

**ARTICLE**    **OPEN**


# Human MAIT cells undergo clonal selection and expansion during thymic maturation and aging

 Myeong-seok Lee<sup>1,5</sup>, Suyeong Park<sup>1,5</sup>, Jung-Hwan Choi<sup>1</sup>, Seon Yong Bae<sup>2</sup>, Han Suk Ryu<sup>3</sup>, Min-Sung Kim<sup>4</sup>, Jae Gun Kwak<sup>2</sup> and You Jeong Lee<sup>1</sup>

© The Author(s) 2025

Mucosal-associated invariant T (MAIT) cells harbor conserved T cell receptors (TCRs) recognizing riboflavin metabolites, yet exhibit substantial diversity similar to conventional memory T cells. However, the mechanisms shaping this diversity related to their thymic ontogeny remain unclear. Here we analyze 37 samples of human thymic MAIT cells across ages and compare them with other unconventional T cells, such as iNKT and  $\gamma\delta$  T cells. We find that CD27 and CD161 serve as common markers distinguishing the maturation stages of unconventional T cells such as MAIT, iNKT and  $V\gamma 9^+ V\delta 2^+$   $\gamma\delta$  T cells. Notably, CD161<sup>+</sup> mature MAIT cells clonally expand proportionally to aging with the upregulation of genes associated with tissue residency. MAIT cell diversity is initially determined by diverse CDR3 $\beta$  sequences, which become reduced upon maturation. Furthermore, 25% of MAIT cells express polyclonal dual TCR $\alpha$  transcripts, suggesting they arise from double-positive thymocytes with random TCR $\alpha$  rearrangement. Collectively, these findings show that thymic MAIT cells undergo dynamic regulation of repertoire selection, similar to conventional T cells.

*Experimental & Molecular Medicine* (2025) 57:1789–1801; <https://doi.org/10.1038/s12276-025-01509-x>

## INTRODUCTION

Human mucosal-associated invariant T (MAIT) cells have a canonical TRAV1-2 (T cell receptor (TCR) Va7.2) chain paired with oligoclonal TRAJ33/20/12 and TRBV20/6 chains, recognizing bacteria-derived vitamin B2 metabolites in the context of MR1<sup>1,2</sup>. In mice, MAIT cells, similar to invariant natural killer T (iNKT) cells, and some  $\gamma\delta$  T cells develop in the thymus as memory cells expressing the transcription factor promyelocytic zinc finger (PLZF) protein<sup>2,3</sup>. In the human thymus, MAIT cells undergo sequential differentiation from CD27<sup>+</sup>CD161<sup>+</sup> stage 1 to CD27<sup>+</sup>CD161<sup>+</sup> stage 2 and finally to CD161<sup>+</sup> stage 3 (ref. 4). Stage 3 mature MAIT cells co-express PLZF, TBET and ROR $\gamma$ t and secrete IFN $\gamma$ , TNF $\alpha$  and IL-17, indicating that serial differentiation accompanies the acquisition of innate phenotype<sup>4</sup>. Recent reports have shown that the clonal diversity of MAIT cells is comparable to that of conventional memory T cells<sup>5</sup>, and polyclonal MAIT cells are present in the periphery<sup>6,7</sup>. However, the mechanisms that shape receptor diversity during thymic development, where they acquire innate phenotypes, remain poorly understood. Furthermore, it is unclear whether MAIT cells undergo clonal selection or expansion in a manner similar to that of conventional memory T cells.

Recent studies have analyzed the transcriptional profiles of human MAIT cells during thymic development, revealing that MAIT and iNKT cells share conserved transcriptional programs across species<sup>5,8–10</sup>. iNKT cells have TRAV10 (TCR Va24) paired with TRAJ18 and TRBV25-1 and are restricted by CD1d presenting lipid

antigens<sup>11–13</sup>. Previously, we and others showed innate T cells in mice share common lineage differentiation programs generating type 1, type 2 and type 17 effector lineages<sup>12,14,15</sup>. In the thymus, human MAIT, iNKT and  $V\gamma 9^+ V\delta 2^+$   $\gamma\delta$  T cells shared distinctive transcriptional programs distinguished from conventional T cells<sup>10</sup>. These studies, however, used human thymic tissues obtained from donors under 2 years old, and it is not well understood whether the developmental programs of MAIT and iNKT cells are stable for a lifetime or whether there are long-term tissue-resident populations formed similarly to mice natural killer T (NKT) cells<sup>16</sup>.

We address the above issues to fill the knowledge gap in the field. For this, we analyzed the thymic ontogeny of human counterparts of mouse innate T cells, focusing on repertoire selection of MAIT and iNKT cells using 37 postnatal thymi ranging from 2 weeks to 56 years of age. We sorted MAIT, NKT and  $\gamma\delta$  T cells from four samples representing different age groups and performed single-cell RNA sequencing (scRNA-seq) combined with TCR repertoire analysis. This approach enabled us to uncover their ontogeny, clonal selection and maturation pathways across age groups. We found that CD27 and CD161 are effective markers for delineating the developmental stages of MAIT and iNKT cells. While mature and immature MAIT and iNKT cells broadly shared transcriptional similarities, they exhibited distinct maturation patterns. By contrast, most of the classical  $\gamma\delta$  T cells were markedly different in their transcriptional profiles. Intriguingly, only CD8<sup>+</sup> MAIT cells efficiently matured into stage 3 cells, while iNKT cells predominantly remained immature at stage 2.

<sup>1</sup>Research Institute of Pharmaceutical Sciences, College of Pharmacy, Seoul National University, Seoul, Republic of Korea. <sup>2</sup>Department of Thoracic and Cardiovascular Surgery, Seoul National University Hospital, College of Medicine, Seoul National University, Seoul, Republic of Korea. <sup>3</sup>Department of Pathology, Seoul National University Hospital, College of Medicine, Seoul National University, Seoul, Republic of Korea. <sup>4</sup>Department of Life Sciences, Pohang University of Science and Technology, Pohang, Republic of Korea. <sup>5</sup>These authors contributed equally: Myeong-seok Lee, Suyeong Park. ✉email: leri11@snu.ac.kr; youjeonglee@snu.ac.kr

Received: 17 January 2025 Revised: 23 April 2025 Accepted: 7 May 2025  
 Published online: 8 August 2025

In the TCR repertoire analysis, thymic MAIT and iNKT cells utilized canonical CDR3 $\alpha$  sequences paired with diverse CDR3 $\beta$  sequences, generating diverse clonotypes from an immature stage. Notably, mature MAIT cells in the adult human thymus underwent clonal expansion and upregulated genes associated with tissue residency. During the MAIT cell maturation, we also observed a reduction in TCR V $\beta$  diversity, indicating that the selection of semi-invariant TCRs is required for their final maturation or long-term residence. Moreover, approximately 25% of human MAIT and iNKT cells possessed an additional polyclonal TCR $\alpha$  transcript, similar to conventional T cells, suggesting they arise from random TCR gene rearrangement in double-positive (DP) thymocytes. However, polyclonal TRAV1-2-negative noncanonical MAIT cells were not detected in the thymus when we examined 27 clonotypes. Collectively, our findings show that the diverse repertoire of human MAIT cells is established in the thymus by pairing public TCR $\alpha$  chain with private TCR $\beta$  chains, and their maturation accompanies clonal selection and expansion similar to conventional memory T cells.

## MATERIALS AND METHODS

### Study design

In this study, we investigated the maturation pathways of innate T cells across different ages using a cohort of 37 postnatal thymi. MAIT, iNKT and  $\gamma\delta$  T cells were obtained via magnetic-based and flow-based cell sorting from thymocytes of four donors spanning various age groups. Their gene expression profiles and clonotypes were analyzed using scRNA-seq along with TCR repertoire analysis. We cloned 27 random polyclonal TCRs into a lentiviral vector to test the presence of polyclonal MAIT cells.

### Human samples

Human thymi and blood samples were obtained from patients who underwent cardiac surgery. Thymectomy was performed only in cases where thymic tissue interrupted operative fields, and its removal was kept to a minimum to improve exposure during corrective cardiovascular surgery. As previously described<sup>17,18</sup>, human thymocytes and peripheral blood mononuclear cells (PBMCs) were isolated by density gradient centrifugation using Ficoll-Paque (GE Healthcare Life Science). This study was reviewed and approved by the Seoul National University Hospital (Seoul, Republic of Korea; H-2206-221-1337) and conducted according to the principles of the Declaration of Helsinki. Informed consent was obtained from all study participants.

### Flow cytometry

Biotinylated PBS57-loaded CD1d monomers and 5-OP-RU-loaded MR1 monomers were obtained from the tetramer facility of the US National Institutes of Health. Tetramer synthesis was described previously<sup>19</sup>. Single-cell suspensions were surface stained, fixed and permeabilized for intracellular staining with an eBioscience Foxp3 staining buffer set as described<sup>20,21</sup>. Cells were analyzed on LSR Fortessa X-20 (BD Biosciences), and data were processed with FlowJo software (TreeStar).

### Cell sorting for scRNA-seq

As previously described<sup>22</sup>, single-cell suspensions of human thymocytes were stained with BV421-conjugated anti-human TCR $\gamma\delta$  or PE-conjugated CD1d, and APC-conjugated MR1 tetramers after blocking with anti-CD36 antibody<sup>23</sup>, then enriched with anti-PE and anti-APC microbeads (Miltenyi) according to the manufacturer's instructions. After sorting, iNKT, MAIT and  $\gamma\delta$  T cells were mixed and analyzed for single-cell sequencing.

### TCR transfection assay

Based on the previous protocol<sup>24</sup>, synthetic TCR $\alpha$  and TCR $\beta$  sequences from MR1 tetramer-bound sorted single T cells were cloned into a pLV-EF1 $\alpha$ -Puro vector (Thermo Fisher Scientific). Human embryonic kidney (HEK293T) cells were cultured overnight on a 24-well plate containing 300  $\mu$ l of Dulbecco's modified Eagle medium supplemented with 10% fetal bovine serum and penicillin–streptomycin at 37°C, and they were subsequently transfected with the pLV-TCR plasmid using FuGENE HD transfection reagent (Promega).

Transfected HEK293T cells were analyzed for tetramer binding by flow cytometry 48 h after transfection.

### 10x Genomics library generation and sequencing

Sequencing libraries were generated using 10x Genomics Chromium Single Cell 5' Reagent Kits (v2 Chemistry Dual Index, following the manufacturer's instructions). Cells were loaded onto the Chromium Controller (10x Genomics) at a concentration of approximately  $0.5 \times 10^6$  cells/ml, with 8000–10,000 cells targeted for sequencing. Library quality and concentration were assessed using a TapeStation (Agilent) and Countess II FL Automated Cell Counter (Thermo Fisher Scientific), respectively. Library generation and sequencing on an Illumina Nova-Seq6000 were performed at Geninus. Sequencing depths were 29,860–32,634 reads per cell for scRNA-seq.

