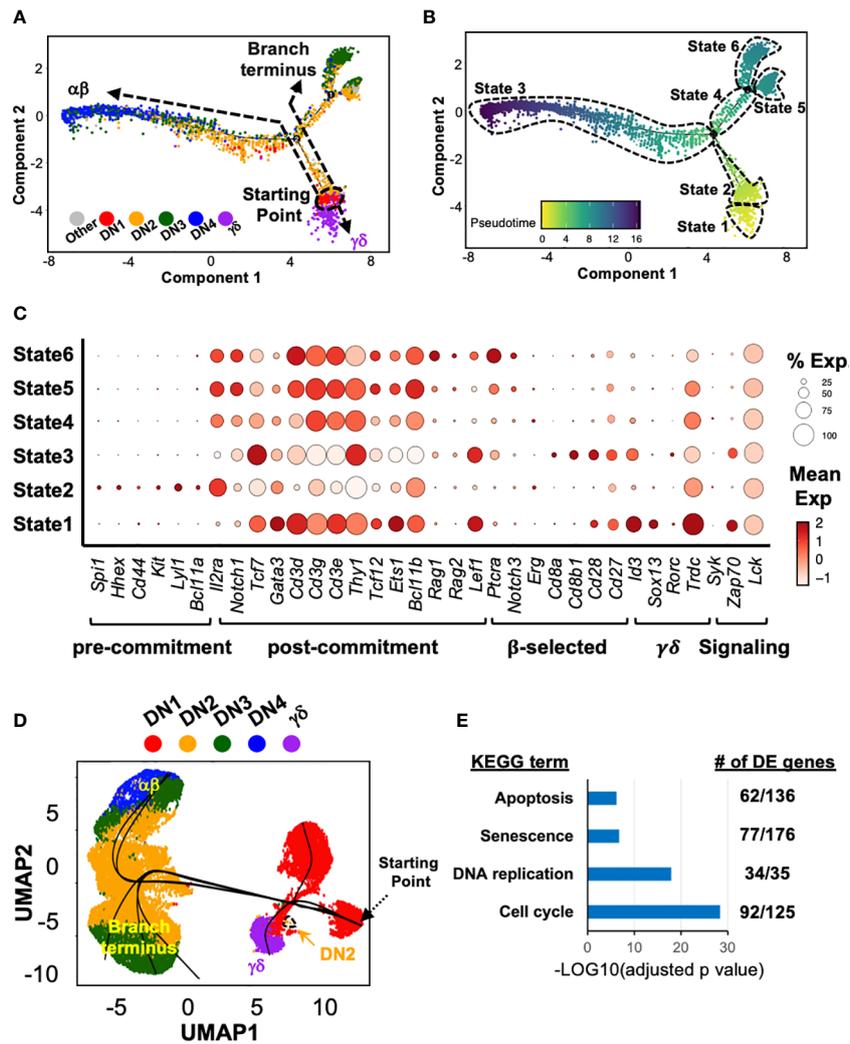


FIGURE 2 Trajectory analysis suggests that at least some $\gamma\delta$ thymocytes develop directly from DN1 thymocytes. (A) Pseudotime analysis of total DN and $\gamma\delta$ thymocytes (first 10x run) with Monocle 2. The cells are color-coded by thymocyte development stage (DN1 to 4 or $\gamma\delta$) based on expression of key marker genes (described in Supplementary Figure 2) by the individual clusters. (B) Six distinct states were identified within this asymmetric trajectory. (C) Dot plots for expression of key genes across the six states. Marker genes are grouped based on usage to assign to the indicated stages in early T cell development. (D) The three 10x datasets were integrated then subjected to Slingshot analysis, with the farthest grouping of DN1 thymocytes assigned as the starting point. (E) Differential gene expression (DE) analysis was performed on States 4, 5 and 6 (combined) compared to State 3. The DE genes were then analyzed for KEGG term enrichment. Shown are four of the terms with significant enrichment. The number of DE genes for each term is also indicated.

DN4 as the endpoint. States 5 and 6 contained T lineage committed cells but terminated with DN3 cells. State 1 corresponded to the $\gamma\delta$ branch. Thus, based simply on the transcriptomic profile of the cells, Monocle 2 appears to infer that $\gamma\delta$ cells develop directly from DN1.

We also performed trajectory analyses with another algorithm, Slingshot (29), on the integrated dataset (Figure 2D). This also suggested that $\gamma\delta$ cells develop directly from DN1. Moreover, the main $\alpha\beta$ pathway, ending with DN4, and the alternate branch that terminated with DN3 cells were also observed. However, this algorithm also predicted a DN1 branch.

We suspected that the alternate branch that terminated with DN3 cells represented cells that had not passed β -selection. We

therefore performed differential gene expression analysis between the State 4/5/6 cells with the State 3 cells. First, *Cd27* expression was upregulated in State 3 but not State 4/5/6 cells, which was previously shown to delineate pre- and post- β -selected DN3 cells (2). Furthermore, there was an enrichment of differentially expressed genes associated with “Apoptosis”, “Senescence”, “DNA replication” and “Cell Cycle” KEGG terms (Figure 2E), among others. Together, these strongly suggest that this branch represents cells that had failed β -selection.

Our *de novo* assembly of a model of early T cell development based on scRNAseq data thus suggests that the current model where $\gamma\delta$ development branches from the DN2b>DN3a stage is likely to be incomplete.

Cellular barcoding in OP9-DL1 cocultures suggests only a partial overlap of the DN1 thymocytes that give rise to $\alpha\beta$ and $\gamma\delta$ T cells

If the $\alpha\beta$ and $\gamma\delta$ lineages can develop independently, we might expect to see each lineage being derived from distinct DN1 thymocytes. To investigate this, we sorted total DN1 ($CD44^+CD25^-CD4^-CD8^-B220^-CD11b^-CD11c^-NK1.1^-TCR\beta^-TCR\gamma\delta^-$) thymocytes, which includes both $CD117^+$ ETPs and the $CD117^{-/lo}$ subpopulations. The cells were tagged with a lentiviral library of unique DNA barcodes (31). Once tagged, that barcode is inherited by the progeny of that cell and thus we can estimate how frequently $\alpha\beta$ cells and $\gamma\delta$ cells originate from the same starting DN1 cell (Figure 3A). If an $\alpha\beta$ cell and a $\gamma\delta$ cell inherit the same

barcode sequence, it means that they were derived from the same progenitor.

First, we performed a time-course of $\alpha\beta$ vs. $\gamma\delta$ differentiation from DN1 thymocytes to determine the optimal times for this analysis (Figures 3B, C). Cell surface CD8 was used as the marker of $\alpha\beta$ lineage cells because the fixation required to detect intracellular TCR β interferes with downstream analyses. Late DN4 to DP-staged cells and the few CD8 (TCR β^+) single positive cells that are generated were captured by this strategy. We also checked CD4 expression (not shown), which correlated entirely with the DP stage. CD4 (TCR β^+) single positive cells do not develop in these cultures due to lack of class II MHC presentation (33). TCR $\gamma\delta^+$ cells first appeared in these DN1 OP9-DL1 cocultures on Day 9 and reached a maximum at Day 12. CD8 $^+$ cells started appearing on Day 14 and reached a maximum at Day 20. Beyond this time point,

