Integrated single-cell sequencing analysis reveals peripheral immune landscape across the human lifespan

Renyang Tong[1#], Ronghong Li[1#], Yichao Zhao[1#], Lingjie Luo[2#], Jianmei Zhong[1], Zheng Li[1], Lin Wei[1], Yifan Chen[1], Jianfeng Shi[1], Yu Gao[1], Mingze Sun[1], Yuyan Lyu[1], Ancai Yuan[1], Lingyue Sun[1], Yuping Guo[1], Honglin Wang[3], Haoyan Chen[4], Lei Chen[3], Bin Li[3], Wei Lin[1], Feng Wang[3], Li Wang[2], Jun Pu[1*]

[1] State Key Laboratory for Oncogenes and Related Genes, Division of Cardiology, School of Medicine, Renji Hospital, Shanghai Cancer Institute, Shanghai Jiao Tong University, Shanghai, China.

[2] State Key Laboratory of Cardiovascular Disease, Fuwai Hospital, National Center for Cardiovascular Diseases, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China.

[3] Shanghai Institute of Immunology, Department of Immunology and Microbiology, State Key Laboratory of Oncogenes and Related Genes, Shanghai Jiao Tong University School of Medicine, Shanghai, China.

[4] State Key Laboratory for Oncogenes and Related Genes; Key Laboratory of Gastroenterology & Hepatology, Ministry of Health; Division of Gastroenterology and Hepatology; Shanghai Cancer Institute; Shanghai Institute of Digestive Disease; Renji Hospital, Shanghai Jiao Tong University School of Medicine, 145 Middle Shandong Road, 200001, Shanghai, China.

[*] To whom correspondence should be addressed:

Jun Pu, MD, Ph.D.; State Key Laboratory for Oncogenes and Related Genes, Division of Cardiology, School of Medicine, Renji Hospital, Shanghai Cancer Institute, Shanghai Jiao Tong University, Shanghai, China. E-mail: pujun310@hotmail.com.

[#]These authors contributed equally to this work

SUMMARY

Systematic understanding of immune dynamics across the entire human lifespan at single-cell resolution is currently lacking. Here, we performed single-cell RNA sequencing (scRNA-seq) and single-cell T cell receptor (TCR)/B cell receptor (BCR) sequencing (scTCR/BCR-seq) on over 380,000 peripheral blood mononuclear cells collected from 45 healthy participants aged 0 to over 90 years. We revealed that the functions of T cell subsets were most susceptible to senescence among all PBMCs, featured by increased NF-κB signaling and IFN-γ responses pathways, while reduced telomere maintenance and energy metabolism. We subsequently explored the rewiring of cell-cell interactions among different immune cells across the lifespan and revealed the major alteration of immune checkpoints in T cells within the cellular interaction networks. By combined analysis of scRNA-seq and scTCR-seq, we revealed that 1) $GNLY^{+}$ CD8 Effector T cells exhibited a high clonal expansion with distinct functional signatures in children and the elderly; 2) Naïve $CD4^{+}$ T and naïve $CD8^{+}$ T cells displayed different aging patterns in both transcriptomes and immune repertoires; and 3) $CD8^{+}$ MAIT cells showed a higher cell abundance and clonal diversity in adolescents. Furthermore, we identified a unique cytotoxic B cell subset enriched in children by scRNA-seq and scBCR-seq analysis and experimental validations. In summary, our work provided valuable insights and rich resources for understanding the development and aging of the human immune system across the lifespan.

Keywords: Lifespan; Single-cell RNA sequencing; single-cell T cell receptor/B cell receptor sequencing; Peripheral blood mononuclear cells; T cells; Cytotoxic B cell

INTRODUCTION

Development and aging run through the human life with gradual maturation and inevitable degeneration of the immune system (Mittelbrunn and Kroemer, 2021; Vijg and Calder, 2004). As a key defense system, immune system protects the host against various exogenous invasions and endogenous challenges (e.g., viruses, bacteria, parasites, and cancer cells) (López-Otín and Kroemer, 2021). Human beings tend to be less vulnerable to foreign pathogens as the immune system gradually matures after birth (Simon et al., 2015). By contrast, immunosenescence leads to a higher incidence of infection, autoimmune diseases and cancer with aging (Montecino-Rodriguez et al., 2013; Nikolich-Zugich, 2018). Moreover, aged immune cells drive aging of solid organs, such as the heart and liver, which is a probable cause of systemic senescence (Yousefzadeh et al., 2021). Yet, little is known about when and how the immune system changes across the entire lifespan. Therefore, there is an urgent need to systematically understand the physiology/pathology of immune development and immune senescence across the human lifespan.

Human peripheral blood mononuclear cells (PBMCs) are a diverse mixture of highly specialized immune cells that are critical components of the immune system. Prior omics studies of human PBMCs have revealed their involvement in immune pathologies [e.g., immunosenescence (Alpert et al., 2019; Mogilenko et al., 2021; Zhang et al., 2019), autoimmune disorders (Nehar-Belaid et al., 2020) and infectious diseases (Ren et al., 2021; Zhu et al., 2020)] as well as in several diseases of solid organs, such as cardio-cerebrovascular diseases (Mazzola et al., 2008; Moore et al., 2005) and acute-on-chronic liver failure (Li et al., 2022). Some significant age-related changes in PBMCs composition and function from healthy donors have been demonstrated using flow cytometry or bulk sequencing. For example, flow cytometry analysis demonstrated that numbers of $CD19^+$ B cells and naive T cells decreased with age, whereas antigen-experienced memory T cell numbers increased with concomitant loss of co-stimulation factors CD27 and CD28 (Fagnoni et al., 2000; Qin et al., 2016). Moreover, bulk RNA sequencing identified a large group of age-associated genes in PBMCs and whole blood, which was used to construct prediction models for “transcriptomic age” (Peters et al., 2015). These findings provided valuable insights into the mechanism of immune aging, whereas these assays based on canonical cell markers and classification criteria might be insufficient to reflect the

sophistication and heterogeneity of peripheral immune cells.

Single-cell RNA sequencing (scRNA-seq) and single-cell T cell receptor (TCR)/B cell receptor (BCR) sequencing (scTCR/BCR-seq) provide an opportunity to comprehensively identify previously undefined cell types at single-cell resolution (Ahmed et al., 2019; See et al., 2017; Villani et al., 2017). It can provide novel details into both development- and aging-related processes by deconvoluting molecular features within specific cell types (Han et al., 2020; Ma et al., 2020). Several scRNA-seq studies on peripheral immune cells conducted at specific ages have yielded valuable findings at single-cell resolution: 1) compared with adult PBMCs from public database, scRNA-seq of neonate umbilical cord blood (UCB) identified an enrichment of a previously unrecognized cytotoxic innate immune cell type, namely $GZMB^+$ natural killer T (NKT) cells, although the specific functions of this UCB-specific immune cells remain unclear (Zhao et al., 2019); 2) compared with young controls (25 to 29 years old), clonal $GZMK^+$ $CD8^+$ T cells were identified as a hallmark of inflammaging in order individuals (62 to 70 years old) (Mogilenko *et al.*, 2021); and 3) compared with senior controls (aged 50 to 89 years old), scRNA-seq data revealed the expansion of cytotoxic $CD4^+$ T cells as a key feature of healthy immunosenescence in very elderly individuals (i.e., supercentenarian reaching 110 years of age) (Hashimoto et al., 2019). These scRNA-seq studies comparing between two specific age groups provided valuable insights into age-related changes in peripheral immune cells at single-cell resolution. However, a systematic understanding of peripheral immune cell dynamics across the entire human lifespan at single-cell resolution is currently lacking.

Here, by combined analysis of scRNA-seq and scTCR/BCR-seq, we systematically investigated dynamic changes in the cell-type composition, transcription profiles, cell-cell interactions and immune repertoires of human peripheral immune cells across the entire lifespan, using PBMC samples from 45 healthy participants aged 0 (neonates) to ≥ 90 y (nonagenarians). Our work provided a rich resource for understanding the development and aging of human peripheral immune cells. We compiled these analyses into a web portal (https://pu-lab.sjtu.edu.cn/shiny) for exploring the data easily.

RESULTS

Overview of single-cell transcriptomes of peripheral immune cells across the entire lifespan

To characterize the dynamics of peripheral immune cell composition and function throughout the human lifespan, 45 healthy participants aged 0 (neonates) to ≥ 90 years (nonagenarians) were recruited, which spanned ten age groups including five neonates (umbilical cord blood, UCB group), three infants aged 1 year (1 y group), five toddlers aged years (2 y group), six preadolescents aged 6 years (6 y group), four adolescents aged 12 years (12 y group), five adolescents aged 18 years (18 y group), four young adults aged 30 years (30 y group), five middle-aged adults aged 50 years (50 y group), five septuagenarians aged 70 years (70 y group), and four nonagenarians aged ≥ 90 years (90 y group). All participant details are shown in Table S1. We performed single-cell RNA, TCR and BCR sequencing simultaneously on PBMCs from each individual, investigated the dynamics of cell type composition and immune cell functions across the lifespan (Figure 1A).