### 10x Genomics raw data processing

FASTQ files were produced by converting BCL files with Illumina bcl2fastq. The gene expression and TCR data from these FASTQ files were analyzed using Cell Ranger (v.6.1.2) for both gene expression and TCR data. Specifically, for TCR analysis, the 'filtered\_contig\_annotations.csv' file was refined to include only high-confidence, complete and productive contigs associated with either TCR $\alpha$  or TCR $\beta$  chains.

### Quality control

Quality control was performed separately for cells from each channel of the Chromium Controller. Filtered feature–barcode matrices from Cell Ranger count were imported into R using Seurat<sup>25</sup> (v.4.3.0.1). Cells with low gene counts and/or a high percentage of mitochondrial reads were removed. Furthermore, cells identified as potential doublets were excluded from the dataset for subsequent analysis. The probability of being a doublet was calculated using the pDoublet values, which were computed with the Scrublet algorithm<sup>26</sup> (v.0.1, python3). The cutoff values for pDoublet, nFeature\_RNA and percent.mito are described in Supplementary Fig. 3. Cells expressing nontarget cell markers were excluded from further analysis.  $\gamma\delta$  T cells were validated with TRDC (T cell receptor delta constant) expression greater than 0.5, and the absence of a TCR $\beta$  chain sequence NKT and MAITs were distinguished on the basis of TCR $\alpha\beta$  chain information. Cells with no TRAC (T cell receptor alpha constant) and/or TRBC (T cell receptor beta constant) were excluded.

### Normalization, integration, dimensionality reduction and clustering

For the combined analysis of four different human thymus samples, data from each donor were independently normalized using the NormalizeData function in Seurat. Highly variable genes were defined as the top 2000 genes with the highest standardized variance, utilizing the FindVariableFeatures function in Seurat. Furthermore, TCR and Ig genes were removed from highly variable genes and subsequently normalized using the ScaleData function in Seurat to prevent any impact on subsequent clustering analyses pertaining to TCR or Ig gene utilization. Principal component analysis was used for dimensionality reduction, and the donor-specific batch correction was executed using Harmony<sup>27</sup> (v0.1.1) considering 30 input principal components. ElbowPlots were consulted to determine the optimal number of principal components for UMAP construction and clustering. Cell clusters were delineated using Seurat's graph-based clustering method.

### Cell type annotation

Only the subset of cells containing TCR sequences in scRNA-seq data was analyzed using the scRepertoire<sup>28</sup> (v.1.7.0). Within this subset, cells were classified as either MAIT or NKT cells based on the usage of canonical Va chain genes, specifically TRAV1-2 or TRAV10, as determined by the TCR sequence data. Cells that did not express these genes were categorized as polyclonal cells. Three distinct maturation stages were determined on the basis of IL18R1, KLRB1 and CD27 expression using Seurat clustering. The distribution of cell types corresponding to each stage was visualized using the dittoBarPlot function of dittoSeq<sup>29</sup> (v.1.6.0).

### Heterogeneity analysis (GEX)

The analysis was focused exclusively on the subset of cells containing TCR sequences in the scRNA-seq data, and the analysis was performed using the CombineTCR function from scRepertoire. Within the functionality of

the CombineTCR function, we removed cells with missing TCR information (NA) and filtered multiplets, selecting the row with the highest read counts for multi TCR $\alpha$  or TCR $\beta$  as the final value.

### Differentially expressed gene (DEG) analysis

To assess the overall expression patterns of the cells, we used the `plot_density` function of `Nebulosa`<sup>30</sup> (v.1.4.0). DEGs that exhibit differential expression between the comparison groups were identified using `Seurat's FindMarkers` function. Subsequently, these DEGs were visualized utilizing the `EnhancedVolcano` function from `EnhancedVolcano` (v.1.13.2). The cutoff criteria for this analysis included a *P*-value threshold of  $10^{-5}$  and a fold change threshold of 1. For the heat map analysis, DEGs were grouped on the basis of their shared functionality. Enriched gene lists were obtained from each group under comparison. Subsequently, heat map analysis was conducted using the `Seurat's DoHeatmap` function. The signature score analysis was conducted using the `AddmoduleScore_UCell` function of the `UCell`<sup>31</sup> (v.1.3.1).

### Pathway analysis

Based on the list of DEGs, pathway analyses were conducted using Hallmark pathways from `msigdb` (v.7.5.1) and Gene Ontology (GO): Biological Process 2023 analysis from `EnrichR`<sup>32</sup> (v.3.2). In the Hallmark pathway analysis, an initial Wilcoxon rank-sum test was performed on all genes using the `wilcox` function of `presto`<sup>33</sup> (v.1.0.0). Genes were ranked in descending order of their AUC score concerning the target group. Subsequently, the `fgsea` function from `fgsea`<sup>34</sup> (v.1.20.0) was used to compute the normalized enrichment score and adjusted *P* value for the genes comprising each pathway based on their ranks. Hallmark pathways with adjusted *P* values less than 0.05 were presented. For the GO: Biological Process 2023 analysis, enriched pathways between groups were compared on the basis of the sorted rank values. This was achieved using `EnrichR's DEenrichRPlot` function, which utilized the Wilcoxon rank-sum test. A log fold change threshold of 0.4 and a *P*-value threshold of 0.05 were used to identify the top ten GO: BP 2023 pathways based on `log10pval`. For biologically relevant pathways, the `plotEnrichment` function of `fgsea` was used to create gene set enrichment analysis plots for visualization.

### Pseudotime analysis

To utilize `Monocle3`<sup>35,36</sup> (v.1.3.1), the initial step involved converting the `Seurat` object into a cell dataset using the `as_cell_data_set` function from `SeuratWrappers` (v.0.3.1). Then, the `LearnGraph` function was used to ascertain the graph structure within the reduced dimension space. The node representing the immature stage was subsequently set as the root using the `order_cells` function. Lastly, pseudotime values were depicted on UMAP embeddings using the `plot_cells` function.

### TCR repertoire analysis

All TCR analyses were conducted on the subset of cells with available gene expression information.

### TCR chain gene usage

To analyze TCR gene usage, the `scRepertoire's Vizgenes` function was used to create heat maps that visualize TCR gene usage across patients. The data were scaled on the basis of the highest frequency among individual patients and cells. In addition, the `ChorDiagram` function from `Circlize`<sup>37</sup> (v.0.4.15) was used to visualize the composition of TRAJ and TRBV by cell type and their interconnections, providing insights into the relationships between different cell types in the TCR repertoire.

### CDR3 amino acid sequence

The length distribution of CDR3 amino acids in TCR $\alpha$  and TCR $\beta$  was assessed using `scRepertoire's lengthContig` function. Subsequently, sequence logos were visualized for amino acid sequences of a specific length with a single peak for each patient using `Weblogo`<sup>38</sup> (v.2.8.2).

### Clonality analysis

TCR gene usage and matching CDR3 nucleotide sequences were designated as the same clone. Clones were categorized into groups based on their frequency of occurrence: single, small, medium, large and hyperexpanded, representing clone repeat counts of 1, 2–5, 6–20, 21–100

and 101–500, respectively. The Shannon index was computed using `scRepertoire's clonalDiversity` function.

### Plots

Plots were generated with the `ggplot2` (v.3.4.2). Fluorescence-activated cell sorting (FACS) plots were produced using `FlowJo` (v.10.9.0 by BD Biosciences). The other plots were constructed using `GraphPad Prism` (v.10.0.03 by GraphPad Software, LLC).

### Statistical analyses

Statistical analysis was performed using `Microsoft Excel` and `GraphPad Prism` version 10.1.1. Values reported in the figures are expressed as the standard error of the mean (s.e.m.), unless otherwise indicated. For normally distributed datasets, we used the paired-sample *t*-test between two comparison groups and one-way analysis of variance among three or more groups. *P* values  $>0.05$  were considered not significant (ns), and *P* values  $<0.05$  were considered significant. \**P*  $<0.05$ , \*\**P*  $<0.01$ , \*\*\**P*  $<0.001$ .

## RESULTS

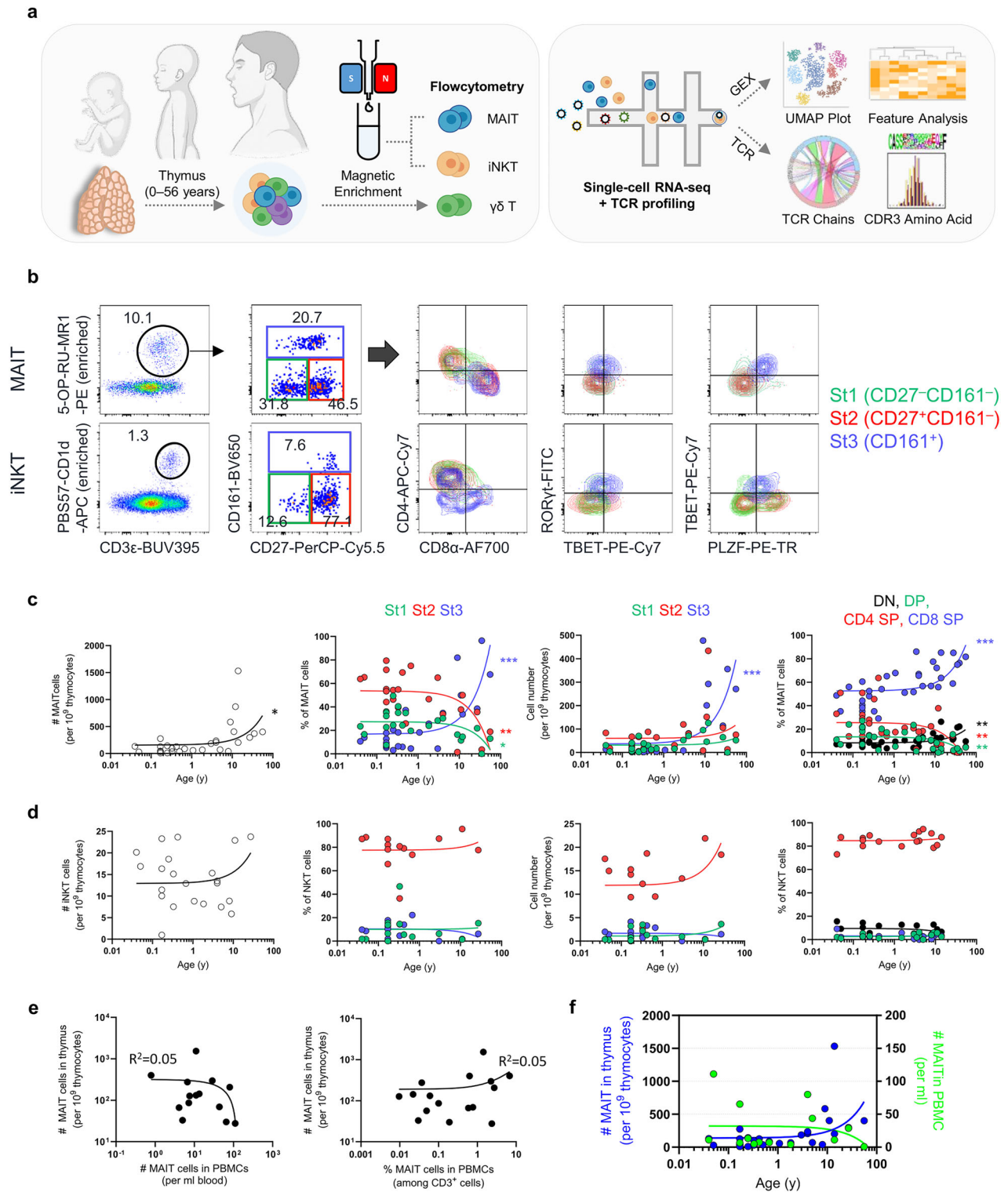
### Mature MAIT cells expand in the thymus with age