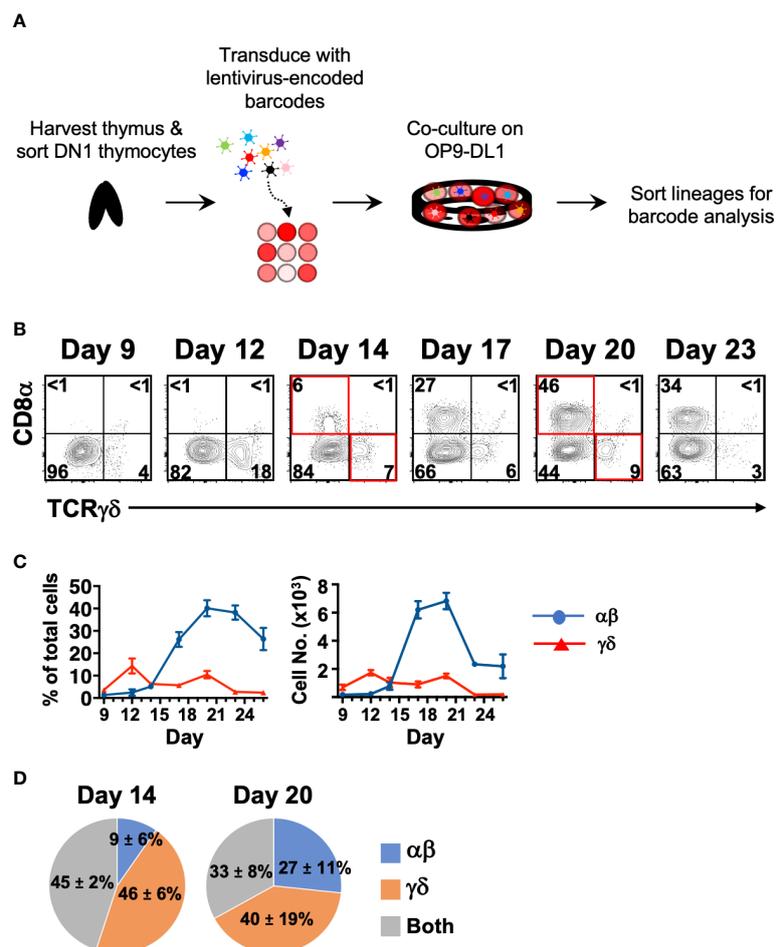


FIGURE 3

Cellular barcoding of DN1 thymocytes suggests that individual cells more frequently give rise to a single lineage than to both $\alpha\beta$ and $\gamma\delta$ cells in OP9-DL1 cocultures. (A) Overview of the experimental setup to track lineage outcomes of DN1 thymocytes. Total DN1 ($CD25^-CD44^+CD4^-CD8^-B220^-CD11b^-CD11c^-NK1.1^-TCR\beta^-TCR\gamma\delta^-$) thymocytes were sorted from the thymus of C57BL/6 mice and tagged with unique genetic barcodes encoded in a lentiviral library. The cells were then differentiated on OP9-DL1 monolayers. (B) Shown is a representative of the $\alpha\beta$ versus $\gamma\delta$ lineage profiles over a time course in these OP9-DL1 cocultures. $\alpha\beta$ lineage cells were identified as $CD8\alpha^+$, which captures cells from late DN4 onwards, while $\gamma\delta$ lineage cells were identified as TCR $\gamma\delta^+$. CD4 expression was also analyzed and was largely concomitant with $CD8\alpha$ expression as most $\alpha\beta$ cells were DP (not shown). For the cellular barcoding analysis, $\alpha\beta$ ($CD90.2^+CD8^+CD4^{+/ -}TCR\gamma\delta^-$) and $\gamma\delta$ ($CD90.2^+TCR\gamma\delta^+$) lineage cells were sorted at Day 14 from half the culture. The remaining half was further differentiated out to Day 20 and then sorted. (C) Shown are the percentages and absolute cell numbers (mean \pm S.E.M.) of 4–7 replicates over the time course starting from 10^5 total DN1 thymocytes. (D) The barcode composition of the $\alpha\beta$ and $\gamma\delta$ populations at Day 14 and 20 were analyzed by Illumina sequencing. Shown is the fraction of unique barcode sequences that were found only in the resulting $\alpha\beta$ cells, $\gamma\delta$ cells or in both populations for that time point. The values indicate the mean \pm S.E.M. of two independent sort/transduction experiments.

mature cells started dying off. We thus chose Days 14 and 20 to sort the two lineages for barcode analysis.

The DNA from the resulting $\alpha\beta$ and $\gamma\delta$ cells were amplified by PCR then Illumina-sequenced (Figure 3D). We found that at Day 14, only 45% of the detected barcode species were sequenced in both $\alpha\beta$ and $\gamma\delta$ populations, while at Day 20, only 33% of detected barcode species were sequenced in both. At both time points, the majority of unique barcode species were detected in only the $\alpha\beta$ or $\gamma\delta$ population, suggesting that these were derived from DN1 thymocytes that gave rise to only one lineage. Together with the trajectory analysis of the scRNAseq data, this suggests that there are different DN1 thymocyte subpopulations, some of which have bipotency for both lineages, while others appear to preferentially develop into either $\alpha\beta$ cells or $\gamma\delta$ cells. Thus, it is possible that at least some $\gamma\delta$ cells can develop directly from DN1 thymocytes

rather than following the canonical pathway and bifurcating from $\alpha\beta$ cells at the DN2>DN3 stage.

Transcriptional heterogeneity of DN1 and DN2 thymocytes

To better delineate the earlier stages of the developmental model, we focused the second 10x run that was performed specifically on DN1 and DN2 thymocytes. This allowed for more precise delineation of these two stages than if analyzing all DN thymocytes together. Unsupervised clustering of the 8,851 cells that pass quality control checks identified 26 clusters (Figure 4A; Supplementary Figure 4A). This was based on a resolution of 2.0, which was the minimum number that resulted in clusters