After filtering out low-quality droplets, more than 380,000 cells were obtained from all samples. The filtered data were integrated, dimension reduced and clustered in an unsupervised method. Based on the unsupervised clustering and expression of signature genes for each subset (Table S2), 31 distinct PBMC subsets were identified and visualized via uniform manifold approximation and projection (UMAP) (Figures 1B and S1A). Specifically, there were four $CD4^+$ T cell subsets, including naïve $CD4^+$ T cells (CD4_Naive_CCR7), effector memory $CD4^+$ T cells (CD4_Tem_ANXA1), type 1 helper $CD4^+$ T cells (CD4_Th_STAT1), and regulatory $CD4^+$ T cells (CD4_Treg_FOXP3); seven $CD8^+$ T cell subsets, including two types of naïve $CD8^+$ T cells (CD8_Naive_LINC02446^{middle} and CD8_Naive_LINC02446^{high}), two types of effector $CD8^+$ T cells (CD8_Effector_GZMK and CD8_Effector_GNLY), tissue resident memory $CD8^+$ T cells (CD8_Trm_ZNF683), mucosal-associated invariant $CD8^+$ T cells (CD8_MAIT_SLC4A10) and exhausted $CD8^+$ T cells (CD8_Exhausted_HAVCR2); one subset of γδ T cells; eight NK cell subsets, including NK cells with high expression of *NCAM1*(encoding CD56) (NK_CD56^{high}_GZMK), and six NK cell subsets with low expression of *NCAM1* (NK_CD56^{low}_GZMK, NK_CD56^{low}_FCGR3A, NK_CD56^{low}_CCL4L2, NK_CD56^{low}_HLA-DRA, NK_CD56^{low}_PTGDS and NK_CD56^{low}_TKTL1), and NKT cells; five BCR-positive cell subsets, including naïve B cells (B_Naive),

memory B cells (B_Memory), atypical memory B cells (B_atypical_Memory), plasma cells (Plasma cell), and a previous undefined, unique B cell subset (we named it as B_BCR$^+$GNLY$^+$, as detailed below); six myeloid cell subsets, including CD14$^+$ classical monocytes (CD14$^+$ Monocytes_VCAN), CD14$^+$ monocytes with high expression of *MX1* (CD14$^+$ Monocytes_MX1), CD16$^+$ monocytes, myeloid dendritic cells (mDCs), plasmacytoid dendritic cells (pDCs), and megakaryocytes (Mega). The representative signature genes of all these cell subsets and their expression levels were shown in heatmaps and feature plots (Figures 1C and S1B).

Among BCR-positive cell subsets, we identified a previously unrecognized B cell type. We named this new-identified B cell type as B_BCR$^+$GNLY$^+$, since it highly expressed general genes encoding B cell markers (e.g., BCR and *JCHAIN*) and canonical cytotoxic genes such as *GNLY* and *NKG7*, but did not express several other canonical B cell marker genes (i.e., *MS4A1*, *CD19* and *CD79B*) (Figures 1C and S1B), indicating that these cells might represent a cytotoxic B cell type. Meanwhile, negative TCR, *TRAC* and *KLRF1* expression excluded the possibility of these cells being T and NK cells. Moreover, we excluded the potential technical artifact (i.e., multiplets) based on the following: 1) the number of detected genes in the B_BCR$^+$GNLY$^+$ cells were relatively lower than that in other cell subsets (Figure S1C); and 2) a stringent doublet filtration standard was performed to exclude potential doublets.

Cell composition dynamics of peripheral immune cells across lifespan

To obtain a global view of the dynamics of PBMC composition throughout human life, the proportion of each of the 31 cell subsets in total PBMCs was calculated for each sample (Figures S2A). Among them, 21 cell subsets exhibited significant differences in their proportion in total PBMCs across different age groups (ANOVA, FDR < 0.05). To further illustrate the age preference of these cell subsets, the ratio of observed cell numbers to randomly expected ones ($R_{O/E}$) (Zhang et al., 2018) was calculated. All these 21 cell subsets exhibited differential age preference (Figure 1D). Among these subsets, the proportion of certain cell subsets was positively correlated with age, as exemplified by CD8_Effector_GZMK cells (Spearman correlation coefficient R = 0.7082, p = 5.26×10^{-8}), CD4_Treg_FOXP3 cells (R = 0.5174, p = 0.0002) and NK_CD56^{low}_HLA-DRA cells (R = 0.4882, p = 0.0006) (Figures 1D and S2B). By contrast, the proportion of certain cell subsets was negatively correlated with age,

such as NK_CD56^{high}_GZMK cells (R = -0.5936, p = 1.72×10^{-5}) and B_Naive cells (R = -0.5579, p = 6.83×10^{-5}) (Figures 1D and S2C). Several cell subsets exhibited an enrichment at certain age stages. For example, CD14^{+} Monocytes_VCAN cells were enriched in UCB group; B_BCR^{+}GNLY^{+} and plasma cells were enriched in children (2 y and 6 y groups); and CD8_MAIT_SLC4A10 cells were enriched in adolescents (12 y and 18 y groups) (Figures 1D, and S2D). Overall, we detected differential correlations of distinct PBMC subsets with age at single-cell resolution.

Global transcriptome dynamics of peripheral immune cells throughout lifespan

To systematically investigate the transcriptome dynamics of PBMC subsets across the lifespan, we compared the expression profiles of each cell subset among different age groups. The differentially expressed genes (DEGs) among different age groups were identified in each subset. Then, the 31 PBMC subsets were ranked based on the number of DEGs (Figure 2A). The top ten cell subsets (DEG numbers > 300) belonged to lymphoid lineage cell subsets (i.e., T, NK and B cell types), indicating that these cell subsets were more susceptible to age. By contrast, not any myeloid lineages (i.e., monocytes, DCs and Mega) ranked the top ten cell subsets with DEG numbers > 300 (Figure 2A). Although eight of the top ten cell subsets belonged to T cells, significant difference in the number of DEGs was observed among different T cell subsets. Among T cell subsets, CD8_Trm_ZNF683 cells expressed the highest DEG numbers with 606 genes, however CD4_Treg_FOXP3 cells expressed the lowest DEG number with 98 genes (Figure 2A). Likewise, among NK cell subsets, NK_CD56^{low}_FCGR3A cells expressed the highest DEG numbers with 460 genes, whereas NK_CD56^{low}_TKTL1 cells expressed the lowest DEG number with 37 genes (Figure 2A). To gain an overview of functional dynamics of PBMCs across the life course, the DEGs in each subset were subjected to functional enrichment analysis. These DEGs were enriched in the biological processes related to antigen processing/presentation (in 29 cell subsets), interferon-gamma (IFN-γ) responses pathway (in 28 cell subsets), lymphocyte activation (in 27 cell subsets), ATP metabolic process (in 20 cell subsets), NF-κB signaling (in 18 cell subsets), and telomere maintenance (in 16 cell subsets) (Figure 2B).

Next, we focused on age-related genes that were monotonically correlated to age [named monotonical DEGs (M-DEGs); Upregulated or Downregulated] (details in Supplementary Methods). Detailed M-DEGs in each cell subset are listed in Table S3.

Then, 31 PBMC subsets were ranked based on the number of M-DEGs (Figures S3A and S3B). All the top ten cell subsets (M-DEGs numbers > 100) belonged to lymphoid lineage cell subsets (i.e., T, NK and B cell types), implied that these lymphoid lineage cell subsets were susceptible to senescence. The top ten cell subsets included eight T cell subsets (CD8_Trm_ZNF683, CD4_Naive_CCR7 and CD8_Naive_LINC02446^{high}, CD8_Naive_LINC02446^{middle}, CD4_Tem_ANXA1, CD4_Th_STAT1, CD8_Effector_GZMK and CD8_Effector_GNLY cells), one NK cell subset (NK_CD56low_FCGR3A cells) and one B cell subset (B_Memory cells). The top three cell subsets with the largest number of M-DEGs were T cell subsets: CD8_Trm_ZNF683, CD4_Naive_CCR7 and CD8_Naive_LINC02446^{high} cells (Figures S3A and S3B). Functional enrichment analysis showed that the upregulated and downregulated M-DEGs were enriched in different biological processes. For instance, upregulated M-DEGs were enriched in NF-κB signaling and IFN-γ responses pathways (Figure S3C), while downregulated M-DEGs were enriched in telomere maintenance and proton transmembrane transport pathways (Figure S3E). To determine the common M-DEGs shared by multiple cell subsets, we ranked all M-DEGs based on the number of cell subsets in which they were involved. Results showed that 21 M-DEGs were upregulated in ≥ 10 cell subsets (Figure 2C), and 34 M-DEGs were downregulated in ≥10 cell subsets (Figure 2D). For example, the expression of genes involved in the IFN-γ signaling pathway (i.e., *HLA-DRB1*, *HLA-DPA1*, *HLA-DPB1* and *B2M*) was significantly upregulated in more than 10 cell subsets (Figures 2C and S3D). By contrast, the expression of genes involved in maintaining telomere length (i.e., *CCT3*, *CCT2* and *CCT5*) and nicotinamide adenine dinucleotide phosphate (NADP) metabolic process (i.e., *TKT* and *PGLS*) was significantly downregulated with age in more than 10 cell subsets (Figures 2D and S3F). These common M-DEGs across different cell subsets suggested that these genes may represent important immune biomarkers across the lifespan. Together, these results suggest that the functions of T cell subsets were more susceptible to the senescence among all PBMCs.