To gain comprehensive insights into the thymic ontogeny of human MAIT cells, we analyzed 37 postnatal human thymi and simultaneously analyzed iNKT and  $\gamma\delta$  T cells for comparison (Fig. 1a). We enriched MAIT and iNKT cells using MR1 and CD1d tetramers or stained whole thymocytes with anti-TCR $\gamma\delta$  antibodies (Supplementary Fig. 1a,b). Analogous to previous studies showing that CD27 and CD161 distinguish stages 1, 2 and 3 MAIT cells<sup>4</sup>, we found that iNKT cells follow a similar maturation pattern to MAIT cells in the thymus (Fig. 1b) and periphery (Supplementary Fig. 1c). Based on this, we designate stage 1 iNKT cells as CD27<sup>+</sup>CD161<sup>+</sup>, stage 2 cells as CD27<sup>+</sup>CD161<sup>+</sup> and stage 3 cells as CD161<sup>+</sup> (Fig. 1b, middle).

Unlike MAIT cells, iNKT cells upregulated PLZF in stages 1 and 2, and subsequently co-expressed PLZF, TBET, and ROR $\gamma$ t, similar to MAIT cells in stage 3 (Fig. 1b). CD4<sup>+</sup> iNKT cells have previously been characterized as immature populations in both the thymus and periphery<sup>39</sup>. The CD4 and CD161 combination was used to distinguish stage 1 (CD4<sup>+</sup>CD161<sup>+</sup>), stage 2 (CD4<sup>+</sup>CD161<sup>+</sup>) and stage 3 (CD4<sup>+</sup>CD161<sup>+</sup>)  $\gamma\delta$  T cells<sup>40</sup>. Therefore, it is possible that iNKT cells follow a developmental scheme similar to that of  $\gamma\delta$  T cells. To check this, we compared CD27–CD161 and CD4–CD161 classification frameworks in iNKT and MAIT cells (Supplementary Fig. 2a). Under the CD4–CD161 scheme, however, stage 1 iNKT cells upregulated PLZF, followed by downregulation at stage 2 and subsequent reexpression at stage 3, which is unlikely to reflect a physiologically relevant maturation trajectory. Given the coinciding expression of CD4 and CD27 in iNKT cells (Supplementary Fig. 2a), CD4<sup>+</sup>CD161<sup>+</sup> iNKT cells are probably indicative of stage 2 cells.

Interestingly, the proportion of MAIT cells, but not iNKT cells, in the human thymus gradually increased with age, mainly due to the expansion of CD8 single-positive (SP) stage 3 cells (Fig. 1c). However, the frequency of stage 1 cells among total thymocytes remained stable across ages, suggesting that immature MAIT cells are continuously generated (Fig. 1c). By contrast, iNKT cells were arrested at stage 2, and the proportions of cells in each stage remained stable up to the mid-50s (Fig. 1d). The expansion of mature MAIT cells in the thymus was unrelated to the periphery, suggesting that thymic expansion is more likely to be driven by *novo* changes rather than systemic effects (Fig. 1e, f).

$\gamma\delta$  T cells are distinguished by their expression of TCR chains, and fetal-derived  $\gamma\delta$  T cells represent the CD161<sup>+</sup> PLZF<sup>+</sup> TBET<sup>+</sup> innate subset, which is infrequently detected in postnatal thymus (Supplementary Fig. 3a,b). These cells were stable in the postnatal thymi, similar to iNKT cells (Supplementary Fig. 3c). Most of the non- $\gamma\delta$  T cells were CD161<sup>+</sup> and did not express



**Fig. 1** Mature MAIT cells expand in the thymus with age. **a** The schematic figure shows experimental procedures for isolating MAIT, iNKT and  $\gamma\delta$  T cells for flow cytometry ( $n=37$ ), and scRNA-seq combined with TCR sequencing from the thymus ( $n=4$ ). **b** Dot plots show representative flow cytometry analysis of human thymus, enriched with 5-OP-RU-loaded MR1 and PBS57-loaded CD1d tetramers. **c, d** The graphs depict the Pearson correlation analyses of the frequencies and numbers of MAIT (**c**) and iNKT (**d**) cells across the donors' ages. **e** The graphs depict the correlations between the frequencies of MAIT cells in the thymus and their matched PBMC samples. **f** The graph shows the absolute numbers of MAIT cells in the thymus and PBMC in various age groups. The lines in the graphs represents the simple linear regression. Numbers indicate frequencies of cells in adjacent gates. St stage, SP single positive, DP double positive. \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ .

PLZF in the thymus (Supplementary Fig. 3), indicating that these are not innate T cells. Later, we referred to these cells as classic  $\gamma\delta$  T cells. Collectively, these features indicate that MAIT cell maturation is partly shared with, but distinct from, iNKT and  $\gamma\delta$  T cell maturation processes. Moreover, mature MAIT cells expand in the thymus in an age-dependent manner.

### Age-dependent kinetics of MAIT cell differentiation at the single-cell level

To gain further insights into the age-dependent expansion of human MAIT cells, we simultaneously sorted MAIT, iNKT and  $\gamma\delta$  T cells and performed scRNA-seq combined with TCR repertoire analysis, as outlined in Fig. 1a and Supplementary Fig. 1a. To encompass various age groups, we collected cells from four donors—a 5-month-old female, a 5-year-old male, an 8-year-old female and a 28-year-old male—yielding a total of 29,117 cells (21,153  $\alpha\beta$  T cells and 8024  $\gamma\delta$  T cells). We obtained  $\alpha\beta$  TCR information from 14,723 cells (a 69.6% detection rate) (Fig. 2a), following quality control (Supplementary Fig. 4a–e). Notably, 2383 cells (16.2%) displayed dual TCR $\alpha$  transcripts. We annotated cell types on the basis of the transcript with higher expression for these cells and identified MAIT ( $n = 11,109$ ), iNKT ( $n = 1885$ ) and polyclonal  $\alpha\beta$  T cells ( $n = 1729$ ) (Supplementary Fig. 4f). For uniform manifold approximation and projection (UMAP) analysis, TCR genes were excluded from the DEG list to avoid TCR-dependent clustering. Seven distinct clusters (C0 to C6) were obtained (Fig. 2a). Supplementary Fig. 5 shows the distribution of cells in each donor, and sequence information is provided in Supplementary Table 1.

Previous studies have shown that MAIT and iNKT cells share similar transcriptional profiles in the thymus of individuals younger than 2 years old<sup>9,10</sup>. Our analysis confirmed that these cell types can be distinguished by their TCRs but not by their clustering in aged samples across ages, despite the expansion of mature cells observed in adult samples (Fig. 2a–d). In the UMAP, clusters C1 and C2 represented immature and mature stages of both MAIT and iNKT cells, distinguished by the simultaneous expression of PLZF (encoded by *ZBTB16*), TBET (encoded by *TBX21*), ROR $\gamma$ t (encoded by *RORC*), IL18R1 (encoded by *IL18R1*) and CD161 (encoded by *KLRB1*) (Fig. 2b). Because we used postnatal thymic, we obtained only about 50  $V\gamma 9^+V\delta 2^+$   $\gamma\delta$  cells (data not shown), and the most of classical  $\gamma\delta$  T cells were not positioned together with MAIT or iNKT cells in UMAP. Consistent with the flow cytometric analysis (Fig. 1b–d), the proportions of mature CD8 SP MAIT cells expanded with aging (Fig. 2d). We generated a dot plot to display key markers of each cluster (Fig. 2e), along with a dot plot depicting the top five differentially upregulated pathways in each Louvain cluster (Fig. 2f). These dot plots further support that each cluster represents unique cell types, and classical  $\gamma\delta$  T cells are distinct from MAIT and iNKT cells. Cluster C0 represents postselection DP MAIT or iNKT cells, and its signature genes overlap with those of C6, which highly expresses Ki-67 (Fig. 2e). Therefore, we designated C6 as stage 0 proliferating immature  $\gamma\delta$  T cells. Overall, these findings indicate that, while the developmental pathways of MAIT and iNKT cells are similar, only mature MAIT cells expand with age, maintaining their innate phenotype.

### MAIT cells are distinguished from iNKT cells for their detailed maturation processes

We separated MAIT and iNKT cells on the basis of TCR annotation and performed UMAP analysis (Fig. 3a–c). Consistent with flow cytometric analysis, CD27 and CD161 differentiated stage 1, 2 and 3 cells in the UMAP, and stage 0 postselection DP cells corresponded to C0 in Fig. 2 (Fig. 3a, c). Pseudotime analysis supported serial differentiation of MAIT and iNKT cells from stage 0 to stage 3 (Fig. 3b). In the density plots, we compared the expression patterns of various markers between MAIT and iNKT cells, revealing distinct developmental patterns between

the two cell types (Fig. 3c). Specifically, iNKT cells co-expressed CD4 and CD27 in the immature cluster and immature iNKT cells prematurely upregulated PLZF before maturation, unlike MAIT cells. These features are consistent with FACS analyses in Fig. 1b and Supplementary Fig. 2 and support the idea that CD4 is not a marker of the most immature iNKT cells.