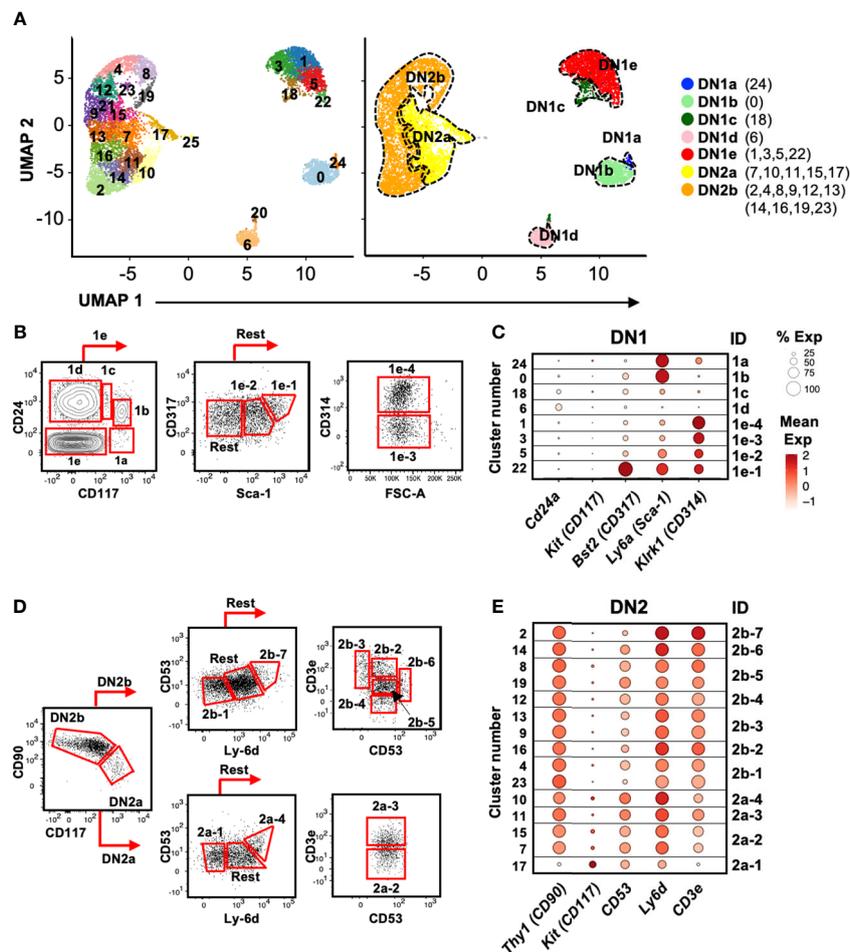


FIGURE 4 Identification of cell surface markers for delineating the DN1 and DN2 subpopulations inferred from scRNAseq. (A) UMAP visualization of 8,851 DN1 and DN2 thymocytes from the second 10x run clustered with Seurat. A minimal resolution of 2.0 was selected such that no cluster contained a mixture of pre- and post-T lineage commitment cells (DN2a versus DN2b). This yielded 26 clusters (left plot), which were then annotated as DN1a, DN1b, DN1c, DN1d, DN1e, DN2a, or DN2b (right plot). (B) The DN1 subpopulations were analyzed for differentially expressed genes encoding cell surface proteins. Antibodies against these proteins were then tested. Shown is the flow cytometric strategy to subdivide the eight DN1 subpopulations using a panel of five antibodies. Total DN1 cells were identified as CD25⁺CD44⁺CD4⁺CD8⁺B220⁺CD11b⁺CD11c⁺NK1.1⁺TCR β ⁺TCR $\gamma\delta$ ⁺. (C) Dot plots showing the expression of the genes encoding cell surface markers used to delineate the DN1 subpopulations. (D) The flow cytometric strategy to subdivide four populations of DN2a cells and seven populations of DN2b cells using a panel of five antibodies. Total DN2 cells were identified as CD25⁺CD44⁺CD4⁺CD8⁺B220⁺CD11b⁺CD11c⁺NK1.1⁺TCR β ⁺TCR $\gamma\delta$ ⁺. (E) Dot plots showing the expression of the genes encoding the cell surface markers used to delineate the DN2 subpopulations.

containing only DN1, DN2a, or DN2b thymocytes. Specifically, we checked that no cluster contained a mixture of cells with DN2a (pre-T-specification) or DN2b (post-T-specification) identity, at least based on expression of conventional stage marker genes. DN1 cells expressed high levels of progenitor markers, including *Hhex*, *Il7r* and *Cd44* but low levels of T-commitment markers, such as *Il2ra*, *Tcf7*, *Notch1*, *Cd24a*, and *Myb* (Supplementary Figure 4B). Late T-lineage commitment genes, including *Bcl11b*, *Rag1*, *Rag2*, *Notch3*, and *Ptcra*, were used to separate the DN2a and DN2b subpopulations (Supplementary Figure 4B). This resulted in eight DN1, five DN2a, and 10 DN2b subpopulations. Some clusters formed distinct subpopulations, particularly DN1 thymocytes, while others appeared to be divisions within a continuum, particularly DN2 thymocytes. Such heterogeneity within these early T cell developmental stages is consistent with that previously reported, at least for DN1 thymocytes (3). There were also three small clusters of non-T cells consisting of doublets and B cells that, for simplicity, were excluded from downstream analysis.

Identification of novel cell surface markers for delineating DN1 and DN2 subpopulations

While various cell surface markers have been used to divide DN1 thymocytes into DN1a to DN1e, and DN2 thymocytes into DN2a and DN2b (3), these are insufficient to delineate the larger number of subpopulations that are suggested by the scRNAseq analysis. We therefore needed to identify useful cell surface markers that could be used for flow cytometry. Differential expression analysis was performed to select features ($p < 0.05$ and twofold difference) for cross-referencing with GO terms for cell surface proteins (not shown). We then tested commercial antibodies against these candidates. Protein does not always correlate with mRNA levels and therefore not all antibodies produced a staining pattern that matched the scRNAseq data. However, we were able to identify a minimal panel that could delineate the eight DN1 subpopulations and most of the DN2 subpopulations.

CD24 and CD117 have previously been shown to divide DN1 thymocytes into five subpopulations (3). DN1a, DN1b, DN1c, and DN1d each corresponded to a single cluster in the scRNAseq analysis, but we identified four clusters (#1, 3, 5, and 22) that corresponded to DN1e thymocytes (Figure 4A). The inclusion of CD314 (NKG2D), CD317 (BST2), and Sca-1 delineated these four DN1e subpopulations (Figures 4B, C).

Differential CD90 and CD117 expression distinguished DN2a from DN2b cells, which were then subdivided further based on CD53, Ly-6d, and CD3e expression (Figures 4D, E). However, unlike DN1 cells that appeared to form distinct subpopulations, DN2 cells did not clearly partition into subpopulations. The pair of DN2a clusters 7/15 could not be separated by cell surface markers, nor could the pairs of DN2b clusters 4/23, 9/13, and 8/19 due to very similar gene expression profiles. Thus, these cluster pairs were combined. The resulting 11 DN2 clusters were annotated as DN2a-1 to 4 and DN2b-1 to 7 for identification purposes (not an indication of developmental order).

Using this panel of antibodies, we reliably identified the eight subpopulations of DN1 cells and 11 DN2 cells over experiments (Supplementary Figure 5), and thus, it was employed to sort subpopulations of DN1 and DN2 thymocytes for functional analyses.

OP9-DL1 cultures reveal that specific subpopulations exhibit a bias towards the $\gamma\delta$ lineage

The transcriptional heterogeneity of the DN1 thymocytes may be an indication that only some subpopulations are true progenitors of T cells, as previously suggested (3). Alternately, these different subpopulations could represent different progenitors of $\alpha\beta$ and $\gamma\delta$ lineages, which would be consistent with the early branch point inferred by pseudotime trajectory analyses (Figures 2A, B). To test this, we sorted the DN1 and DN2 subpopulations using our antibody panels and accessed their $\alpha\beta$ vs. $\gamma\delta$ lineage potential in OP9-DL1 cultures after 14 and 20 days.

At Day 14, there were primarily $\alpha\beta$ lineage or undifferentiated cells in DN1a, DN1b, and DN1c cultures, with very few TCR $\gamma\delta^+$ cells produced (Figures 5A–C). DN1d and all DN1e cultures only generated TCR $\gamma\delta^+$ cells. Notably, the number of TCR $\gamma\delta^+$ cells produced in all DN1e cultures was substantially greater than the number of TCR $\gamma\delta^+$ produced in DN1a and DN1b cultures (Figure 5C).

At 20 days, there was a substantial increase in the percentage and number of $\alpha\beta$ cells produced in DN1a and DN1b cultures, while DN1c had given rise to both $\gamma\delta$ cells and $\alpha\beta$ cells. DN1d and DN1e subpopulations continued to exhibit a bias toward the $\gamma\delta$ lineage, with DN1e-1 and DN1e-2 producing the highest percentage and number of TCR $\gamma\delta^+$ cells. In terms of absolute numbers, DN1a and DN1b subpopulations produced $\alpha\beta$ cells at a greater rate than the production of $\gamma\delta$ cells from the DN1c, DN1d, or DN1e subpopulations. Starting from 10^3 DN1a thymocytes, almost 2×10^4 $\alpha\beta$ cells were present after 20 days, whereas only $2\text{--}3 \times 10^3$ $\gamma\delta$ cells had been produced in all four DN1e cultures after this time (Figure 5C). This of course could be a reflection of $\gamma\delta$ cells differentiating and dying more quickly, but it also implies that DN1a and DN1b subpopulations are proliferating more quickly (3) and preferentially producing $\alpha\beta$ cells.

At Day 14, all DN2b subpopulations and DN2a-1 had produced a high percentage and number of $\alpha\beta$ cells and few TCR $\gamma\delta^+$ cells, while the rest of DN2a subpopulations produced mostly TCR $\gamma\delta^+$ cells and some $\alpha\beta$ cells (Figures 5D, E). By Day 20, all DN2 subpopulations had produced $\alpha\beta$ lineage cells, while TCR $\gamma\delta^+$ cells had been lost from the DN2a cultures. There was also a dramatic reduction in cell numbers in the DN2b cultures, which may be a result of the cells undergoing cell death after reaching the DP stage because these cultures poorly support the later stages of $\alpha\beta$ maturation (33).