Rewiring of cell-cell interaction among peripheral immune cells across the whole lifespan

Immunological processes involve interactions between various immune cell types to coordinate behavior of immune cells during an immune response (Hao et al., 2020;

Rieckmann et al., 2017). To characterize dynamics of cell-cell interactions (CCIs) among immune cells across the entire lifespan, we calculated the CCI probabilities [reflecting the communication probability between two cells (Efremova et al., 2020; Jin et al., 2021)] among all pairs of cell subsets based on the expression of known ligands and receptors in each sample (details in Supplementary Methods). Human ligand-receptor interaction pairs were downloaded from the CellTalkDB database (Shao et al., 2021). The varied significantly CCIs among different age groups (ANOVA, Bonferroni-adjusted $p < 0.05$) were determined by comparing the CCI probabilities of all cell subset pairs in each sample. Using hierarchal clustering, these varied significantly CCIs were clustered into five clusters, four of which were enriched in specific age groups (Figure 3A). The interaction hubs of each cluster were different (Figures 3A-3E). The results showed that 1) CCI probabilities of cluster 1 peaked in the UCB group (Figure 3A), and the interaction hub of this cluster was $CD14^+$ monocytes (Figure 3B); 2) CCI probabilities of cluster 2 peaked in children (1 y, 2 y, and 6 y groups) (Figure 3A), and their hubs changed to CD8_Exhausted_HAVCR2 and $B_BCR^+GNLY^+$ cells (Figure 3C); 3) CCI probabilities of cluster 3 peaked in middle-aged adults (50 y group) (Figure 3A), and the interaction hubs of this cluster were CD8_Effector_GZMK and CD4_Tem_ANAX1 cells (Figure 3D); and 4) CCI probabilities of cluster 4 peaked in the elderly (70 y and 90 y groups) (Figure 3A), with its hubs changed to NK_CD56^{low}_HLA-DRA and CD8_Trm_ZNF283 cells (Figure 3E).

Furthermore, more than 500 differentially expressed ligands and receptors among different age groups (ANOVA, $p < 0.001$) were identified in all cell subsets (detailed in Table S4). Functional enrichment analysis showed that these differentially expressed ligands and receptors were enriched in “leukocyte migration”, “leukocyte cell-cell adhesion”, and “T cell activation” pathways (Figure S4A). Consequently, a CCI network was constructed, which was mediated by differentially expressed ligand-receptor pairs (Figure 3F). The 19 known immune checkpoint targets in cancer immunotherapy (Zhang et al., 2020b) were identified among these differentially expressed ligands and receptors and circled with red dashed lines in Figure 3F. Next, the topological characteristics of the CCIs network were examined using the software Cytoscape (Shannon et al., 2003), which enables visualization, rendering, and analysis of networks in one computational environment. The results showed that both the node degree and the betweenness centrality of immune checkpoint receptors were higher

than those of non-immune-checkpoint receptors ($p = 0.01$ and $p = 0.023$, respectively) (Figure 3G). The higher nodes degree indicated that the nodes were likely to be hubs and therefore involved in more competing interactions, while the higher betweenness centrality of immune checkpoint receptors implied that they acted as bridges connecting different network components and controlling communication (Ahmed et al., 2018; Wang et al., 2015). These specific topological patterns reflected the functional importance of immune checkpoint receptors in the age-modulated CCI network, which was mediated by ligand-receptor pairs across peripheral immune cell subsets. Moreover, the expression of these immune checkpoints was significantly altered across the lifespan in multiple cell subsets (Figures 3H), particularly in T cell subsets (Figures 3I, 3J and S4B). For example, the expression of co-stimulatory checkpoint *CD27*, a promising immune-modulatory target in cancer treatment (Buchan et al., 2018), showed a decreased trend with age in ten T cell subsets (Figure 3I). Similarly, the expression of co-stimulatory checkpoint *ICOS* (Figure 3I) and co-inhibitory checkpoint *PDCD1* (Figure 3J) peaked in children (1 y, 2 y and 6 y groups) and exhibited a decreased trend with age in four T cell subsets. Conversely, the expression of co-inhibitory checkpoint *CTLA4* showed an increased trend with age in both naïve $CD4^+$ T and naïve $CD8^+$ T cells (Figure 3J). The characteristic changes of immune checkpoints across the lifespan indicated that the T cells functions influenced by immune checkpoints might play an essential role during normal human development and aging. Our results reveal major alterations of immune checkpoints at single-cell resolution across the human lifespan.

Clonal expansion of effector T cells in young and old individuals

The characteristic changes of immune checkpoints in T cells were involved in the regulation of T cell activation and proliferation. Thus, we explored the clonal expansion among individual T cells and their V(D)J genes features across the entire lifespan by reconstructed paired TCR sequences from single-cell sequencing data. We obtained matched TCR and transcriptome information from more than 70% of cells across all T cell subsets (Figure S5A). Notably, among ten age groups, high T-cell clonal expansions (high TCE; clone size>50 cells) was obvious in five age groups: 2 y and 6 y (children), 12 y (early teenagers), and 70 y and 90 y (old people) (Figure 4A). We compared the usage of V(D)J genes across these five high-TCE age groups. The *TRAV10*, *TRAV12-2*, *TRAV30*, *TRAV38-2/DV8*, *TRAJ35*, *TRAJ49*, *TRBV4-1*, *TRBV4-2*,

TRBV6-4, *TRBV7-3*, *TRBV12-4*, *TRBV12-5*, *TRBV19*, and *TRBJ1-3* were differentially used in the five age groups (ANOVA, FDR < 0.05) (Figure 4B). The selective usage of V(D)J genes may be the features of T cell responses in different age groups and may be associated with development and aging of human immune system. Using hierarchal clustering, five high-TCE age groups were clustered into two distinct groups based on these differentially used V(D)J genes: 1) 2 y–12 y group (young high-TCE group), and 2) 70 y–90 y group (old high-TCE group) (Figure 4B). Whereas, no significant difference was observed in the length distribution of complementarity-determining region-3 sequences in TCR α and TCR β chains among these five age groups (Figure S5B).

Moreover, we observed different degrees of clonal expansion among T cell subsets, with the highest proportions of highly clonal expanded cells (clone size > 50 cells) being CD8_Effector_GNLY cells (Figure 4C). Among eleven T cell subsets, CD8_Effector_GNLY cells contributed substantially to almost all highly clonal expanded cells in both young and old high-TCE groups (Figures 4D and S5C). Meanwhile, proportion analysis of T cell subsets in total T cells showed that there were remarkable two peaks in the variation curve of CD8_Effector_GNLY cell proportion in T cells throughout the lifespan (early peak: 2 y–12 y; late peak: 70 y–90 y) (Figure S5D). This change pattern of CD8_Effector_GNLY cell proportion in T cells were well in line with the results obtained by the clonal analysis that CD8_Effector_GNLY cells exhibited clonal expansion in both groups (Figure 4D).