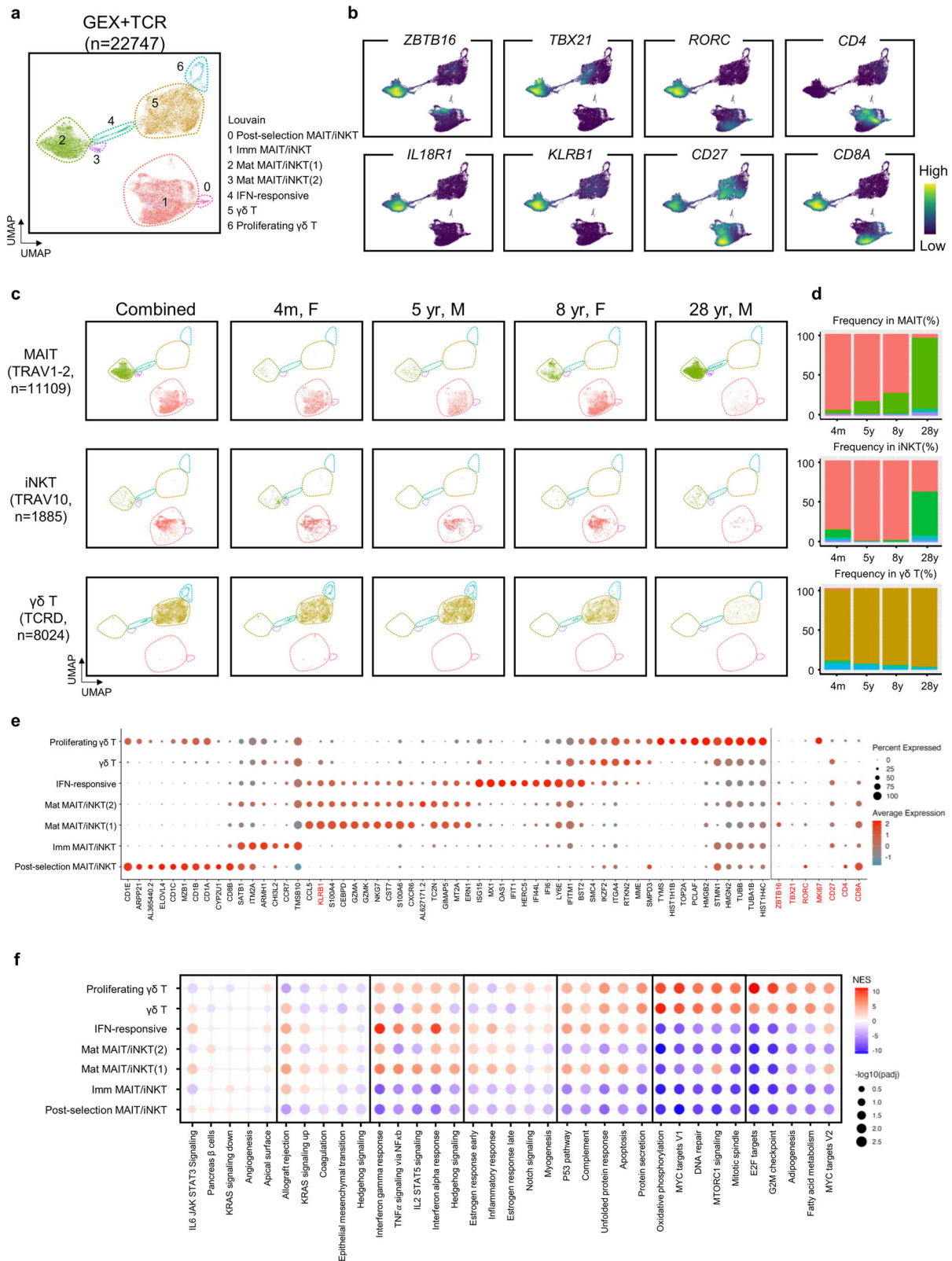
A heat map for MAIT and iNKT cells showed markers associated with cytokine, chemokine, cytotoxicity, chemokine and TCR signaling differentially expressed between each stage (Supplementary Fig. 6a). Mature MAIT and iNKT cells had higher expressions of genes associated with cytokine, chemokines and cytotoxic molecules, whereas immature cells upregulated genes associated with TCR signaling (Supplementary Fig. 6a). We further validated the expressions of HLA-ABC, GZMB, GZMK, CCR7 and TOX by flow cytometry (Supplementary Fig. 6b).

For more detailed analysis, we generated UMAPs for each cell type (Fig. 3d). In MAIT cells, we identified eight clusters, which could be grouped into two main categories: one composed of stage 0, 1 and 2 cells (C1, C2, C3 and C7), and the other composed of stage 3 cells (C0, C3, C4 and C6) expressing *PLZF*, *TBX21*, *RORC*, *IL18R1*, *KLRB1*, *CD8*, *PRF1* and *CD27* (Fig. 3d, e, top). This pattern was clearly illustrated in a heat map (Supplementary Fig. 7, left) and a volcano plot, which showed that stage 1 and 2 MAIT cells had only a few DEGs (Supplementary Fig. 8a, top). For a more robust analysis, we integrated two public datasets (GSE239558 and GSE189485), which contained respective 5994 and 1,1316 human MAIT cells in the thymus<sup>8,9</sup> (Supplementary Fig. 9). The expression patterns of key markers in immature and mature cells, along with pseudotime analysis, confirmed consistent maturation trajectories across all three datasets (Supplementary Fig. 9).

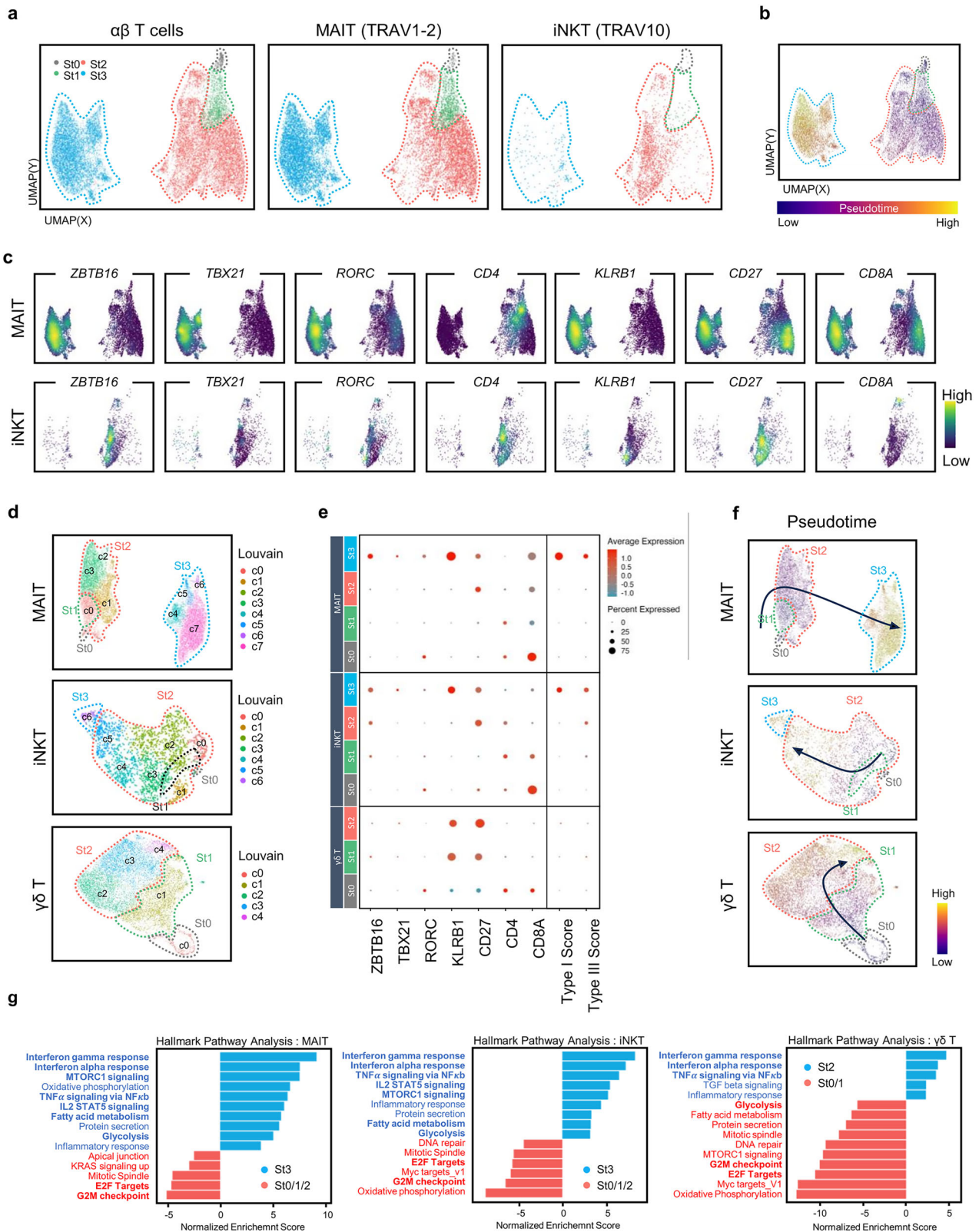
In iNKT cells, we identified seven clusters, in which CD161<sup>+</sup> mature cells (C6) were adjacent to clusters containing immature cells expressing *CD4* and *CD27* (Fig. 3d, e, middle), consistent with the flow cytometric analysis (Fig. 1b). Similar to MAIT cells, stage 1 and 2 iNKT cells had relatively few DEGs, while stage 3 cells were more distinct (Supplementary Fig. 8a, middle).

In classical  $\gamma\delta$  T cells, five clusters were identified (C0 to C4), corresponding to DP stage 0 (C0), CD27<sup>+</sup> stage 1 cells (C1) and CD161<sup>dim</sup> stage 2 cells (C2 to C4) (Fig. 3d, e, bottom). Pseudotime analysis supported the differentiation of C0 (DP, stage 0) cells into C1 (CD27<sup>+</sup>CD161<sup>+</sup> stage 1) cells and C2 to C4 (CD27<sup>+</sup>CD161<sup>dim</sup> stage 2) cells (Fig. 3d–f, bottom). Volcano plots also revealed a substantial number of DEGs between stages 0, 1 and 2 (Supplementary Fig. 8a, bottom). However, CD161 expression on  $\gamma\delta$  T cells was lower than MAIT and iNKT cells (Fig. 2b), consistent with the CD161<sup>dim</sup> pattern observed in flow cytometry of non- $V\gamma 9^+V\delta 2^+$   $\gamma\delta$  T cells (Supplementary Fig. 3a). These findings suggest that CD161 expression also marks the maturation of  $\gamma\delta$  T cells, albeit at lower levels than in MAIT or iNKT cells. As we used postnatal thymi, most of the  $\gamma\delta$  T cells were not  $V\gamma 9^+V\delta 2^+$ , which are fetal-derived cells sharing a transcriptional nature with MAIT and iNKT cells<sup>10</sup>. Non- $V\gamma 9^+V\delta 2^+$   $\gamma\delta$  T cells were distinct from MAIT or iNKT cells in the UMAP and did not fully acquire markers such as PLZF, TBET and ROR $\gamma$ t, suggesting they do not belong to the innate lineage (Fig. 2).