Fetal thymic organs cultures confirm the lineage bias of DN1 subpopulations

While OP9-DL1 cultures are a well-characterized system for analyzing T cell development, it is possible that the lineages biases

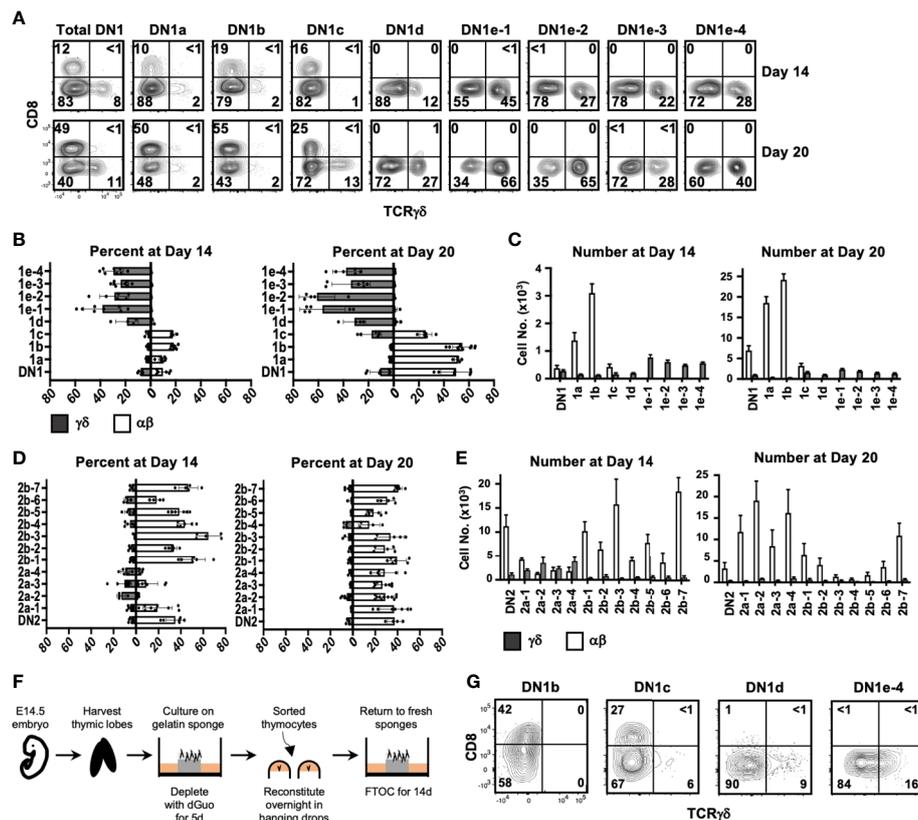


FIGURE 5

Assessing the $\alpha\beta$ or $\gamma\delta$ potential of DN1 subpopulations. (A) Sorted DN1 subpopulations were cultured on OP9-DL1 monolayers to assess their lineage potential. The cultures were analyzed at 14 and 20 days of culture by flow cytometry. $\alpha\beta$ lineage cells were identified as CD8⁺ (capturing late DN4 and DP cells) while $\gamma\delta$ lineage cells were identified as TCR $\gamma\delta$ ⁺. Representative flow cytometric plots are shown. (B, C) Pooled data analyzing $\alpha\beta$ versus $\gamma\delta$ differentiation from sorted DN1 subpopulations. The means \pm S.E.M. of four to nine replicates performed over four independent experiments are shown. See [Supplementary Tables 2, 3](#) for p-value calculations. (D, E) Pooled data analyzing $\alpha\beta$ versus $\gamma\delta$ differentiation from sorted DN2 subpopulations. The means \pm S.E.M. of four to nine replicates performed over four independent experiments are shown. See [Supplementary Tables 4, 5](#) for p-value calculations. (F) Overview of the experimental setup to repopulate dGuo-depleted FTOCs with sorted thymocytes. (G) The indicated DN1 subpopulations were sorted from the thymus of adult mice and introduced into dGuo-depleted E14.5 fetal thymic lobes. The reconstituted lobes were cultured for 14 days before analysis by flow cytometry for CD8 versus TCR $\gamma\delta$ expression. Shown is a representative from three independent experiments.

observed for the different DN1 subpopulations may be exaggerated in this model. We therefore also assessed the differentiation of select DN1 subpopulations in fetal thymic organ cultures (FTOCs), by seeding the sorted subpopulations into dGuo-depleted fetal thymic lobes ([Figure 5F](#)). The lobes were analyzed at Day 14 after repopulation. Like in the OP9-DL1 co-cultures, DN1b cells preferentially produced $\alpha\beta$ cells, DN1c cells primarily produced $\alpha\beta$ cells, and some $\gamma\delta$ cells, while DN1d and DN1e-4 cells only produced TCR $\gamma\delta$ ⁺ cells ([Figure 5G](#)).

Thus, both OP9-DL1 and FTOC systems suggest that the DN1d and DN1e subpopulations are biased toward the $\gamma\delta$ lineage. Moreover, although DN1a, DN1b, and DN1c can produce both $\alpha\beta$ cells and $\gamma\delta$ cells, they are heavily biased toward the $\alpha\beta$ lineage and produce these cells in large numbers. The DN1d and DN1e subpopulations, together, make up more than three-quarters of DN1 thymocytes. Thus, on a per cell basis, it appears that more $\gamma\delta$ cells are produced from CD117⁺ DN1 thymocytes than from what are traditionally considered the ETPs (DN1a and DN1b).

Gene expression analysis of DN1d and DN1e subpopulations suggest potential relationships with distinct mature $\gamma\delta$ subsets

Consistent with two group of $\gamma\delta$ progenitors identified by Sagar and colleagues (34), $\gamma\delta$ thymocytes clearly segregate into two clusters (#11 and 19) in our scRNAseq dataset ([Figure 1C](#)). We therefore wanted to determine how these relate to the different DN1 subpopulations that produced $\gamma\delta$ cells. Differential gene expression analysis revealed substantial differences between the two $\gamma\delta$ clusters, with 93 genes expressed at significantly higher levels by cluster 11 cells, while 117 genes were expressed at significantly higher levels by cluster 19 cells ([Figure 6A](#)). Genes that were highly expressed by cluster 11 include *Gzma*, *Blk*, *Maf*, *Sox13*, *Etv5*, *Gata3*, *Ccr9*, *Rorc*, *Sox4*, *Tcf12*, *Lgals9*, *Cmak4*, and *Bcl11b*, which are all associated with the IL-17 effector phenotype (34). This suggests that cluster 11 cells are probably the $\gamma\delta$ thymocytes that mature into the $\gamma\delta 17$

subset in the periphery. Numerous interferon-related genes, such as *Stat1*, were highly expressed by cluster 19 cells, suggesting that these cells probably mature into the IFN γ -producing $\gamma\delta$ subset. This is consistent with a previous study that suggested the eventual effector function of $\gamma\delta$ T cells may already be acquired in the thymus (21).

Next, to investigate the relationship between these $\gamma\delta$ thymocyte subpopulations and the DN1 subpopulations that differentiate into TCR $\gamma\delta^+$ cells, the DN1c, DN1d, and DN1e subpopulations were analyzed for the expression of the 210 differential genes that distinguished the two $\gamma\delta$ thymocyte populations. This revealed a similar transcriptional profile between DN1d and cluster 11 $\gamma\delta$ thymocytes, while DN1c and all the DN1e subpopulations overlapped significantly with cluster 19 $\gamma\delta$ thymocytes (Figure 6B). However, there were clearly also differences between the DN1e subpopulations.

Differential expression of transcription factors was notable. These are likely to be important because as regulators of gene expression they could potentially play key roles in hardwiring $\gamma\delta$ effector outcomes in DN1 thymocytes. *Sox13* was highly expressed by DN1d cells (Figure 6B), which was previously shown to be important for the differentiation of a subset of DN1 thymocytes into IL-17-producing $\gamma\delta$ T cells (18). Interestingly, DN1e subpopulations also express transcription factors associated with IL-17-producing $\gamma\delta$ T cells. Notably, DN1e-1 and DN1e-2 cells expressed high levels of *Maf*, while *Gata3* was highly expressed by both DN1e-1 and DN1d cells (Figure 6B). We thus predict that the

foundation of the $\gamma\delta$ effector transcriptional network may already be in place in DN1 subpopulations, with DN1d going on to develop into IL-17-producing cells and DN1e potentially producing both effector subsets.