To explore the functional features underlying this clonal expansion in young and old high-TCE groups, we performed DEGs analysis on these highly clonal expanded CD8_Effector_GNLY cells. Then, these DEGs were subjected to functional enrichment analysis using clusterProfiler packages. The DEGs analysis showed that the expression of interferon related genes (*IFITM2* and *ISG20*), anti-apoptosis genes *TSC22D3* (Asselin-Labat et al., 2004) and *IL7R* (Barata et al., 2019) was significantly upregulated in old high-TCE group (Figures 4E and 4F). Functional enrichment analysis showed that the genes highly expressed in old high-TCE group were enriched in “response to interferon gamma”, “response to type I interferon”, “response to virus” and “negative regulation of T cell apoptotic process” pathways (Figure 4G). By contrast, the DEGs analysis showed that the expression of cytotoxic genes (*GZMB*,

GNLY and *PRF1*) and energy metabolism-related genes (*COX5B*, *COX5A*, *NDUFAB1*, *NDUFB10*, *NDUFS6*, *ATP5F1B*, *VCP*, *UQCRC1*, *NDUF3* and *SDHA*) was significantly upregulated in young high-TCE group (Figures 4E-F). Functional enrichment analysis showed that the genes highly expressed in young high-TCE group were enriched in "cell killing", "oxidative phosphorylation", "PD-L1 expression and PD-1 checkpoint pathway in cancer", and "T cell activation" pathways (Figures 4H and S5E). These results suggested that highly expressed genes in highly clonal expanded CD8_Effector_GNLY cells from the old people were involved in the biological processes related with chronic inflammation (chronic inflammation has been implied in the aging of solid organs) (Sayed et al., 2021; Yousefzadeh *et al.*, 2021). While, the highly expressed cytotoxic genes in highly clonal expanded CD8_Effector_GNLY cells from the young people indicate that these cells possessed a powerful killing function against infected and cancerous cells.

To further illustrate the properties of highly clonal expanded CD8_Effector_GNLY cells in young and old high-TCE groups, we performed Gene Set Enrichment Analysis (GSEA) with all expressed genes. Compared with the young high-TCE group, multiple inflammatory diseases-related gene sets, such as asthma, systemic lupus erythematosus (SLE) and rheumatoid arthritis, were significantly enriched in highly clonal expanded CD8_Effector_GNLY cells in the old people (Figures 4I and 4J). By contrast, proteasome, PI3K-AKT-mTOR signaling, and IL6-JAK-STAT3 signaling pathways were significantly enriched in highly clonal expanded CD8_Effector_GNLY cells in the young people (Figure 4K). Taken together, although both groups displayed clonal expansions of CD8_Effector_GNLY cells, there were distinct functions underlying these expanded T cells between young and old high-TCE groups.

Different aging patterns in naïve CD4$^+$ T and naïve CD8$^+$ T cells

A key feature of age-related immune erosion is the loss of naïve T cells (Goronzy and Weyand, 2019; Nikolich-Žugich, 2018). Proportion analysis of T cell subsets in total T cells showed that the proportions of all the three naïve T cell subsets [including two types of naïve CD8$^+$ T cells (CD8_Naïve_LINC02446^{high} and CD8_Naïve_LINC02446^{middle}) and one type of naïve CD4$^+$ T cells (i.e.,CD4_Naïve_CCR7)] exhibited a decreased trend with age (Figure S5D), which was consistent with earlier flow cytometry results (van den Broek et al., 2018). We

further compared the dynamic expression patterns of M-DEGs in naïve $CD4^+$ T and naïve $CD8^+$ T cells. The most of upregulated M-DEGs expression in naïve $CD8^+$ T cells was sharply increased with age after 50 years old (i.e., 101 of 137 upregulated M-DEGs in CD8_Naïve_LINC02446^{high} cells, 69 of 115 upregulated M-DEGs in CD8_Naïve_LINC02446^{middle} cells) (Figures 5A, S6A and S6C), while the expression of all 99 upregulated M-DEGs in naïve $CD4^+$ T cells (i.e., CD4_Naïve_CCR7) exhibited a progressive increase with age (Figure 5B). By contrast, no difference was observed in dynamic expression patterns of downregulated M-DEGs in naïve $CD4^+$ T and naïve $CD8^+$ T cells (Figures S6B and S6C). Moreover, functional enrichment analysis showed that the upregulated M-DEGs in naïve $CD4^+$ T and naïve $CD8^+$ T cells were significantly enriched in the biological processes related to inflammatory pathways, including "positive regulation of I-kappaB kinase/NF-kappaB signaling" and "response to interferon gamma" pathways (Figure S3C). We then analyzed the expression of these two inflammatory pathways in naïve $CD4^+$ T and naïve $CD8^+$ T cells across the whole lifespan. The results showed that the expression of these two inflammatory pathways in naïve $CD8^+$ T cells (CD8_Naïve_LINC02446^{high} and CD8_Naïve_LINC02446^{middle}) were sharply upregulated after 50 years old (Figures 5C, S6D and S6E), while displayed a gradual increase in naïve $CD4^+$ T cells (CD4_Naïve_CCR7) (Figures 5C and S6D). Consistently, the variation of TCR diversity and "TCR V(D)J recombination" pathway expression was corelated with the abovementioned functional signatures in naïve $CD4^+$ T and $CD8^+$ T cells. Both the TCR diversity (Figures 5D and S6F) and "TCR V(D)J recombination" pathway expression (Figures 5E and S6G) in naïve $CD8^+$ T cells (CD8_Naïve_LINC02446^{high} and CD8_Naïve_LINC02446^{middle}) displayed an abrupt decrease after 50 years old, while those in naïve $CD4^+$ T cells (CD4_Naïve_CCR7) showed a gradual decrease with age (Figures 5D and 5E). These results revealed the distinct aging-associated transcriptional patterns and repertoire dynamics in naïve $CD4^+$ T and naïve $CD8^+$ T cells.

Activation of $CD8^+$ MAIT cells in adolescents

MAIT cells are unique innate-like T cells that bridge innate and adaptive immunity (Toubal et al., 2019). Accumulating evidence indicated that MAIT cells could rapidly produce cytokines and cytotoxic effectors via both TCR-dependent and -independent pathways (Legoux et al., 2020; Provine and Klenerman, 2020) in response to a broad

range of bacteria and viruses, such as Mycobacterium tuberculosis (TB) (Ruibal et al., 2021) and Hepatitis B virus (Dias et al., 2019). Proportion analysis of T cell subsets showed that the proportion of CD8_MAIT_SLC4A10 cells in total T cells peaked at the adolescent stage (12 y and 18y groups) during the whole lifespan (Figure S5D), which was consistent with recent flow cytometry studies (Chen et al., 2019; Nicholas A et al., 2018). In addition, clonal analysis showed that the TCR diversity of CD8_MAIT_SLC4A10 cells peaked at the adolescent stage (12 y and 18y; Figure 5F), implying that CD8_MAIT_SLC4A10 cells from adolescents might more efficiently fend off pathogen invasion compared with other age groups. To further characterize the function of CD8_MAIT_SLC4A10 cells in adolescents, we compared the gene expression profiles of CD8_MAIT_SLC4A10 cells between adolescent (including 12 y and 18y) and other age groups. The upregulated DEGs in adolescent group were enriched in the biological processes related to T cell activation, lymphocyte proliferation and cytokine production (Figures 5G and 5H), while those downregulated DEGs in adolescents were enriched in the biological processes related to neutrophil activation and cell adhesion (Figure S6H). We further depicted the functional properties of CD8_MAIT_SLC4A10 cells in adolescents by GSEA. The GSEA results showed that the TGF-beta signaling, an immunosuppressive signaling, displayed a significant down-regulation in CD8_MAIT_SLC4A10 cells in adolescents (Figure S5I). To further determine the function of upregulated DEGs in adolescents, we downloaded the $CD8^+$ MAIT cell transcriptome data of healthy individuals and latent tuberculosis infection patients (Pomaznoy et al., 2020). T cell proliferation and activation related genes including *IL7R*, *CD48*, *SIT1*, *CCR2*, *CD40LG*, *RORC* and *JAML*, highly expressed in adolescents (Figures 5I and 5J). Then, by Gene Set Variation Analysis (GSVA), we calculated the enrichment score of these T cell proliferation and activation related genes in MAIT cells derived from healthy individuals and latent tuberculosis infection patients. The GSVA enrichment score of healthy individuals was significantly higher than that of patients with latent tuberculosis infection (Figure 5K), possibly because of hypo-immune status in latent tuberculosis infection patients (Gong and Wu, 2021). According to the WHO 2021 Global Tuberculosis Report (https://www.who.int/), the TB incidence and prevalence in adolescents were lower than in adults (> 25 y), echoing our findings of enhanced MAIT cell functions in adolescents. Altogether, these results indicated that

CD8_MAIT_SLC4A10 cells in adolescents were more resistant to pathogen invasion than those in other age groups.