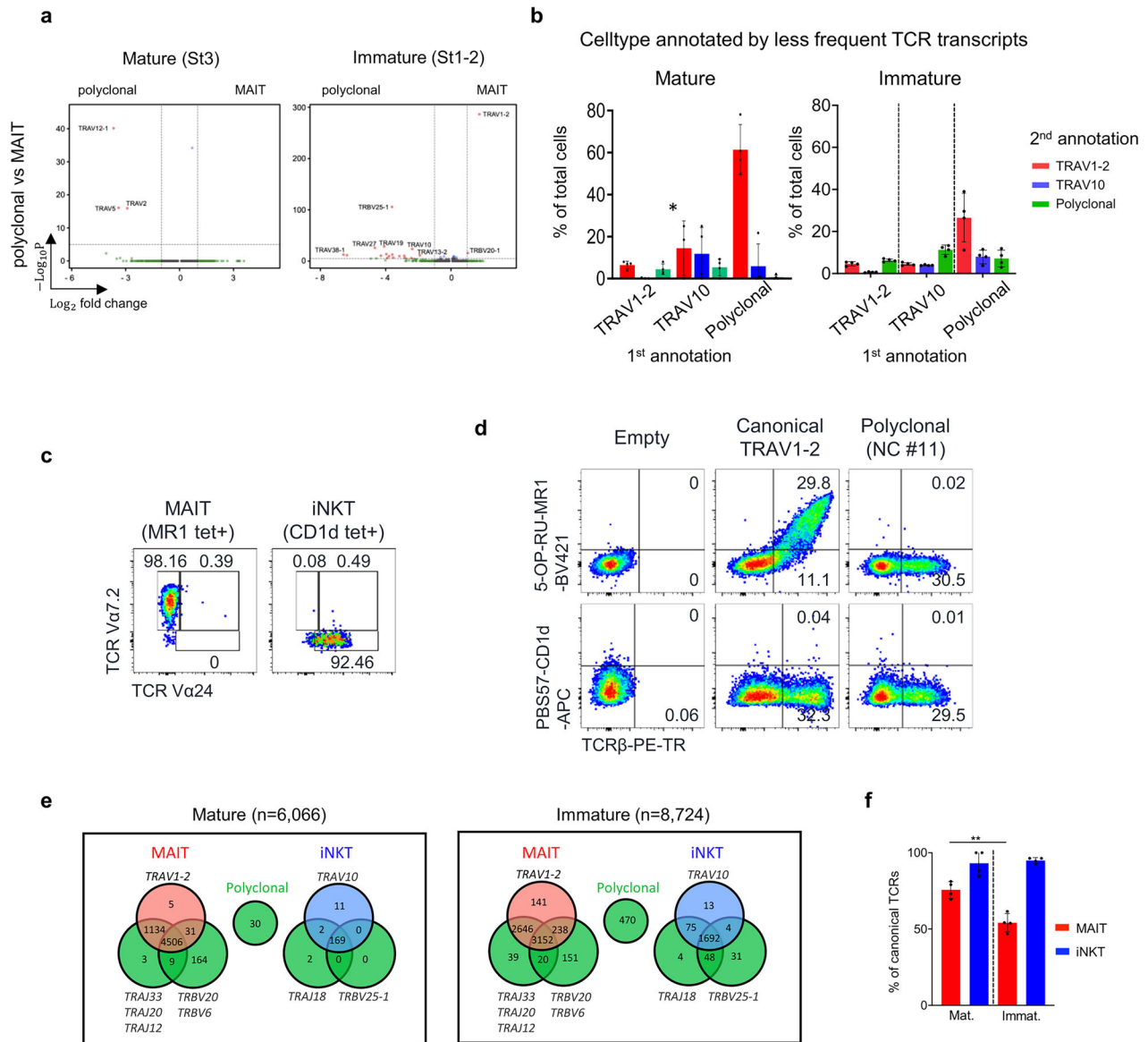
Pathway analyses revealed that MAIT and iNKT cells shared common maturation pathways (Fig. 3g and Supplementary Fig. 8b). In a Hallmark pathway analysis, mature MAIT and iNKT cells, compared with immature counterparts, upregulated pathways such as MTORC1 response, TNF $\alpha$  signaling, IL2-STAT5 signaling, fatty acid metabolism and glycolysis, while they downregulated E2F targets and G2M checkpoints (Fig. 3g). In GO pathway analysis, immature MAIT and iNKT cells were enriched in pathways related to antigen receptor-mediated signaling and T cell differentiation within the thymus, while mature cells upregulated pathways associated with cytokine-mediated signaling and cytotoxicity (Supplementary Fig. 8b). However, classical  $\gamma\delta$  T cells were not enriched with 'cytokine-mediated signaling in mature



**Fig. 2** Age-dependent kinetics of MAIT cell differentiation at the single-cell level. **a** The UMAP plot shows combined NKT, MAIT and  $\gamma\delta$  T cells with visualization based on the Louvain clusters. **b** The density plots illustrate the expression of indicated markers. **c** The UMAP plot in **a** was separated by donors and cell types. **d** Frequency plots depict the distribution of stages across individual donors and indicated cell types. **e, f** The dot plots show each Louvain cluster's top ten DEGs and manually selected markers (red) (**e**) and top five enriched pathways (**f**).



**Fig. 3 MAIT cells are distinguished from iNKT cells for their detailed maturation processes.** **a–c** Combined analysis of MAIT and iNKT cells. UMAP plots visualize the total, MAIT and iNKT cells (**a**). The density plots illustrate pseudotime progression (**b**) and the expression levels of indicated markers (**c**). **d** UMAP plots display unique cell clusters in each cell type. Stages were annotated according to CD27 and CD161 expression patterns. **e** Dot plots show the differential expression of the indicated markers and type-immunity-related markers for each Louvain cluster of each cell type. **f** Pseudotime analysis shows the maturation pathways of each cell type. **g** Bar graphs display the enriched pathways for mature cells (blue) compared with immature cells (red), with an adjusted  $P$  value of less than 0.05. St, stage.



**Fig. 4** MAIT cells with canonical TCRs are selected upon maturation. **a** Volcano plots display the DEGs between MAIT and polyclonal cells in mature and immature stages. **b** Graphs show the frequency of cell type shifts when annotated with less frequent TCR transcripts in mature and immature cells. \*One sample with four cells was excluded from the analysis. **c** Total thymocytes were enriched using 5-RO-RU-loaded MR1 or PBS57-loaded CD1d tetramer, followed by staining with anti-TCR V $\alpha$ 7.2 and V $\alpha$ 24 antibodies. **d** HEK293T cells were transfected with empty vectors, canonical MAIT TCRs (MAITc) or polyclonal TCRs (NC, noncanonical) and stained with MR1 and CD1d tetramers. Representative dot plots show a representative example of 27 NC samples. **e** Venn diagrams show the number of cells using the indicated combination of TCR chains. **f** The graph shows the frequency cells with all canonical combinations of TRAV, TRAJ and TRBV genes among total MAIT (TRAV1-2<sup>+</sup>) or iNKT (TRAV10<sup>+</sup>) cells. Numbers indicate the frequency of cells in each quadrant or adjacent gates. Error bars indicate s.e.m. \*\*\* $P < 0.01$ .

cells', and 'T cell differentiation in thymus' in immature cells, again suggesting they follow different maturation pathways (Supplementary Fig. 8b).

Collectively, these analyses indicate MAIT and iNKT cells share common maturation patterns, but are distinct in their detailed maturation processes.

#### MAIT cells with canonical TCRs are selected upon maturation

We obtained 1729 polyclonal  $\alpha\beta$  T cells from single-cell sequencing in both mature and immature cell clusters (Supplementary Fig. 5), displaying highly diverse usage of TCR genes (Supplementary Fig. 10). Because these cells had a minimal number of DEGs with canonical MAIT cells except TCR genes (Fig. 4a), we considered the dual TCR $\alpha$  transcripts, previously reported in 15% of human thymocytes<sup>41</sup>, and MR1 tetramer<sup>low</sup>

peripheral MAIT cells<sup>42</sup>. Supplementary Fig. 4g shows that an average of 50% of polyclonal T cells have second TCR $\alpha$  transcripts, and further analysis confirms that 65% of mature cells and 30% of immature cells have TRAV1-2 or TRAV10 as minor transcripts (Fig. 4b). Given that the detection rate for the TCR $\alpha$  gene in single-cell sequencing is about 70%, we estimate that 92% of mature cells and 42% of immature cells of polyclonal T cells have canonical MAIT or iNKT TCRs.

Notably, 10% of TRAV10<sup>+</sup> iNKT cells expressed TRAV1-2 as minor transcripts (Fig. 4b). However, flow cytometry analysis showed that very few cells co-expressed TCR V $\alpha$ 7.2 (TRAV1-2) and V $\alpha$ 24 (TRAV10) (Fig. 4c), suggesting that only one of the dual TCR transcripts is expressed on the cell surface. Therefore, there remains the possibility that some polyclonal TCRs represent polyclonal MAIT or iNKT cells in the thymus, as a previous report

identified TRAV36<sup>+</sup> noncanonical MAIT cells in the periphery<sup>7</sup>. To convincingly test this possibility, we cloned 28 TCRs, 27 random polyclonal ones and 1 canonical MAIT one, into a lentiviral vector (Supplementary Table 2) and expressed them in HEK293T cells to assess their MR1 and CD1d tetramer binding (Fig. 4d and Supplementary Fig. 11). We confirm that none of the 27 TCRs showed binding to the tetramers, demonstrating that these TCRs do not recognize MR1 or CD1d. These results demonstrate that polyclonal TCRs represent dual TCR $\alpha$  transcripts of either MAIT or iNKT cells. Consequently, we reannotated polyclonal  $\alpha\beta$  T cells as MAIT or iNKT cells if they contained additional TRAV1-2 or TRAV10 transcripts, respectively, as summarized in Fig. 4e.

To further investigate the TCR usage, we analyzed TCR Ja and V $\beta$  combinations among canonical TCR $\alpha$ -expressing MAIT and iNKT cells (Fig. 4f). We found that iNKT cells displayed a strict canonical pairing of TCR Va10, Ja18 and V $\beta$ 25-1, reflecting a highly conserved selection process. By contrast, MAIT cells exhibited developmental differences in TCR usage, with immature MAIT cells showing less frequent use of TCR V $\beta$ 6 and V $\beta$ 20 compared with their mature counterparts. This observation suggests that canonical combinations of MAIT TCRs are selectively favored during the final stages of maturation, potentially ensuring functional optimization and long-term survival of mature MAIT cells.