Different DN1 subpopulations can give rise to different $\gamma\delta$ effector subsets

To determine if different DN1d and DN1e subpopulations might differentiate into distinct effector $\gamma\delta$ T cell subsets, the TCR $\gamma\delta^+$ cells that developed in OP9-DL1 co-cultures were analyzed for intracellular IL-17A and IFN γ expression and for expression of specific V γ chains (Figure 7A). Cytokine production is highly associated with specific V γ chain usage, with V γ 1.1 $^+$ cells enriched for IFN γ and V γ 2 $^+$ cells enriched for IL-17A production (19, 35).

DN1c thymocytes generated both V γ 1.1 $^+$ and V γ 2 $^+$ cells, with only a low percentage of V γ 1.1 $^+$ cells expressing IFN γ (Figures 7B–D). DN1d primarily produced V γ 2 $^+$ cells that express IL-17A. Similarly, DN1e-1 and DN1e-2 were biased towards a V γ 2 $^+$ IL-17A $^+$ $\gamma\delta$ effector subset. On the other hand, only DN1e-3 and DN1e-4 exhibited the plasticity to generate both IL-17A and IFN γ -expressing cells (Figures 7B–D). This suggests that the foundation of $\gamma\delta$ effector programs is already in place in DN1 thymocytes. Indeed, the differential gene expression analysis clearly showed that each of

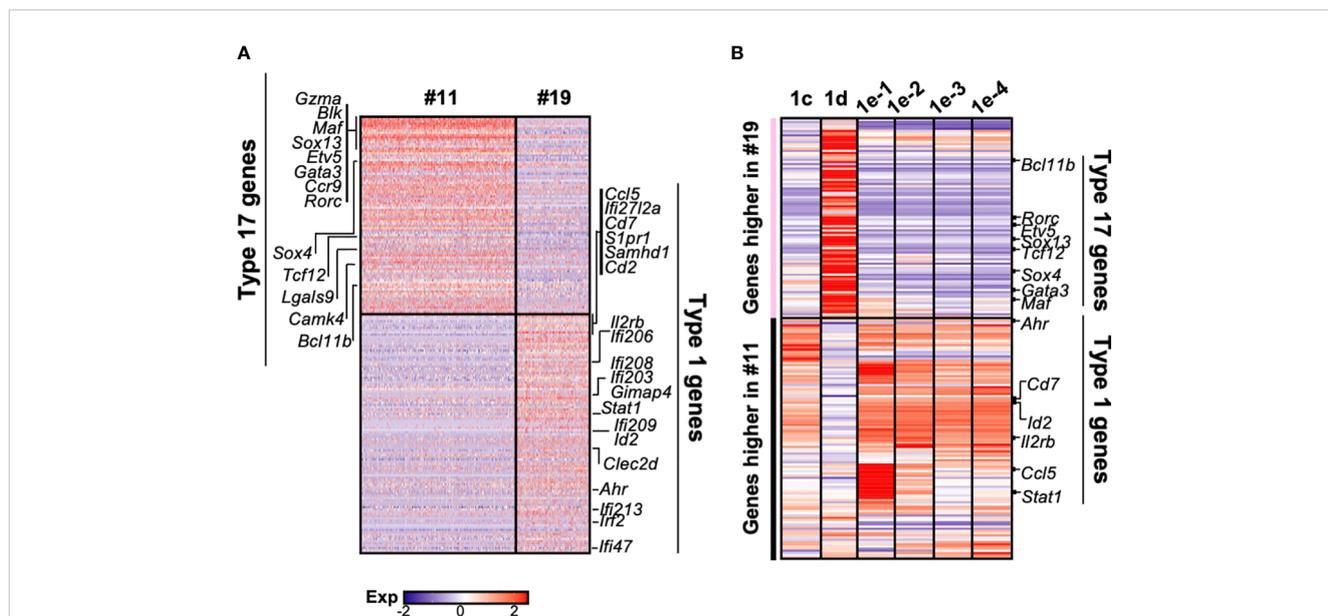


FIGURE 6

The gene expression profiles of the different DN1 subpopulations correlate with distinct $\gamma\delta$ effector subsets. (A) Heatmap showing genes differentially expressed (p -value < 0.05 and twofold difference) between the two $\gamma\delta$ thymocyte subpopulations (clusters 11 and 19) identified in the scRNAseq analysis of DN and $\gamma\delta$ thymocytes in Figure 1C. Each column is an individual cell in the dataset while each row is a differentially expressed gene. Genes associated with either type 17 immune responses (IL-17A-production) or type 1 immune responses (IFN γ production) are indicated. (B) Analysis of DN1 subpopulations for expression of the 210 genes differentially expressed between the two $\gamma\delta$ thymocyte populations. The genes are grouped based on higher expression in the cluster 11 or 19 $\gamma\delta$ thymocytes. Indicated are some of the transcription factors and cell surface receptors that are associated with either type 1 or 17 immune responses.