To further investigate the transcriptional regulatory mechanism of CD8_MAIT_SLC4A10 cells in adolescents, we predicted the core transcription factors (TFs) using SCENIC (Aibar et al., 2017) . We identified a set of TF regulons including RORC (99g), RUNX2 (66g), NR1D1 (38g), MAF (36g) and ELK3 (19g) by hierarchical clustering heatmap, whose regulon area under the curve values reached a peak in adolescents (Figures 5L, 5M and S6J). Moreover, four of five predicted TFs reached their expression peak in adolescents, including *RORC*, *RUNX2*, *NR1D1* and *ELK3* (Figure S6K). These TFs could directly regulate T-cell proliferation and activation-related genes highly expressed in adolescents, including *CD48*, *CD40LG*, *CCR2* and *IL7R* (Figure S6L). The transcription factor RORγt (encoded by *RORC*) has been reported to play crucial roles in MAIT cell maturation and effector functions by regulating IL-17 production (Guggino et al., 2017; Ivanov et al., 2006; Toubal *et al.*, 2019). Together, these results implied that these TFs (including *RORC*, *RUNX2*, *NR1D1* and *ELK3*) might exert important roles in regulating the activation of CD8_MAIT_SLC4A10 cells during adolescence.

Characteristics of the new-identified cytotoxic B cell subset (B_BCR$^+$GNLY$^+$ cells)

Immune cells with cytotoxic functions in the peripheral blood of healthy people generally included NK and cytotoxic T cells (Taylor and Cohen, 1992). However, we found a unique B cell subset (B_BCR$^+$GNLY$^+$ cells) that simultaneously expressed cytotoxic genes, such as *GNLY* and *NKG7*, *JCHAIN* and BCR genes such as immunoglobulins heavy or light chain coding genes (Figures 1C and S1B). The B_BCR$^+$GNLY$^+$ cells were enriched in children (2 y and 6 y groups) (Figures 1D and S2A). To explore the functional difference between B_BCR$^+$GNLY$^+$ cells and the other four B cell subsets, we identified the signature genes of each of the five B cell subsets. Functional enrichment analysis showed that the signature genes of B_BCR$^+$GNLY$^+$ cells were enriched in “cytolysis” and “cell killing” pathways (Figures 6A and S7A).

To further validate the existence of B_BCR$^+$GNLY$^+$ cells, we performed the immunofluorescence assay to detect the protein expression of signature genes of B_BCR$^+$GNLY$^+$ cells, including B cell markers, J chain (encoded by *JCHAIN*), two

kinds of BCRs IgM and IgA, and granulysin (cytotoxic protein, encoded by *GNLY*). Immunofluorescence assay showed that double-positive cells (J chain and granulysin, IgM and granulysin or IgA and granulysin) were identified in the peripheral blood from three healthy 2-year-old children (Figure 6B). Moreover, flow cytometry was performed to detect the percentage of double-positive cells (IgM and granulysin or IgA and granulysin) in the peripheral blood from children and adults. The results showed that double-positive cells (IgM and granulysin or IgA and granulysin) in the peripheral blood in children (2-6 years old) were significantly higher than in adults (25-50 years old) (Figures 6C and S7C). Additionally, we reanalyzed the published scRNA-seq data of PBMCs derived from three children (aged 1.9, 2.1 and 4.7 y) (Wang et al., 2021), and identified the existence of this cytotoxic B cell subset, whose proportion in PBMCs was 22.08% (1.9 y), 11.00% (2.1 y) and 8.56% (4.7 y), respectively (Figure S7D). The possible reason of neglected identification of the B_BCR$^+$GNLY$^+$ cells in this publication should have been that their cell identities were defined by comparison to reference datasets of known cell types (Wang *et al.*, 2021). Collectively, our results identified an unrecognized B_BCR$^+$GNLY$^+$ cell type that should belong to the B-cell lineage with cytotoxic functions.

To further investigate the relationship between B_BCR$^+$GNLY$^+$ cells and other B cells, we calculated the similarity scores among gene expression profiles of these B cell subsets, and evaluated the distribution of IgA, IgD, IgG, IgM and J chain in each B cell subset. The results showed that B_BCR$^+$GNLY$^+$ cells were most similar to plasma cells both in terms of the transcriptome profiles (Figure 6D) and the distribution of immunoglobulin related genes (Figures 6E). In addition, we explored the potential developmental trajectories of these B cell subsets based on transcriptional similarities using pseudotime trajectory analysis by Monocle 2 (Qiu et al., 2017). B cell subsets formed into a pseudotime trajectory that began with the naïve B cells, followed by memory B cells, B_BCR$^+$GNLY$^+$ cells and ended with plasma cells (Fig. S7D). Plasma cells are the terminally differentiated and antibody-secreting B cells (Oracki et al., 2010). Pseudotime trajectory analysis results suggested that B _BCR$^+$GNLY$^+$ cells were closely linked with plasma cells on the developmental trajectories of B cells. Moreover, to gain insights into the clonal relationship among five B cell subsets, we reconstructed paired BCR sequences from single-cell BCR sequencing data. Briefly, the detection rate of BCRs was high in all B cell subsets: 96.84% (B_Naive), 96.46% (Plasma cells), 91.26%

(B_atypical_Memory), 85.78% (B_Memory) and 68.33% (B_BCR$^+$GNLY$^+$) (Figure S7E). We calculated the average ratio of BCR clones shared between each pair of these B cell subsets. The results showed that plasma cells (22.68%) and memory B cells (10.35%) were the top two cell subsets sharing the clones with B_BCR$^+$GNLY$^+$ cells (Figure 6F). Two representatively shared BCR clones between these B cell subsets in one sample were shown in the UAMP plots (Figure 6G).

Next, to explore the transcriptional regulatory mechanism of cell differentiation in all B-cell subsets, we predicted the core TFs in each B-cell subset using SCENIC (Aibar *et al.*, 2017). Notably, consistent with the similar gene expression profiles between the B_BCR$^+$GNLY$^+$ cells and plasma cells (Figure 6D), these two B cell subsets were similar in the transcriptional regulatory mechanism, as revealed by clustering patterns in the TF regulons heatmap (Figure S7E). The top 10 TF regulons in each B cell subset were shown in Figures 6H. The UMAP plots showed that the activities of these regulons were highly specific to each B cell subset (Figure S7F). The results showed that the top 10 TFs in B_BCR$^+$GNLY$^+$ and plasma cells, respectively, were different. The top 3 TFs in B_BCR$^+$GNLY$^+$ cells, including *STAT1* (Zimmerer et al., 2007, Putz et al., 2016), *FOS* (Bernard et al., 1999) and *IRF1* (Ksienzyk et al., 2011), which have been reported to be associated with cell cytotoxicity, might be essential for the cytotoxicity of B_BCR$^+$GNLY$^+$ cells. Additionally, in the top 10 TFs of B_BCR$^+$GNLY$^+$ cells, *PRDM1* has a central role in determining and shaping the secretory arm of mature B cell differentiation and in promoting immunoglobulin synthesis (Cattoretti et al., 2005). Next, we conducted hierarchical clustering on the expression of top 10 TFs in each B cell subset using heatmap (Figure 6I). The results showed that the B_BCR$^+$GNLY$^+$ and plasma cells had similar expressions of *PRDM1*, *CREB3L2*, *XBP1*, *EZH2* and *HMGB2* (Figure 6I), which play crucial roles in plasma cells differentiation and V(D)J rearrangement (Herviou et al., 2019; Little et al., 2015; Reimold et al., 2001). Taken together, these results indicated that although B_BCR$^+$GNLY$^+$ cells resemble plasma cells in gene expression profiles, distribution of immunoglobulin isotypes, BCR clonotype and transcriptional regulatory mechanism, this new-identified B cell type has its own distinct characteristics.

Discussion

The human immune system dramatically changes across the lifespan (He and Sharpless, 2017; López-Otín and Kroemer, 2021). However, little is known about when and how the immune system changes across the entire human lifespan. Current study provided novel insights into the age-specific alteration in cell-type composition, transcription profiles, cell-cell interactions and immune repertoires of PBMCs across the entire human lifespan at single-cell resolution. Several novel contributions were made. First, we revealed that the functions of T cell subsets were the most susceptible to senescence among all PBMCs and were characterized by increased activation of NF-κB signaling and IFN-γ responses pathways, reduced telomere maintenance and energy metabolism, and significantly altered immune checkpoints in T cells within the cellular interaction networks. Second, scRNA-seq and immune repertoire sequencing disclosed unique transcriptional and clonal expansion features of T cell subsets [CD8_Effector_GNLY cells exhibiting a bimodal distribution across the human lifespan, with an early peak in 2 y–12 y age groups and a late peak in 70 y–90 age groups; naïve CD4+ T and naïve CD8+ T cells exhibiting distinct aging-associated transcriptional patterns and repertoire dynamics; CD8_MAIT_SLC4A10 cells exhibiting peak abundance and clonal diversity in adolescents (12 y and 18 y groups)]. Furthermore, we identified a unique cytotoxic B cell subset (B_BCR^{+}GNLY^{+} cells) that simultaneously expressed cytotoxic genes and BCR genes, and was enriched in children compared to other age groups. Overall, our dataset provides temporal resolution of immune cell transcriptomes across the entire human lifespan, and facilitates in-depth exploration of the immune system at specific ages.