Overall, these findings highlight the critical role of MR1 or CD1d recognition in defining the specificity of canonical TCR $\alpha$  chains. Moreover, our data suggest that MAIT cells undergo a process of clonal selection during maturation, favoring specific TCR combinations that may be crucial for their functional stability and survival.

### The diverse repertoire of MAIT and iNKT cells is shaped in the thymus

The above findings show that the diversity of MAIT and iNKT cells arises from oligoclonal TCR $\beta$  chains paired with canonical TCR $\alpha$  chains. To further investigate their TCR usage, we analyzed the lengths and sequences of CDR3 loci (Fig. 5a). As expected, the canonical TRAV genes in MAIT and iNKT cells exhibited uniform CDR3 $\alpha$  lengths with highly conserved sequences, whereas their CDR3 $\beta$  regions displayed significant diversity in both lengths and sequences (Fig. 5a), present in both mature and immature populations (Supplementary Fig. 12). Furthermore, no significant changes were observed in the preferential usage of TCR genes across different age groups (Fig. 5b). The chord diagram illustrates that the combinations of TCR Ja and V $\beta$  usage are highly conserved in both MAIT and iNKT cells (Fig. 5c).

A previous study has shown that TRAJ usage influences the pairing between TRAV and TRBV chains of MAIT cells in PBMCs<sup>5</sup>. We also observed that TRAJ12 and TRAJ20 preferentially bind to TRBV6-4 in the thymus (Fig. 5d). In iNKT cells, which predominantly utilize TRAJ18, TRAJ chains other than TRAJ18 exhibited minimal pairing with TRBV25-1 (Fig. 5e). Moreover, the promiscuous pairing of canonical TCR $\alpha$  chains, previously described in the periphery<sup>5</sup>, was also detected in the thymus, where approximately 30% of MAIT cells exhibited more than one pairing with TCR $\beta$  chains (Fig. 5f). Although not statistically significant, iNKT cells showed a more limited pattern of pairing, consistent with their strict usage of canonical TCRs (Figs. 4e and 5f).

In both MAIT and iNKT cells, CDR3 $\alpha$  nucleotide sequences were public, shared across all four donors (Fig. 5g, top left) in both immature and mature cells (Supplementary Fig. 13). Conversely, CDR3 $\beta$  sequences were private (Fig. 5g, top middle), and the combinations of CDR3 $\alpha$  and CDR3 $\beta$  were not shared between donors (Fig. 5g, top right), aligning with previously described peripheral diversity in human MAIT cells<sup>5</sup>. These patterns were also observed in iNKT cells (Fig. 5g, bottom), confirming that the diversity of iNKT cells is similarly shaped.

Collectively, these results indicate that the diverse repertoire of MAIT cells is shaped during thymic development rather than by peripheral selection, and iNKT cells follow similar trends.

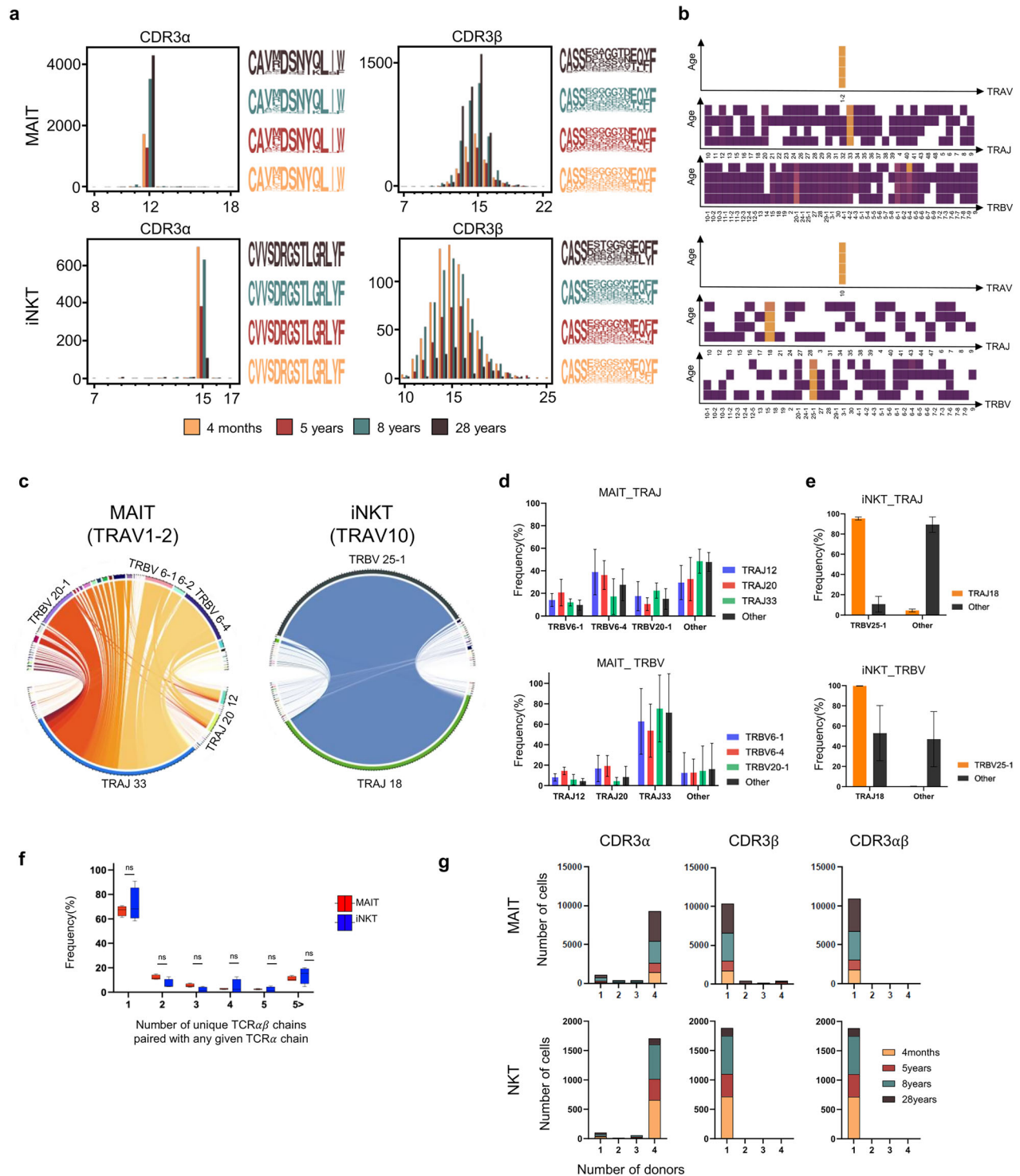
### MAIT cells clonally expand in the adult human thymus

Finally, we investigated the clonal expansion of MAIT and iNKT cells in the thymus (Fig. 6). A clonotype is defined as a group of cells sharing identical nucleotide sequences across six CDR regions, indicating their derivation from a single progenitor cell. We observed that nearly all cells in stages 0–2, and approximately half of the cells in stage 3, were clonally unique, indicating that immature cells do not proliferate after positive selection (Fig. 6a). Notably, the fractions of single clonotypes showed a decreasing trend with aging (Fig. 6a, top right). Furthermore, certain clonotypes expanded over 100-fold in stage 3, predominantly originating from the thymus of a 28-year-old donor (Fig. 6a, top middle). Consequently, the Shannon diversity index, which measures clonal diversity, was significantly reduced in stage 3 MAIT cells from this donor. A slight reduction in the Shannon index was also observed in stage 3 MAIT cells from an 8-year-old donor, confirming the result in multiple donors (Fig. 6b). Importantly, these findings are consistent with our flow cytometric analysis, which showed that stage 3 MAIT cells expand in proportion to aging (Fig. 1c). Stage 3 MAIT and iNKT cells exhibited higher tissue residency scores than their immature counterparts or classical  $\gamma\delta$  T cells (Fig. 6c) across all donors (data not shown). The clonal expansion of MAIT cells was correlated with aging, whereas iNKT cells showed a clonal pattern similar to MAIT cells in young MAIT donors (Fig. 6d). In summary, our results indicate that mature MAIT cells undergo clonal expansion in the human thymus. A comprehensive summary of our findings is shown in Fig. 6e, which illustrates the distinct maturation patterns of MAIT and iNKT cells in the human thymus.

### DISCUSSION

In this study, we conducted an extensive analysis of the developmental landscape of MAIT cells across different ages, examining their associations with iNKT cells and classical  $\gamma\delta$  T cells in the human thymus. We found that mature MAIT cells clonally expand in the human thymus, and they are distinct from iNKT cells in several features: iNKT cells express PLZF and CD4 at immature stages and progress to stage 3 less efficiently. In addition, we defined markers CD27 (not CD4) and CD161 as representing the maturation stages of iNKT cells. Our data also showed that the diverse repertoire of MAIT and iNKT cells is established in the thymus by pairing public TCR $\alpha$  with private oligoclonal TCR $\beta$  chains in the immature stage. We also provide evidence that MAIT cells undergo clonal selection upon maturation and proliferation with aging in the thymus, leading to the accumulation of expanded clonotypes in adult donors. Overall, these results provide new insights into the developmental landscape of human MAIT cells in the thymus.

We demonstrated that stage 3 MAIT cells underwent robust clonal expansion in the thymus of a 28-year-old donor and, to a lesser extent, in an 8-year-old donor (Fig. 6). Our findings are consistent with demographic data that used multiple samples indicating that stage 3 CD8 SP MAIT cells expand within the thymus (Fig. 1c). This suggests that the clonal expansion of mature MAIT cells may be an age-associated phenomenon, potentially reflecting their role in the adaptive immune response as individuals age. We also found that the frequency of stage 1 MAIT cells among total thymocytes remains constant in adults (Fig. 1c). These features are in contrast to the finding that peripheral MAIT cells decrease after adulthood (Fig. 1f and ref. 43). CD8 expression facilitates MR1 tetramer binding<sup>44</sup>, raising the technical issues of detecting mature MAIT cells. However, it is unlikely that this happens only in aged donors. Rather, it is possible that the homeostatic expansion of MAIT cells in the blood is