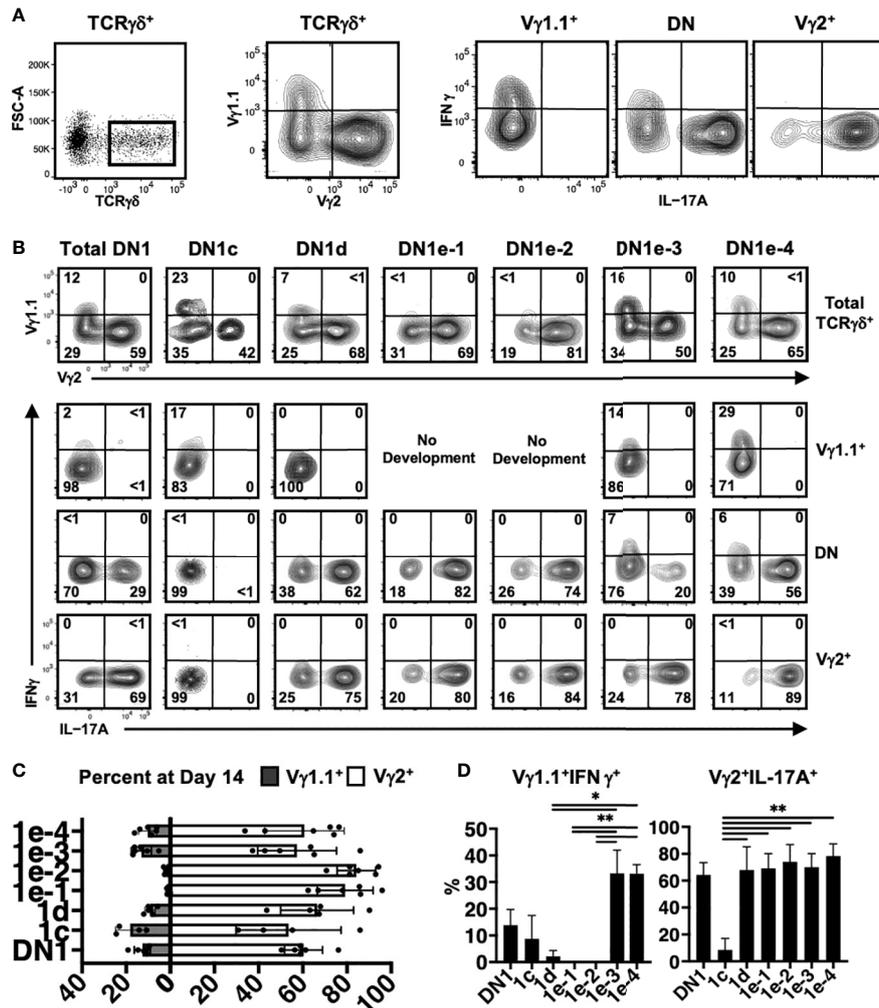


FIGURE 7
 Analysis of $\gamma\delta$ effector outcomes from DN1 subpopulations. (A) Gating strategy for analyzing the phenotype of $\gamma\delta$ T cells generated from DN1 subpopulations after culturing on OP9-DL1 monolayers for 14 days. $\text{TCR}\gamma\delta^+$ ($\text{CD4}^+\text{CD8}^-\text{TCR}\beta^+$) cells were first divided based on $\text{V}\gamma 1.1$ versus $\text{V}\gamma 2$ expression. The three subpopulations, including $\text{V}\gamma 1.1^+\text{V}\gamma 2^-$ double negative (DN) cells were then analyzed for intracellular IL-17A and IFN γ expression. (B) Shown are the flow cytometric plots from a representative experiment. The top row shows the $\text{V}\gamma 1.1$ versus $\text{V}\gamma 2$ expression on gated $\text{TCR}\gamma\delta^+$ cells. The $\text{V}\gamma 1.1^+$, $\text{V}\gamma 2^+$ and double negative (DN) cells were then analyzed for IL-17A versus IFN γ expression in the bottom three rows. (C) Pooled data analyzing the percentage of $\text{V}\gamma 1.1$ versus $\text{V}\gamma 2$ cells differentiated from sorted DN1 subpopulations. The means \pm S.E.M of four to six replicates performed over three independent experiments is shown. See [Supplementary Table 6](#) for p-value calculations. (D) Pooled data analyzing percentage of $\text{V}\gamma 1.1^+\text{IFN}\gamma^+$ (left) and $\text{V}\gamma 2^+\text{IL-17A}^+$ (right) cells out of total $\text{TCR}\gamma\delta^+$ cells. The means \pm S.E.M is shown (* $p < 0.05$, ** $p < 0.01$).

$\gamma\delta$ -producing DN1 subpopulations are transcriptionally distinct from each other, although DN1e-3 and DN1e-4 cells did appear somewhat similar to each other (Figure 6B).

Discussion

Our study has confirmed that, at a transcriptional and functional level, DN1 thymocytes are indeed a heterogeneous population of cells (3). We also confirmed that DN1a and DN1b thymocytes (CD117^+ DN1 fraction), which together have been considered the true ETPs, can give rise to both $\alpha\beta$ and $\gamma\delta$ lineages. However, we showed that $\gamma\delta$ cells can be derived from multiple progenitor sources, with the CD117^- DN1 subpopulations more efficient at producing $\gamma\delta$ cells than DN1a and DN1b

subpopulations. That being said, $\alpha\beta$ cells are produced in greater numbers because DN1a and DN1b thymocytes proliferate much more than the other DN1 subpopulations (3), and these two DN1 subpopulations only produce very few $\text{TCR}\gamma\delta^+$ cells compared to $\alpha\beta$ lineage cells. This would explain why $\text{TCR}\gamma\delta^+$ thymocytes are greatly outnumbered by $\alpha\beta$ lineage thymocytes (DN4 onwards) within the thymus.

While CD117^+ DN1 cells have been considered bipotent, a previous study demonstrated the existence of lineage-restricted precursors. Spidale and colleagues showed that neonatal IL-17-producing $\gamma\delta$ T cells develop from a subset of Sox13 -expressing DN1d thymocytes and that this lineage is determined by a cell intrinsic program that is independent of TCR signaling (18). Our study has now expanded the subdivision of DN1 thymocytes to define eight subpopulations. We showed that not only are DN1d

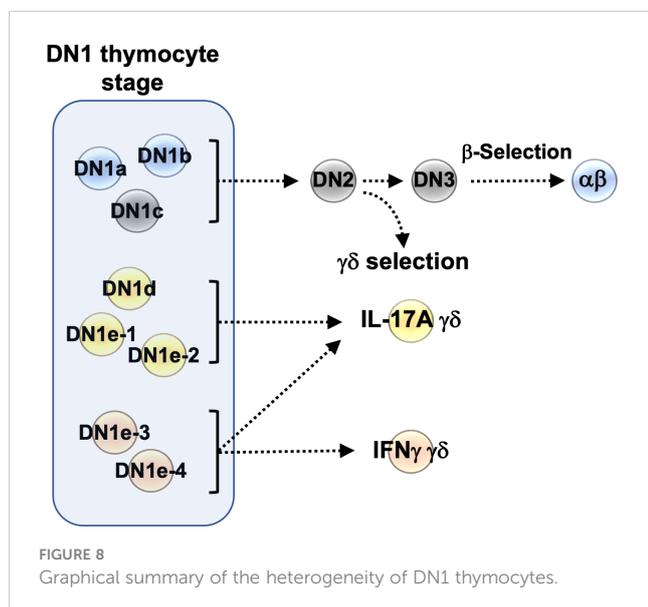
thymocytes primed towards the $\gamma\delta$ lineages but also the four DN1e subpopulations. This does not mean lineage commitment is determined entirely at this point because the progression from DN1 to fully mature $\gamma\delta$ thymocytes still requires TCR signaling to select cells that express a functional TCR $\gamma\delta$ dimer. Indeed, Scaramuzzino and colleagues showed that TCR signaling is required for $\gamma\delta$ maturation because LAT-deficient DN3 cells that express a $\gamma\delta$ TCR are unable to completely activate the $\gamma\delta$ lineage program, including expression of *Sox13*, *Maf*, and *Cpa3* (36).

Evidence suggests that TCR signal strength is an important determinant of lineage outcome in bipotential precursors. Strong signals that activate the ERK-EGR-ID3 pathway have been shown to drive $\gamma\delta$ T cell differentiation, whereas a weak signal promotes the $\alpha\beta$ fate (37–39). However, it is still unclear whether TCR signal strength is dependent on instructive extrinsic signals or simply a result of stochastic selection of the TCR chains that are expressed. The instructive model proposes that TCR $\gamma\delta$ signals compete with pre-TCR(β) signals and that the lineage decision is determined by cell-specific interactions that activate key transcription factors, which in turn instructs the gene expression program (15). In contrast, the stochastic selective model postulates that lineage commitment is determined prior to the onset of TCR gene rearrangement, and it is only when a thymocyte receives the appropriate TCR signal that matches its hardwired identity that it actually progresses along the $\alpha\beta$ or $\gamma\delta$ developmental pathway. DN1d and DN1e thymocytes are clearly committing to $\gamma\delta$ lineage cells prior to the expression of the TCR. The fact that a DN3 thymocyte expressing both a pre-TCR and $\gamma\delta$ TCR results in an even stronger signal that drives $\gamma\delta$ differentiation (14) also points towards stochastic selection. This does not rule out the role of selection because the appropriate antigen presenting cell may still be required for $\gamma\delta$ maturation.

Not only do a large fraction of $\gamma\delta$ thymocytes appear to be derived from distinct DN1 thymocytes prior to TCR signaling, but we also observed compartmentalization of effector outcomes (Figure 8). Unlike $\alpha\beta$ T cells, $\gamma\delta$ T cells are thought to acquire

their effector potential in the thymus rather than upon antigen exposure in secondary lymphoid organs (40). Although none of the DN1 subpopulations expressed definitive markers of specific mature T cell populations, we observed substantial transcriptomic overlap between the IL-17-primed DN1 subpopulations with mature IL-17-expressing $\gamma\delta$ thymocytes and between the IFN γ -primed DN1 subpopulations with mature IFN γ -expressing $\gamma\delta$ thymocytes. Expression of key transcription factors was notable. We showed that DN1d thymocytes express many of the transcription factors that are expressed by mature IL-17-producing $\gamma\delta$ thymocytes but not IFN γ -producing thymocytes, like *Bcl11b*, *Etv5*, *Sox13*, *Rcf7*, *Rorc*, and *Maf*. SOX13 has previously been shown to be an important lineage determining factor for the neonatal IL-17A-producing cells (18). It is thought to act in concert with other transcription factors like BC11B, TCF7, ROR $\gamma\delta$ and c-MAF to specify the IL-17A-effector program in $\gamma\delta$ T cells (18, 41).