The immune system is immature at birth and gradually develops under exposition of postnatal environmental stimuli (Brodin and Davis, 2017). In particular, the early childhood is a critical period for immune development (Hill et al., 2020). The maternally-derived antibodies taper off after birth (Kollmann et al., 2017; Saso and Kampmann, 2017), while the self-derived immune mediators such as T cells from thymus, B cells and innate immune cells from bone marrow augment over time (Goronzy et al., 2019). Notably, we found a remarkable feature in this critical developmental period: a unique cytotoxic B cell subset (B_BCR^{+}GNLY^{+}) expanded significantly in 2 and 6-year-old age groups. These B_BCR^{+}GNLY^{+} cells were characterized by the enriched expression of genes with cytotoxic activity including *GNLY*, *GZMB* and *NKG7*, and the suppression of general B cell marker genes *CD19*,

CD79A and *MS4A1*. In addition, trajectory analysis and transcriptional profiles revealed close resemblance of these B_BCR$^+$GNLY$^+$ cells to the plasma cells. Current findings revealed that the number of B_BCR$^+$GNLY$^+$ cells was extremely low in neonates and 1-year old children, with an obvious increase in 2 to 6-year-old children followed by gradual decrease afterwards. Similarly, previous work reported B cells expressing cytotoxic genes including GZMB$^+$ B cells and plasma cells (GZMB$^+$ PCs). The GZMB$^+$ B cells were characterized by CD19 and granzyme B expression, but not granzyme A or granulysin expression, and were detected in the peripheral blood of patients with SLE (Hagn et al., 2010) and remitting multiple sclerosis (Boldrini et al., 2022), and lymph nodes of patients with breast cancer (Arabpour et al., 2019) through flow cytometry assays. More recently, granzyme B-expressing plasma cells (GZMB$^+$ PCs) were detected in human tonsils using immunofluorescence technology (Xu et al., 2014). These GZMB$^+$ PCs expressed BCR light chains and cytotoxic protein granzyme B, while the expression of the cytotoxic protein granulysin and CD19 was hardly detected in them. However, these GZMB$^+$ B cells detected by flow cytometry and GZMB$^+$ PCs discovered by immunofluorescence technology were significantly different from the B_BCR$^+$GNLY$^+$ cells discovered through single-cell analysis in healthy population in the present study. First, B_BCR$^+$GNLY$^+$ cells were characterized by high expression of BCR genes, *JCHAIN* and *GNLY* (encoding granulysin) but negative for *CD19*. By contrast, GZMB$^+$ B cells were CD19-positive but granulysin-negative, and GZMB$^+$ PCs were BCR light chains-positive but CD19- and granulysin-negative. Second, the expression of *GZMB* (encoding granzyme B) in our B_BCR$^+$GNLY$^+$ cells was lower than the expression of *GNLY* ($p < 2.2\times10^{-16}$) (Figure S7H). These findings indicated that the B_BCR$^+$GNLY$^+$ cells discovered through the present single-cell analysis might be a unique cluster of B cells in healthy people, particularly in children. However, the expansion mechanisms of B_BCR$^+$GNLY$^+$ cells in childhood remain undetermined. A previous flow cytometry study reported that tick-borne encephalitis vaccination in healthy subjects could induce cytotoxic phenotypes in B cells with expression granzyme B (Hagn et al., 2009). Thus, we retrospectively reviewed the vaccinations recorded in these children recruited in our study. Interestingly, we noticed that 2 to 6-year-old children routinely received more intense vaccination (e.g., booster vaccinations, immunization with multivalent vaccines and live attenuated vaccines) than other age groups (Table S7).

These finding implied that vaccination might be one of the mechanisms underlying B_BCR$^+$GNLY$^+$ cell expansion. However, the mechanisms underlying B_BCR$^+$GNLY$^+$ cell expansion merits further exploration. Further studies are warranted to explore the function of these new-identified B cells in health and disease states.

Moving from childhood to adolescence, our results revealed that peripheral immunity presents another scene. The immune system undergoes dramatic changes beyond adolescence and sexual maturity (Vallejo, 2011). We found that the immune maturation during adolescence was characterized by the expansion of MAIT cells with higher diversity of TCR compared with other age groups. A previous study demonstrated that commensal bacteria mediated the MAIT cell maturation and its recirculation into thymus (Legoux et al., 2019), hinting a critical role of MAIT cells in TCR and immune system maturation. During adolescence, the dramatic changes of hormones lead to the microenvironment alteration of mucosal membrane and skin, and the variation of commensal bacterial species and numbers (Kundu et al., 2017), which regulate the proliferation of MAIT cells. MAIT cells contribute to host defense against infection (Toubal *et al.*, 2019) including tuberculosis (Malka-Ruimy et al., 2019) and influenza (van Wilgenburg et al., 2018). According to the WHO 2021 Global Tuberculosis Report, the TB incidence and prevalence in adolescents were lower than in adults (> 25 y). However, until now, understanding of peripheral circulating MAIT cells in healthy individuals across the life course is still lacking. To our knowledge, the present findings for the first time delineated functional significance of MAIT cells in adolescence through the transcriptional profiling of this cell type across the lifespan.

The immune function declines with aging, termed “immunosenescence” (Nikolich-Zugich, 2018). However, different types of immune cells display differential susceptibility to aging. Our results revealed that the functions of T cell subsets were the most susceptible to senescence among all PBMCs, which were in a good agreement with the mice aging study by Matthew J et al. (Yousefzadeh *et al.*, 2021). Moreover, functional enrichment analysis of cell type-specific age-related genes suggested that the elevation of pro-inflammatory signaling and the attenuation of signaling related to telomere maintenance and energy metabolism cooperatively contributed to immunosenescence and inflammaging. Furthermore, we discovered that naïve CD4$^+$ and naïve CD8$^+$ T cells exhibited distinct aging patterns. Both naïve

$CD4^+$ T and naïve $CD8^+$ T cells arise from the thymus, but the behavior of naive T cells is different in $CD4^+$ and $CD8^+$ compartments. The expression of age-related genes and TCR diversity in naive $CD8^+$ T cells exhibited sharp variations after 50 years old. By contrast, naïve $CD4^+$ T cells showed progressive changes in expression of age-related genes and TCR diversity across the lifespan. Although the collapse of TCR diversity in naïve T cells has been considered a strong predictor of poor health in old people (Qi et al., 2014), the variation of TCR diversity with age in naïve $CD4^+$ T and naïve $CD8^+$ T cells is not very clear. Even though previous studies have reported that TCR diversity of naïve T cells displayed age-associated decline by bulk RNA sequencing (Britanova et al., 2014; Britanova et al., 2016), we depicted the different dynamics of TCR diversity across the lifespan in naïve $CD4^+$ T and naïve $CD8^+$ T cells for the first time. These results implied that the underlying mechanisms of aging in naïve $CD4^+$ T and naïve $CD8^+$ T cells are not identical, and provided valuable insights into the dysfunction of naïve T cells with age.

The proportion of CD8_Effector_GNLY cells exhibited a bimodal distribution across the human lifespan, with an early peak in 2 y–12 y age groups and a late peak in 70 y–90 age groups (Figure S5D). Consistently, clonal analysis indicated that CD8_Effector_GNLY cells exhibited clonal expansion in these two groups. Whereas, molecular features underlying the clonal expansion of effector T cells in young and old were different. The highly clonal expanded CD8_Effector_GNLY cells in young people showed high expression of cytotoxic genes, indicating that it possesses the powerful killing function against infected and cancerous cells. By contrast, genes highly expressed in highly clonal expanded CD8_Effector_GNLY cells from the old people were enriched in the biological processes related with chronic inflammation, which was associated with onset of inflammatory diseases, such as asthma (Leuti et al., 2020) and rheumatoid arthritis (Chen et al., 2019). Considering the variation of physiology between young and old people, and environmental stimuli in childhood, such as frequent vaccinations, our results may reflect distinct mechanisms between age-related clonal expansion in old people (e.g., age-related) and clonal expansion in young people (e.g., external stimulus-related). Not only that, our results demonstrated that age played a crucial role in modulating the expression of immune checkpoints on peripheral blood lymphocytes, particularly in T cell subsets. Meanwhile, the dynamics of immune checkpoint including *CD27*, *ICOS*, *PDCD1* and *CLTA4* echoed the clonal expansion of effector T cells in young and old people. The CD27 (van Gisbergen et al.,

2011) and ICOS (Dong et al., 2001) generally participate in T cells activation and survival, while the PD-1 (encoded by *PDCD1*) and CTLA-4 are the critical molecules in T cell-response regulation (Wei et al., 2017). Although PD-1 and CTLA-4 are both negative regulator of T cell immune response, the CTLA-4 inhibits T cell activity at the beginning of immune response (Bagchi et al., 2021). While, the PD-1 suppresses T cell activity and amplification in the late stage of immune response (Bagchi et al., 2021; Buchbinder and Desai, 2016). The high expression of *CD27*, *ICOS*, *PDCD1* in children, and *CLTA4* in old people reflects distinct features of T cell response at different ages.