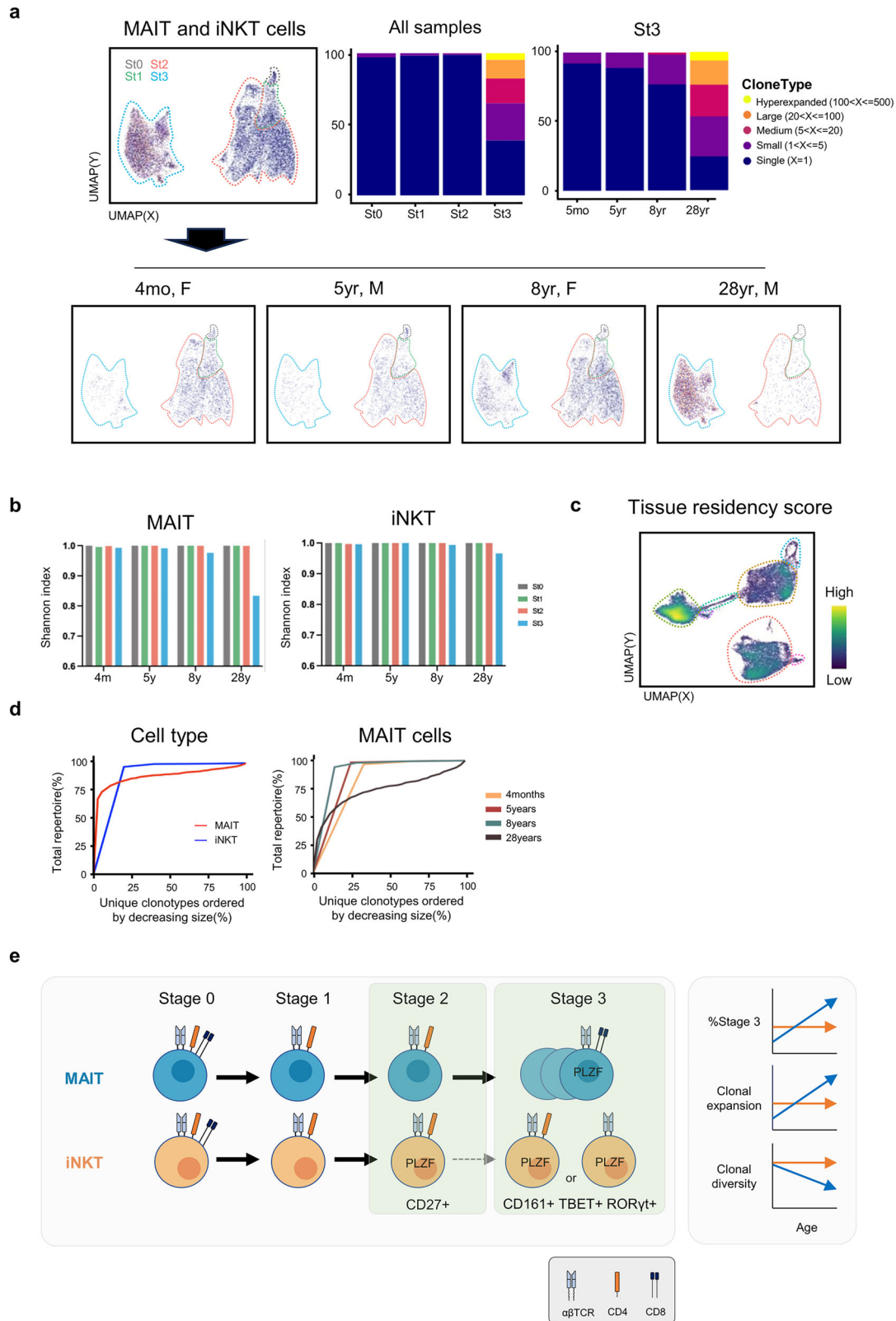


**Fig. 5 The diverse repertoire of MAIT and iNKT cells is shaped in the thymus.** **a** Graphs show the distribution of CDR3 $\alpha/\beta$  amino acid lengths for the indicated cell types and their corresponding amino acid sequences. **b** Heat maps illustrate TCR gene usage in each donor. **c** Chord diagrams depict the connectivity patterns between TRAJ and TRBV in MAIT and iNKT cells. **d, e** Graphs show the frequencies of each TRAJ and TRBV combination in MAIT (**d**) and iNKT (**e**) cells. **f** The graph shows the frequencies of cells having unique TCR $\alpha\beta$  chains paired with any given TCR $\alpha$  chain. **g** Graphs show the number of shared clones of MAIT and iNKT cells between donors based on CDR3 $\alpha$ , CDR3 $\beta$  or the CDR3 $\alpha\beta$  combination. Error bars indicate s.e.m.

less robust than that of conventional T cells or that MAIT cells become more tissue resident and do not leave the thymus in adults. Mature murine thymic NKT cells persist in the thymus<sup>16</sup>, and our data suggest human MAIT cells have similar trends, suggesting they

form a resident memory population, which warrants further investigation into their biological significance.

While our study provides detailed insights into the thymic ontogeny of MAIT cells, several limitations should be noted. The



**Fig. 6 MAIT cells clonally expand in the adult human thymus.** **a** UMAP plots show the distribution of the repeated clonotypes in total (top, right) and each donor (bottom) among total MAIT and iNKT cells. The bar graphs show the fraction of repeated clones in each stage (top, middle) and donor (top, right). **b** The graphs show Shannon index illustrating TCR diversity in each donor and cell type. **c** The density plot shows the tissue residency score. **d** Line plots show the clonality colored by cell type (left) and donor (right). **e** The figure summarizes the proposed model of thymic development of human unconventional T cells.

limited number of adult thymic samples restricts our ability to fully capture age-related dynamics in single-cell analysis with statistical significance. Although stage 3 MAIT cell expansion was prominent in a donor aged 28 and, to a lesser extent, in the 8-year-old donor, further validation in elderly individuals is needed. In addition, while we considered TRAV1-2- and TRAV10-positive cells as MAIT and iNKT cells, respectively, functional validation for these TCRs binding to tetramers remains necessary to confirm their repertoire selection during maturation (Fig. 4e, f).

MAIT and iNKT cells share developmental markers yet exhibit divergent maturation and clonal expansion patterns, particularly with aging. Only MAIT cells undergo robust clonal expansion and acquire tissue-residency signatures in adults. One speculation on these differences is that human thymi continuously provide ligands for MAIT TCRs, while such ligands for iNKT cells are limited. In contrast to humans, murine NKT1 cells in the thymus exhibit similar features, becoming long-lived tissue-resident populations<sup>16</sup>, whereas MAIT cells are rare in mice. The differences in self or nonself ligands expressed in the thymus could be a leading factor determining cell-type-specific residency in different species.

Functionally, MAIT cells contribute to antimicrobial defense and are implicated in chronic inflammation and autoimmune diseases<sup>45</sup>. Their capacity to produce IFN- $\gamma$  and IL-17 and persist as tissue-resident populations positions them as key players in mucosal immunity. One potential implication of clonally expanded mature MAIT cells in the adult thymus is that they might provide inflammatory cytokines to developing immature conventional T cells in the thymus, potentially conditioning them to become more responsive in the periphery. This speculation requires further investigation in the future.

Previous studies on MAIT cells either did not annotate TCR information<sup>8,9</sup> or excluded all T cells having dual TCR $\alpha$  transcripts<sup>5</sup>. So far, the nature of dual TCR $\alpha$  transcripts or polyclonal repertoire of MAIT and iNKT cells has not been thoroughly explored. Here, we demonstrated that most polyclonal TCRs do not bind to MR1 or CD1d tetramers (Fig. 4c and Supplementary Fig. 11). This suggests that the expression of canonical TCR transcripts of MAIT and iNKT cells is sometimes surpassed by those of polyclonal TCRs. In theory, dual TCR $\alpha$  transcripts are present in about 30–35% of human T cells, resulting from the simultaneous recombination of two TCR $\alpha$  alleles following  $\beta$ -selection, which occurs sequentially with stringent allelic exclusion<sup>46–48</sup>. Given that the detection rate of TCR $\alpha$  sequences in scRNA-seq is approximately 70%, the frequency of dual TCR $\alpha$  transcripts detected in single-cell sequencing is calculated to be ~19–22%. A recent meta-analysis found that 15% of human thymocytes express dual TCR $\alpha$  transcripts<sup>41</sup>, and our result was 16.2% (corresponding to 25% in the actual value), falling between the theoretical and observed values of conventional T cells. This suggests that human MAIT and iNKT cells are derived from DP thymocytes undergoing random TCR rearrangement, similar to murine iNKT cells<sup>49</sup>. Quality control results also support that our findings are not due to the incomplete exclusion of doublets, and we estimate the contamination frequency to be less than 5% when assessing the proportions of polyclonal cells lacking canonical TCR expression (Supplementary Fig. 4).

A previous study has shown that iNKT cells are rare in postnatal human thymus and almost undetectable using anti-TCR antibodies<sup>50</sup>. However, recent studies using CD1d tetramer-based enrichment have provided insight into the transcriptional characteristics of iNKT cells in both the human thymus<sup>8,9,51</sup> and blood<sup>52</sup>. These studies revealed transcriptional similarities between iNKT and MAIT cells, and our research further identified unique features of iNKT and  $\gamma\delta$  T cells, highlighting the heterogeneity of human unconventional T cells. In mice, the absence of iNKT or  $\gamma\delta$  T cells leads to a threefold increase in MAIT cell numbers<sup>12,53</sup>, suggesting a mutually suppressive relationship.

In humans, it remains unclear whether the developmental arrest of iNKT and  $\gamma\delta$  T cells at stage 2 in the thymus is influenced by MAIT cells or if these cells require extrathymic environments for final maturation.

Collectively, our findings suggest that human MAIT cells undergo clonal selection and expansion during thymic maturation unlike other types of innate T cell. These characteristics align more closely with conventional T cells selected by agonistic self-peptides despite the unique ability of MAIT cells to recognize 5-OP-RU, a metabolite derived from intestinal bacteria. These insights provide a deeper understanding of the mechanisms underlying thymic selection and the shaping of the T cell repertoire of unconventional T cells.

## DATA AVAILABILITY

Sequencing data generated in this study have been deposited in NCBI's Gene Expression Omnibus (GEO) and are accessible through GEO SuperSeries accession number [GSE254975](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE254975). The MSigDB database and WebLogo generator are publicly available.

## CODE AVAILABILITY

No custom software was generated for this manuscript. The code developed for the analyses performed in this study is available upon request. The software used for all analyses is listed in the relevant sections of the 'Materials and methods'.