While DN1e-1 and DN1e-2 thymocytes were also biased towards a V γ 2⁺ IL-17A-producing $\gamma\delta$ T cell fate, only DN1e-3 and DN1e-4 thymocytes displayed the plasticity to also produce V γ 1.1⁺ IFN γ ⁺ cells. This plasticity is likely to involve the differential expression of transcription factors that contribute to distinct effector fates. Although all four DN1e subpopulations expressed *Stat1*, DN1e-1 and DN1e-2 also expressed *Maf*. c-MAF is known to positively regulate IL-17A-producing $\gamma\delta$ T cell development (42), whereas a lack of c-MAF expression by $\gamma\delta$ T cells correlates with increased IFN γ expression (42–44). Furthermore DN1e-1 cells were found to express *Gata3*, encoding another important regulator of IFN γ expression (45). The interplay between transcription factors may thus be key to lineage decisions. Thus, critical components of the IL-17 or IFN γ -producing $\gamma\delta$ T cell transcriptional programs appear to be already in place within distinct DN1 subpopulations, suggesting that predetermination contributes to $\gamma\delta$ effector subset differentiation. Interestingly, Shibata et al. previously showed that the DN2 thymocyte stage contains a heterogeneous mixture of $\gamma\delta$ T cell precursors that give rise to either IL-17 producers or non-producers (46), thus further suggesting that $\gamma\delta$ effector outcomes may indeed be hardwired from these stages in T cell development. Further analysis of chromatin states and epigenetic mechanisms associated with these transcription factors will likely be valuable for revealing the regulatory cascades that drive the different $\gamma\delta$ effector outcomes.



Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary Material.

Ethics statement

The animal study was reviewed and approved by St Vincent's Hospital Animal Ethics Committee.

Author contributions

Conceptualization, SO, SN, DG, and MC; Methodology, SO, XL, ST, ML, JS, SB, and MC; Software, XL; Validation: SO and XL; Formal Analysis, SO and XL; Investigation, SO, XL, ST, ML, JS, and SB; Data Curation, XL and ST; Writing - Original Draft, SO, XL, and ST; Writing - Review and Editing, SN, DG, and MC; Visualization, SO and XL; Supervision, MC; Funding Acquisition, MC. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by grants and fellowships from the National Health and Medical Research Council, Australia (1078763, 1090236, 1145888, and 1158024 to DG and 1079586, 1117154, 1122384, and 1122395 to MC), Cancer Council Victoria (1102104 to DG), Diabetes Australia (Y20G-CHOM to MC), and U.S. Department of Defense (W81XWH-19-1-0728 to MC) and the Victorian State Government Operational Infrastructure Support and the Independent Research Institutes Infrastructure Support Scheme of the National Health and Medical Research Council, Australia.

Acknowledgments

We thank Lucy Kloboucnik in the St Vincent's Hospital Melbourne's BioResources Centre for expert animal husbandry,

References

- Ciofani M, Zuniga-Pflucker JC. Determining gammadelta versus alpha T cell development. *Nat Rev Immunol* (2010) 10(9):657–63. doi: 10.1038/nri2820
- Taghon T, Yui MA, Pant R, Diamond RA, Rothenberg EV. Developmental and molecular characterization of emerging β - and $\gamma\delta$ -selected pre-T cells in the adult mouse thymus. *Immunity* (2006) 24(1):53–64. doi: 10.1016/j.immuni.2005.11.012
- Porritt HE, Rumpf LL, Tabrizifard S, Schmitt TM, Zuniga-Pflucker JC, Petrie HT. Heterogeneity among DN1 prothymocytes reveals multiple progenitors with different capacities to generate T cell and non-T cell lineages. *Immunity* (2004) 20(6):735–45. doi: 10.1016/j.immuni.2004.05.004
- Matsuzaki Y, Gytoku J, Ogawa M, Nishikawa S, Katsura Y, Gachelin G, et al. Characterization of c-kit positive intrathymic stem cells that are restricted to lymphoid differentiation. *J Exp Med* (1993) 178(4):1283–92. doi: 10.1084/jem.178.4.1283
- Benz C, Martins VC, Radtke F, Bleul CC. The stream of precursors that colonizes the thymus proceeds selectively through the early T lineage precursor stage of T cell development. *J Exp Med* (2008) 205(5):1187–99. doi: 10.1084/jem.20072168
- Ikawa T, Kawamoto H, Fujimoto S, Katsura Y. Commitment of common T/Natural killer (NK) progenitors to unipotent T and NK progenitors in the murine fetal thymus revealed by a single progenitor assay. *J Exp Med* (1999) 190(11):1617–26. doi: 10.1084/jem.190.11.1617
- Michie AM, Carlyle JR, Schmitt TM, Ljutic B, Cho SK, Fong Q, et al. Clonal characterization of a bipotent T cell and NK cell progenitor in the mouse fetal thymus. *J Immunol* (2000) 164(4):1730–3. doi: 10.4049/jimmunol.164.4.1730
- Zhou W, Yui MA, Williams BA, Yun J, Wold BJ, Cai L, et al. Single-cell analysis reveals regulatory gene expression dynamics leading to lineage commitment in early T cell development. *Cell Syst* (2019) 9(4):321–37.e9. doi: 10.1016/j.cels.2019.09.008
- Kreslavsky T, Gleimer M, Garbe AI, von Boehmer H. $\alpha\beta$ versus $\gamma\delta$ fate choice: counting the T-cell lineages at the branch point. *Immunol Rev* (2010) 238(1):169–81. doi: 10.1111/j.1600-065X.2010.00947.x
- Ciofani M, Knowles GC, Wiest DL, von Boehmer H, Zuniga-Pflucker JC. Stage-specific and differential notch dependency at the $\alpha\beta$ and $\gamma\delta$ T lineage bifurcation. *Immunity* (2006) 25(1):105–16. doi: 10.1016/j.immuni.2006.05.010
- Fehling HJ, Krotkova A, Saint-Ruf C, von Boehmer H. Crucial role of the pre-t-cell receptor α in development of $\alpha\beta$ but not $\gamma\delta$ T cells. *Nature* (1995) 375(6534):795–8. doi: 10.1038/375795a0
- Oh S, Gray DHD, Chong MMW. Single-cell RNA sequencing approaches for tracing T cell development. *J Immunol* (2021) 207(2):363–70. doi: 10.4049/jimmunol.2100408
- Kreslavsky T, Garbe AI, Krueger A, von Boehmer H. T Cell receptor-instructed $\alpha\beta$ versus $\gamma\delta$ lineage commitment revealed by single-cell analysis. *J Exp Med* (2008) 205(5):1173–86. doi: 10.1084/jem.20072425
- Zarin P, Wong GW, Mohtashami M, Wiest DL, Zuniga-Pflucker JC. Enforcement of $\gamma\delta$ -lineage commitment by the pre-t-cell receptor in precursors with weak $\gamma\delta$ -TCR signals. *Proc Natl Acad Sci U S A* (2014) 111(15):5658–63. doi: 10.1073/pnas.1312872111
- Carpenter AC, Bosselut R. Decision checkpoints in the thymus. *Nat Immunol* (2010) 11(8):666–73. doi: 10.1038/ni.1887
- Luche H, Ardouin L, Teo P, See P, Henri S, Merad M, et al. The earliest intrathymic precursors of CD8 α^+ thymic dendritic cells correspond to myeloid-type double-negative 1c cells. *Eur J Immunol* (2011) 41(8):2165–75. doi: 10.1002/eji.201141728
- Moore AJ, Sarmiento J, Mohtashami M, Braunstein M, Zuniga-Pflucker JC, Anderson MK. Transcriptional priming of intrathymic precursors for dendritic cell development. *Development* (2012) 139(2):373–84. doi: 10.1242/dev.069344
- Spidale NA, Sylvia K, Narayan K, Miu B, Frascoli M, Melichar HJ, et al. Interleukin-17-Producing $\gamma\delta$ T cells originate from SOX13 $^+$ progenitors that are

and Anthony Di Carluccio from St Vincent's Institute's Cytometry Facility and Casey Anttila from Walter and Eliza Hall Institute's Cytometry Facility for assistance with cell sorting and the scRNAseq libraries. We also thank Dr. David Izon (St Vincent's Institute) for the OP9-DL1 cell line, who originally obtained them from Dr. Juan Carlos Zúñiga-Pflucker (University of Toronto).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fimmu.2023.1106652/full#supplementary-material>