In conclusion, through our study, significant progresses have been made in understanding the age-related cellular and molecular variations in PBMCs during the development, maturation and aging of the immune system with greater details that were not appreciated previously. This dataset provides a valuable resource on circulating immune cells of different ages (available at https://pu-lab.sjtu.edu.cn/shiny for quick browsing), allowing for systematic elucidation of age-specific hallmarks of circulating immune cells across the lifespan, and serving as a reference for PBMCs dynamics during development, maturation and aging.

ACKNOWLEDGMENTS

We thank the volunteers from the Pudong Natural Population-based Cohort for their contribution to the study. This work was supported by grants from the National Science Fund for Distinguished Young Scholars (81625002), the National Natural Science Foundation of China (81930007, 81470389, 81500221, 82071852, 82003013 and 81800307), the National High Level Talents Special Support Plan, the Shanghai Outstanding Academic Leaders Program (18XD1402400), Innovative research team of high-level local universities in Shanghai (SSMU-ZDCX20180200, SSMU-ZDCX20212100, SSMU-ZDCX20210601), Shanghai Municipal Education Commission Gao Feng Clinical Medicine Grant Support (20152209), Sponsored by Shanghai Sailing Program (21YF1445000), Shanghai Municipal Health Commission (20184Y0371) and the National Key R&D program of China (2021YFC2502300).

AUTHOR CONTRIBUTIONS

J.P. conceived the study. R.Y.T., Z.L., J.F.S., A.C.Y., Y.P.G. and L.Y.S. recruited volunteers. R.Y.T., J.F.S., Y.G and Y.F.C. collected samples from donors. R.Y.T., J.F.S. and Y.Y. L. completed the experiments. R.H.L., R.Y.T., Y.C.Z, L.J.L., J.M.Z. and L.W. finished the data analysis. R.Y.T. and R.H.L drafted the initial version. Y.C.Z., Y.F.C. and L.J.L. proofed the manuscript. J.P., L.W., F.W., W.L, B.L, L.C, H.Y.C, and H.L.W provided guidance and constructive comments for the final version. All the authors read and approved the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

STAR METHODS

Participants and ethics

All human samples were obtained in accordance with protocols approved by the Institutional Review Board at Shanghai Jiao Tong University. All donors were enrolled with approval by the Ethics Committee of Ren Ji Hospital (KY2019-136), School of Medicine, Shanghai Jiao Tong University. After meeting inclusion criteria in accordance with the WHO classification and providing written informed consent in accordance with the criteria set by the Declaration of Helsinki (2013) (World Medical, 2013), 45 healthy volunteers from neonates to the elderly were recruited in the Pudong Natural Population-based Cohort, Shanghai, China. To avoid possible interference by other confounding factors, all participants were free of clinical symptoms including fever, cough, headache, sore throat, malaise, loss of smell, runny nose, abdominal pain, diarrhea, joint pain, wheezing or dyspnea and free of vaccination within three months before sampling. Meanwhile, the participants with hypertension, diabetes mellitus, autoimmune disease, cancer, renal or hepatic dysfunction, hematological disorders, pregnancy, lactation, steroid usage and smoking were excluded. They were assigned to 10 groups including group UCB standing for the neonates (umbilical cord blood, UCB); group 1 y, 2 y, 6 y standing for children aged 1, 2 and 6 years old respectively; group 12 y, 18 y standing for adolescents aged 12 and 18 years old; group 30 y, 50 y standing for adults aged 30 and 50 years old; group 70 y standing for septuagenarian aged 70 years old and group 90 y standing for nonagenarian aged ≥ 90 years old. Each group consisted of three to five individuals that included both males and females. Sample details are shown Table S1.

Sample collection, preparation, and storage

Blood samples were collected from the peripheral vein of each donor and processed within 2 hours of collection except umbilical cord blood (UCB) obtained from placental cords. PBMCs were separated using Vacutainer CPT tubes (BD) according to the manufacture's protocols (Corkum et al., 2015). The viability of PBMCs in each sample was confirmed to be >90% by Trypan Blue staining. Then, PBMCs were immediately used for scRNA library construction and sequencing or were cryopreserved at □80°C in 10% dimethylsulfoxide in fetal bovine serum.

ScRNA-seq library construction and sequencing

To produce single-cell gel beads in emulsion (GEMs), cell suspensions were quantified with 400-600 living cells per microliter determined by a Count Star. Then, a Chromium single cell controller (10× Genomics) was used to prepare cell suspensions as GEMs with a Single Cell 5□ Library and Gel Bead kit (10× Genomics, 1000006) and Chromium Single Cell A Chip kit (10× Genomics, 120236) following the manufacturer's protocols. About 15,000 cells were added to each channel and the recovered target cells was estimated to be about 8,000 cells. Single-cell RNA libraries were constructed the Chromium Single Cell 5□ v1.1 Reagent (10× Genomics, Human T Cell:1000005; Human B Cell: 1000016) and a unique sample index was included in each sequencing library. Finally, sequencing was performed with a sequencing depth of at least 100,000 reads per cell with a pair-end 150 bp (PE150) reading strategy using the Illumina platform (CapitalBio Technology, Beijing, China).

Single-cell RNA sequencing data processing

All single-cell sequencing data were mapped to the human reference genome GRCh38 (10× Genomics, Version 3.0.0.) using STAR (version 2.5.1b). Unique molecular identifier (UMI) matrices were established for each sample using the CellRanger (10× Genomics, Version 3.0.2) pipeline and subjected to the standard processing pipeline that included quality control, normalization, dimension reduction and clustering by Seurat (Version 3.2.0). Specifically, cells with less than 400 expressed genes, or >10% mitochondrial content were excluded. Genes in the following analysis were confirmed with at least one feature count over three cells. To remove potential doublets, cells detected genes above 2500 are filtered out. Additionally, we applied Scrublet package (Wolock et al., 2019) in python to identify potential doublets. The doublet score for each single cell and the threshold based on the bimodal distribution was calculated using default parameters. The expected doublet rate was set to be 0.075, and cells predicted to be doublets or with doubletScore larger than 0.25 were filtered. Based on these measures, gene expression profiles were normalized for each cell under SCT-transform using the Seurat package. Subsequently, the 'FindVariableFeatures' function facilitated the selection of variable genes following the default operation. The gene expression matrixes were integrated together using "FindIntegrationAnchors" and the "IntegrateData" functions was

adopted to correct the batch effects among samples by projecting all genes onto a low dimensional subspace based on canonical correlation analysis. The integrated data were saved for following analysis. Furthermore, a shared nearest neighbor graph was constructed based on the Euclidean distance in the low dimensional subspace spanned by the selected significant principal components. The "FindClusters" function was used to cluster these cells at an appropriate resolution. Finally, the "RunUMAP" function facilitated visualization of all these cells on the plane based on the two-dimensional uniform manifold approximation and projection (UMAP). Following the above pipeline, high-quality scRNA-seq data of ~380k cells derived from 45 samples were remained.

Cell types annotation and cluster markers identification

The "FindAllMarkers" function was used to find markers for each cluster. Clusters were then classified and annotated based on the signature gene expression of each cell subset.

All assigned cell subsets were based on the entire cluster instead of individual cells. Moreover, the variety within the same cluster was rectified by dominant prediction for the whole cluster. A full list of canonical and signature genes for each subset is deposited in Table S2.

Identification of DEGs in peripheral immune cells

To investigate the effect of age on peripheral immune cells, we identified DEGs among ten age groups for each PBMC subset via following steps: 1) For each cell subset in one sample, the synthetic bulk expression value for each gene was generated by dividing the gene-specific UMI count with the total number of UMIs in the cell and multiplying by 10000 ("CP10K"). 2) Ribosome and mitochondrial genes were removed in all the samples. 3) The genes with expression more than 1 (CP10K > 1) in at least three samples were kept and analyzed by one-way analysis of variance (one-way ANOVA). 4) The genes with p value < 0.001 were considered to be DEGs. Age-related genes (monotonically correlated with age; M-DEGs) were further determined using Spearman Rho analysis (Rho ≥ 0.5, p < 0.01) based on all DEGs. CP10K were defined as:

$$\mathrm{CP10K}_{i,k} = \frac{\sum_j e_{i,j,k}}{\sum_i \sum_j e_{i,j,k}} \cdot 10^4$$

e is the number of UMIs of i gene defined by the unique molecular barcode of j cell in group k.