## REFERENCES

- Legoux, F. et al. Microbial metabolites control the thymic development of mucosal-associated invariant T cells. *Science* **366**, 494–499 (2019).
- Provine, N. M. & Klenerman, P. MAIT cells in health and disease. *Annu. Rev. Immunol.* **38**, 203–228 (2020).
- Mayassi, T., Barreiro, L. B., Rossjohn, J. & Jabri, B. A multilayered immune system through the lens of unconventional T cells. *Nature* **595**, 501–510 (2021).
- Koay, H. F. et al. A three-stage intrathymic development pathway for the mucosal-associated invariant T cell lineage. *Nat. Immunol.* **17**, 1300–1311 (2016).
- Garner, L. C. et al. Single-cell analysis of human MAIT cell transcriptional, functional and clonal diversity. *Nat. Immunol.* **24**, 1565–1578 (2023).
- Meermeier, E. W. et al. Human TRAV1-2-negative MR1-restricted T cells detect *S. pyogenes* and alternatives to MAIT riboflavin-based antigens. *Nat. Commun.* **7**, 12506 (2016).
- Koay, H. F. et al. Diverse MR1-restricted T cells in mice and humans. *Nat. Commun.* **10**, 2243 (2019).
- Chandra, S. et al. Transcriptomes and metabolism define mouse and human MAIT cell populations. *Sci. Immunol.* **8**, eabn8531 (2023).
- Bugaut, H. et al. A conserved transcriptional program for MAIT cells across mammalian evolution. *J. Exp. Med.* <https://doi.org/10.1084/jem.20231487> (2024).
- Loh, L. et al. Unraveling the phenotypic states of human innate-like T cells: comparative insights with conventional T cells and mouse models. *Cell Rep.* <https://doi.org/10.1016/j.celrep.2024.114705> (2024).
- Kwon, D. I. & Lee, Y. J. Lineage differentiation program of invariant natural killer T cells. *Immune Netw.* **17**, 365–377 (2017).
- Lee, M. et al. Single-cell RNA sequencing identifies shared differentiation paths of mouse thymic innate T cells. *Nat. Commun.* **11**, 4367 (2020).
- Pellicci, D. G., Koay, H. F. & Berzins, S. P. Thymic development of unconventional T cells: how NKT cells, MAIT cells and gammadelta T cells emerge. *Nat. Rev. Immunol.* **20**, 756–770 (2020).
- Harsha Krovi, S. et al. Thymic iNKT single cell analyses unmask the common developmental program of mouse innate T cells. *Nat. Commun.* **11**, 6238 (2020).
- Krovi, S. H., Loh, L., Spengler, A., Brunetti, T. & Gapin, L. Current insights in mouse iNKT and MAIT cell development using single cell transcriptomics data. *Semin. Immunol.* **60**, 101658 (2022).
- Berzins, S. P., McNab, F. W., Jones, C. M., Smyth, M. J. & Godfrey, D. I. Long-term retention of mature NK1.1<sup>+</sup> NKT cells in the thymus. *J. Immunol.* **176**, 4059–4065 (2006).
- Lee, Y. J. et al. Generation of PLZF<sup>+</sup> CD4<sup>+</sup> T cells via MHC class II-dependent thymocyte-thymocyte interaction is a physiological process in humans. *J. Exp. Med.* **207**, 237–246 (2010).
- Min, H. S. et al. MHC class II-restricted interaction between thymocytes plays an essential role in the production of innate CD8<sup>+</sup> T cells. *J. Immunol.* **186**, 5749–5757 (2011).

19. Yang, C., Kwon, D. I., Kim, M., Im, S. H. & Lee, Y. J. Commensal Microbiome Expands Ty $\delta$ 17 cells in the lung and promotes particulate matter-induced acute neutrophilia. *Front. Immunol.* **12**, 645741 (2021).
20. Kwon, D. I. et al. Homeostatic serum IgE is secreted by plasma cells in the thymus and enhances mast cell survival. *Nat. Commun.* **13**, 1418 (2022).
21. Lee, E., Kim, M. & Lee, Y. J. Selective expansion of Tregs using the IL-2 cytokine antibody complex does not reverse established alopecia areata in C3H/HeJ mice. *Front. Immunol.* **13**, 874778 (2022).
22. Lee, M. et al. Dietary antigens suppress the proliferation of type 2 innate lymphoid cells by restraining homeostatic IL-25 production. *Sci. Rep.* **12**, 7443 (2022).
23. Gherardin, N. A. et al. CD36 family members are TCR-independent ligands for CD1 antigen-presenting molecules. *Sci. Immunol.* <https://doi.org/10.1126/sciimmunol.abg4176> (2021).
24. Bettini, M. L., Bettini, M., Nakayama, M., Guy, C. S. & Vignali, D. A. Generation of T cell receptor-retrogenic mice: improved retroviral-mediated stem cell gene transfer. *Nat. Protoc.* **8**, 1837–1840 (2013).
25. Hao, Y. et al. Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587 (2021).
26. Wolock, S. L., Lopez, R. & Klein, A. M. Scrublet: computational identification of cell doublets in single-cell transcriptomic data. *Cell Syst.* **8**, 281–291 (2019).
27. Korsunsky, I. et al. Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat. Methods* **16**, 1289–1296 (2019).
28. Borchering, N., Bormann, N. L. & Kraus, G. scRepertoire: an R-based toolkit for single-cell immune receptor analysis. *F1000Res* **9**, 47 (2020).
29. Bunis, D. G., Andrews, J., Fragiadakis, G. K., Burt, T. D. & Sirota, M. dittoSeq: universal user-friendly single-cell and bulk RNA sequencing visualization toolkit. *Bioinformatics* **36**, 5535–5536 (2021).
30. Alquicira-Hernandez, J. & Powell, J. E. Nebulosa recovers single-cell gene expression signals by kernel density estimation. *Bioinformatics* **37**, 2485–2487 (2021).
31. Andreatta, M. & Carmona, S. J. UCell: robust and scalable single-cell gene signature scoring. *Comput. Struct. Biotechnol. J.* **19**, 3796–3798 (2021).
32. Chen, E. Y. et al. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinform.* **14**, 128 (2013).
33. Korsunsky, I., Nathan, A., Millard, N. & Raychaudhuri, S. Presto scales Wilcoxon and auROC analyses to millions of observations. Preprint at *bioRxiv* <https://doi.org/10.1101/653253> (2019).
34. Korotkevich, G. et al. Fast gene set enrichment analysis. Preprint at *bioRxiv* <https://doi.org/10.1101/060012> (2021).
35. Trapnell, C. et al. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat. Biotechnol.* **32**, 381–386 (2014).
36. Qiu, X. et al. Single-cell mRNA quantification and differential analysis with Census. *Nat. Methods* **14**, 309–315 (2017).
37. Gu, Z., Gu, L., Eils, R., Schlesner, M. & Brors, B. circlize Implements and enhances circular visualization in R. *Bioinformatics* **30**, 2811–2812 (2014).
38. Crooks, G. E., Hon, G., Chandonia, J. M. & Brenner, S. E. WebLogo: a sequence logo generator. *Genome Res.* **14**, 1188–1190 (2004).
39. Baev, D. V. et al. Distinct homeostatic requirements of CD4<sup>+</sup> and CD4<sup>-</sup> subsets of Valpha24-invariant natural killer T cells in humans. *Blood* **104**, 4150–4156 (2004).
40. Perriman, L. et al. A three-stage developmental pathway for human Vgamma9Vdelta2 T cells within the postnatal thymus. *Sci. Immunol.* **8**, eabo4365 (2023).
41. Zhu, L. et al. scRNA-seq revealed the special TCR beta & alpha V(D)J allelic inclusion rearrangement and the high proportion dual (or more) TCR-expressing cells. *Cell Death Dis.* **14**, 487 (2023).
42. Suliman, S. et al. Dual TCR-alpha expression on mucosal-associated invariant T cells as a potential confounder of TCR interpretation. *J. Immunol.* **208**, 1389–1395 (2022).
43. Kurioka, A. & Klenerman, P. Aging unconventionally: gammadelta T cells, iNKT cells, and MAIT cells in aging. *Semin. Immunol.* **69**, 101816 (2023).
44. Souter, M. N. T. et al. CD8 coreceptor engagement of MR1 enhances antigen responsiveness by human MAIT and other MR1-reactive T cells. *J. Exp. Med.* <https://doi.org/10.1084/jem.20210828> (2022).
45. McWilliam, H. E. G. & Villadangos, J. A. MR1 antigen presentation to MAIT cells and other MR1-restricted T cells. *Nat. Rev. Immunol.* <https://doi.org/10.1038/s41577-023-00934-1> (2023).
46. Huang, C. & Kanagawa, O. Ordered and coordinated rearrangement of the TCR alpha locus: role of secondary rearrangement in thymic selection. *J. Immunol.* **166**, 2597–2601 (2001).
47. Dupic, T., Marcou, Q., Walczak, A. M. & Mora, T. Genesis of the q $\beta$  T-cell receptor. *PLoS Comput Biol.* **15**, e1006874 (2019).
48. Schuldt, N. J. & Binstadt, B. A. Dual TCR T cells: identity crisis or multitaskers? *J. Immunol.* **202**, 637–644 (2019).
49. Egawa, T. et al. Genetic evidence supporting selection of the Valpha14i NKT cell lineage from double-positive thymocyte precursors. *Immunity* **22**, 705–716 (2005).
50. Sandberg, J. K., Stoddart, C. A., Brilot, F., Jordan, K. A. & Nixon, D. F. Development of innate CD4<sup>+</sup> alpha-chain variable gene segment 24 (Va24) natural killer T cells in the early human fetal thymus is regulated by IL-7. *Proc. Natl Acad. Sci. USA* **101**, 7058–7063 (2004).
51. Maas-Bauer, K. et al. Single-cell transcriptomics reveal different maturation stages and sublineages commitment of human thymic invariant natural killer T cells. *J. Leukoc. Biol.* <https://doi.org/10.1093/jleuko/qiad113> (2023).
52. Zhou, L. et al. Single-cell RNA-seq analysis uncovers distinct functional human NKT cell sub-populations in peripheral blood. *Front. Cell Dev. Biol.* **8**, 384 (2020).
53. Xu, C. et al. Expansion of MAIT cells in the combined absence of NKT and  $\gamma\delta$ -T cells. *Mucosal Immunol.* **16**, 446–461 (2023).

## ACKNOWLEDGEMENTS

This work was supported by the Korean Ministry of Science, Information/Communication Technology and Future Planning (RS-2022-NR067484 and RS-2024-00346548), by a grant from the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (RS-2024-00438759) and by the four phases of the Brain Korea 21 Program in 2024 by the Ministry of Education, Korea.

## AUTHOR CONTRIBUTIONS

M.L. and S.P. designed and performed experiments, analyzed data and wrote the draft; S.P. analyzed scRNA-seq data; S.Y.B., H.S.R. and J.G.K. collected human samples; M.-S.K. performed TCR cloning; J.G.K. and Y.J.L. wrote the manuscript; Y.J.L. conceptualized the research. J.-H.C. revised and finalized the scRNA-seq data analysis.

## COMPETING INTERESTS

The authors declare no competing interests.

## DECLARATION OF GENERATIVE AI AND AI-ASSISTED TECHNOLOGIES IN THE WRITING PROCESS

During the preparation of this work, the authors used chatGPT4.0 to improve language and readability. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

## ADDITIONAL INFORMATION

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s12276-025-01509-x>.

**Correspondence** and requests for materials should be addressed to Jae Gun Kwak or You Jeong Lee.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**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 licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence 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 licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2025