- independent of $\gamma\delta$ TCR signaling. *Immunity* (2018) 49(5):857–72.e5. doi: 10.1016/j.immuni.2018.09.010
19. Narayan K, Sylvia KE, Malhotra N, Yin CC, Martens G, Vallerskog T, et al. Intrathymic programming of effector fates in three molecularly distinct $\gamma\delta$ T cell subtypes. *Nat Immunol* (2012) 13(5):511–8. doi: 10.1038/ni.2247
20. Fahl SP, Coffey F, Wiest DL. Origins of $\gamma\delta$ T cell effector subsets: a riddle wrapped in an enigma. *J Immunol* (2014) 193(9):4289–94. doi: 10.4049/jimmunol.1401813
21. Sumaria N, Grandjean CL, Silva-Santos B, Pennington DJ. Strong TCRgammadelta signaling prohibits thymic development of IL-17A-Secreting $\gamma\delta$ T cells. *Cell Rep* (2017) 19(12):2469–76. doi: 10.1016/j.celrep.2017.05.071
22. Zheng GX, Terry JM, Belgrader P, Ryvkin P, Bent ZW, Wilson R, et al. Massively parallel digital transcriptional profiling of single cells. *Nat Commun* (2017) 8:14049. doi: 10.1038/ncomms14049
23. Zwiener I, Frisch B, Binder H. Transforming RNA-seq data to improve the performance of prognostic gene signatures. *PLoS One* (2014) 9(1):e85150. doi: 10.1371/journal.pone.0085150
24. McGinnis CS, Murrow LM, Gartner ZJ. DoubletFinder: Doublet detection in single-cell RNA sequencing data using artificial nearest neighbors. *Cell Syst* (2019) 8(4):329–37.e4. doi: 10.1016/j.cels.2019.03.003
25. Macosko EZ, Basu A, Satija R, Nemes J, Shekhar K, Goldman M, et al. Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* (2015) 161(5):1202–14. doi: 10.1016/j.cell.2015.05.002
26. Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM3rd, et al. Comprehensive integration of single-cell data. *Cell* (2019) 177(7):1888–902.e21. doi: 10.1016/j.cell.2019.05.031
27. Hafemeister C, Satija R. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome Biol* (2019) 20(1):296. doi: 10.1186/s13059-019-1874-1
28. Qiu X, Mao Q, Tang Y, Wang L, Chawla R, Pliner HA, et al. Reversed graph embedding resolves complex single-cell trajectories. *Nat Methods* (2017) 14(10):979–82. doi: 10.1038/nmeth.4402
29. Street K, Risso D, Fletcher RB, Das D, Ngai J, Yosef N, et al. Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics. *BMC Genomics* (2018) 19(1):477. doi: 10.1186/s12864-018-4772-0
30. Holmes R, Zuniga-Pflucker JC. The OP9-DL1 system: generation of T-lymphocytes from embryonic or hematopoietic stem cells *in vitro*. *Cold Spring Harb Protoc* (2009) 2009(2):prot5156. doi: 10.1101/pdb.prot5156
31. Naik SH, Perie L, Swart E, Gerlach C, van Rooij N, de Boer RJ, et al. Diverse and heritable lineage imprinting of early haematopoietic progenitors. *Nature* (2013) 496(7444):229–32. doi: 10.1038/nature12013
32. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* (2010) 26(1):139–40. doi: 10.1093/bioinformatics/btp616
33. Schmitt TM, Zuniga-Pflucker JC. Induction of T cell development from hematopoietic progenitor cells by delta-like-1 *in vitro*. *Immunity* (2002) 17(6):749–56. doi: 10.1016/S1074-7613(02)00474-0
34. Sagar, Pokrovskii M, Herman JS, Naik S, Sock E, Zeis P, et al. Deciphering the regulatory landscape of fetal and adult $\gamma\delta$ T-cell development at single-cell resolution. *EMBO J* (2020) 39(13):e104159. doi: 10.15252/embj.2019104159
35. O'Brien RL, Born WK. Gammadelta T cell subsets: a link between TCR and function? *Semin Immunol* (2010) 22(4):193–8. doi: 10.1016/j.smim.2010.03.006
36. Scaramuzzino S, Potier D, Ordioni R, Grenot P, Payet-Bornet D, Luche H, et al. Single-cell transcriptomics uncovers an instructive T-cell receptor role in adult $\gamma\delta$ T-cell lineage commitment. *EMBO J* (2022) 41(5):e110023. doi: 10.15252/embj.2021110023
37. Hayes SM, Li L, Love PE. TCR signal strength influences $\alpha\beta/\gamma\delta$ lineage fate. *Immunity* (2005) 22(5):583–93. doi: 10.1016/j.immuni.2005.03.014
38. Lauritsen JP, Wong GW, Lee SY, Lefebvre JM, Ciofani M, Rhodes M, et al. Marked induction of the helix-loop-helix protein Id3 promotes the gammadelta T cell fate and renders their functional maturation notch independent. *Immunity* (2009) 31(4):565–75. doi: 10.1016/j.immuni.2009.07.010
39. Munoz-Ruiz M, Sumaria N, Pennington DJ, Silva-Santos B. Thymic determinants of $\gamma\delta$ T cell differentiation. *Trends Immunol* (2017) 38(5):336–44. doi: 10.1016/j.it.2017.01.007
40. Ribot JC, deBarros A, Pang DJ, Neves JF, Peperzak V, Roberts SJ, et al. CD27 is a thymic determinant of the balance between interferon- γ - and interleukin 17-producing $\gamma\delta$ T cell subsets. *Nat Immunol* (2009) 10(4):427–36. doi: 10.1038/ni.1717
41. Malhotra N, Narayan K, Cho OH, Sylvia KE, Yin C, Melichar H, et al. A network of high-mobility group box transcription factors programs innate interleukin-17 production. *Immunity* (2013) 38(4):681–93. doi: 10.1016/j.immuni.2013.01.010
42. Zuberbuehler MK, Parker ME, Wheaton JD, Espinosa JR, Salzler HR, Park E, et al. The transcription factor c-maf is essential for the commitment of IL-17-producing $\gamma\delta$ T cells. *Nat Immunol* (2019) 20(1):73–85. doi: 10.1038/s41590-018-0274-0
43. Gabrysova L, Alvarez-Martinez M, Luisier R, Cox LS, Sodenkamp J, Hosking C, et al. C-maf controls immune responses by regulating disease-specific gene networks and repressing IL-2 in CD4⁺ T cells. *Nat Immunol* (2018) 19(5):497–507. doi: 10.1038/s41590-018-0083-5
44. Iwata S, Mikami Y, Sun HW, Brooks SR, Jankovic D, Hirahara K, et al. The transcription factor T-bet limits amplification of type I IFN transcriptome and circuitry in T helper 1 cells. *Immunity* (2017) 46(6):983–91.e4. doi: 10.1016/j.immuni.2017.05.005
45. Yagi R, Junttila IS, Wei G, Urban JF Jr., Zhao K, Paul WE, et al. The transcription factor GATA3 actively represses RUNX3 protein-regulated production of interferon- γ . *Immunity* (2010) 32(4):507–17. doi: 10.1016/j.immuni.2010.04.004
46. Shibata K, Yamada H, Nakamura M, Hatano S, Katsuragi Y, Kominami R, et al. IFN- γ -producing and IL-17-producing gammadelta T cells differentiate at distinct developmental stages in murine fetal thymus. *J Immunol*. (2014) 192(5):2210–8.