Next, to further investigate the expression profile of CD8_MAIT_SLC4A10 cells of adolescents, Seurat's FindMarkers function (parameters: only.pos = F, min.pct = 0.25 logfc.threshold = 0.25, test = 'wilcox') was run comparing adolescents against non-adolescents. Meanwhile, the same parameter settings were used to identify the DEGs in clonal expanded CD8_Effector_GNLY cells between young and old group with Seurat's FindMarkers function. The DEGs were filtered using a minimum $\log_2$ (fold change) of 0.25 and a adjust p value of 0.05. The volcano plots of DEGs were performed by ggplot2 in R (version 3.6.0).

Functional enrichment analysis and GSVA scoring

Functional enrichment analysis of DEGs were conducted using R package clusterProfiler (Yu et al., 2012) following the default parameters. R package REVIGO was applied to summarize Gene Ontology terms based on semantic similarity measures (Li et al., 2017; Supek et al., 2011). Gene set enrichment analysis (GSEA) was performed using the Java software GSEA 4.1.0 (Subramanian et al., 2005) based on the dataset 'h.all.v7.1.sumbols.gmt [Hallmarks]', 'c2.cp.kegg.v7.1.symbols.gmt [Curated]' and 'c5.go.bp.v7.4.symbols.gmt', collected from the Molecular Signatures Database (MSigDB) (Liberzon et al., 2011). The expression of signature genes in MAIT cells derived from healthy control (HC) and latent tuberculosis infection (LTBI) patients were quantified using GSVA package (Hänzelmann et al., 2013). The dataset on healthy individuals and latent tuberculosis infection patients came from GEO database (GSE132932) (Pomaznoy et al., 2020). Individuals GSVA scores were plotted as a scatter dot plot with the ggplots2 package. Next, a comparison of the GSVA enrichment score of healthy individuals and latent tuberculosis infection patients was performed using a two-sided unpaired Wilcoxon test, and a p-value < 0.05 was considered statistically significant.

Cell-cell interaction analysis

Manually curated human ligand-receptor interaction pairs were downloaded from the CellTalkDB database (Shao et al., 2021). For each sample, the interaction strength was calculated between cell subset i and cell subset j. For each ligand-receptor pair LR, the ligand and receptor are L and R, respectively. For cell subset i and cell subset

j, their interaction strength was defined as:

$$S_{i,j} = \sum_{k=1}^{N} L_{k,i} * R_{k,j} + \sum_{k=1}^{N} R_{k,i} * L_{k,j}$$

$L_{k,i}$ and $R_{k,i}$ represent the total UMI count of ligand L_k and receptor R_k expressed in cell subset *i*, and $L_{k,j}$ and $R_{k,j}$ represent the total UMI count of ligand L_k and receptor R_k expressed in cell subset *j* and *N* is the total number of LR pairs. For each sample, the relative cell-cell interaction probabilities between cell subset ***i*** and cell subset ***j*** was defined as:

$$P_{i,j} = \frac{S_{i,j}}{\sum_{\substack{1 \le i \le (M-1) \\ i<j \le M}} S_{i,j}}$$

Then, ANOVA was used to calculate the significance p values of $P_{i,j}$ among samples from different age groups. CCIs with Bonferroni-adjusted p value < 0.05 were selected as CCIs that varied significantly with age. ANOVA was used to identify ligands and receptors differentially expressed among different age groups ($p < 0.001$). Addition, experimentally validated immune checkpoint targets in cancer immunotherapy were downloaded from CKTTD database (Zhang et al., 2020), which were analyzed for their expression dynamics in different cell subsets across the human lifespan.

TCR and BCR repertoire sequencing and analysis

In addition to the single-cell transcriptome data, single-cell TCR and BCR repertoire libraries of all samples (10× Genomics) were prepared and profiled with paired-end sequencing (2□×□150□bp) on the Illumina HiSeq2500 platform (Illumina, San Diego, CA). The CellRanger V(D)J pipeline (10× Genomics, Version 3.0.2) was adopted to mark the TCR and BCR expression and clonal types. The TCR and BCR repertoire sequencing data were filtered with the following criteria: productive is “True”; high_confidence is “True”; umis ≥1; raw_consensus_id is not “None”. To identify the TCR clonotype for each T cell, only cells with at least one TCR α chain and one TCR β chain were used. For a given T cell, if there were two or more TCR α or TCR β chains assembled, the highest expressed (UMI or reads) TCR α or TCR β chains were regarded as the dominant TCR α or TCR β chain in the cell. Each unique TCR α(s)-TCR β(s) pair (CDR3 nucleotide sequences and rearranged VJ genes included) was defined as a clonotype. BCR clonotypes were identified in a similar

manner to TCR clonotypes. Only cells with at least one productive heavy chain (IGH) and one productive light chain (IGL or IGK) were kept for further analysis. Each unique IGH-IGL/IGK pair was defined as a clonotype. For a given B cell, if there were two or more IGH or IGL/IGK assembled, the highest expressed (UMI or reads) IGH or IGL/IGK chain was defined as the dominant IGH or IGL/IGK chain in the cell.

For CDR3 amino acid length analysis, filtered TCR α and β chains were separately calculated with "nchar", "as.dat.frame", "unlist", "table", "sum" and "round" functions. For the TCR clonal status analysis, "as.dat.frame", "unlist" and "table" functions were used to calculate clonotype counts of filtered cells. Cells expressing the same clonotype were counted. The counts were classified into eight classes to identify the cellular clonal expansion status, i.e., >100, 51 to 100, 31 to 50, 21 to 30, 11 to 20, 6 to 10, 2 to 5, and a unique clone without expansion.

Moreover, based on the barcode mapping, clonal type information from CellRanger V(D)J was integrated into the meta data of the Seurat object for further analysis. To estimate TCR diversity, Shannon's entropy was measured using the function 'H' of the philentropy package as follows:

$$H = -\sum_{x} p(x) * \log_2[p(x)]$$

Here, the TCR among T with TCR identified is represented as p(x). Changes in clone diversity with age were fitted by loess regression. The Spearman correlation coefficient and significance were also calculated from the "cor.test" function in stats package. The p value<0.05 was considered statistically significant.

Definition of cell state scores

Cellular modular scores were used to assess the cell state during the whole lifespan. The "AddModuleScore" function in Seurat was used to calculate the modular score of pathways with default settings. Response to interferon gamma (GO:0034341), Positive regulation of I-kappaB kinase/NF-kappaB (GO:0043123), V(D)J recombination (GO:0033151) were used to define the IFN-γ response, NK-κB signaling pathway, TCR V(D)J recombination. The cell scores for each pathway were calculated based on the average expression of genes from the abovementioned gene set. Expression signal changes with age were fitted by generalized loess regression. The Spearman correlation coefficient and significance were also calculated from the 'cor.test' function in stats package. A p-value < 0.05 was considered statistically

significant.

Comparison of immune cell subsets transcriptomes

MetaNeighbor analysis (Crow et al., 2018) was performed through neighbor voting based on the spearman correlation among immune cell subset pseudo-cells to compare immune cell subsets transcriptomes. Pseudo-cell was made by aggregating 10 cells randomly selected from the same cell type derived from the same group to reduce the effects of outliers and noise (Han *et al.*, 2020). Then, mean ROC curve (receiver operating characteristic curve) scores were acquired from MetaNeighbor. The pheatmap package in R was adopted to visualize transcriptome similarity.

Transcription factor analysis

After arranging the input by the gene expression matrix, the SCENIC (Aibar *et al.*, 2017) package in R facilitated evaluation of transcription factors among immune cell subsets following the default parameters. The pheatmap and ggplot2 packages in R were adopted to visualize the expression profile of transcription factors.

Immunofluorescence staining

PBMCs were separated using vacutainer CPT tubes (BD). The PBMCs were gently resuspended with 1-2ml 1×PBS to 1×10^6 cells /ml. Then, the cells were prepared on SuperFrost plus slides using a Shandon Cytospin 4 (Fisher Scientific) (Bender et al., 2020). After fixation in 4% paraformaldehyde, the PBMCs were permeabilized with 0.1% Triton X□100 and blocked in PBS with 10% BSA at room temperature. To detect the expression of IgM, IgA, J chain and granulysin (encoded by GNLY), fixed PBMCs were incubated with primary antibodies against IgM (1:500, Thermo), IgA (1:500, Santa Cruz), J chain (1:500, Thermo), or granulysin (1:500, Abcam) overnight at 4 °C, followed by incubation with an Alexa Fluor 488-conjugated secondary antibody (1:200, Thermo) or Alexa Fluor 555-conjugated secondary antibody (1:200, Thermo). Counterstaining was performed with 4’-6-diamidino-2-phenylindole (DAPI). The stained PBMCs were viewed by fluorescence microscopy (TCS SP8, Leica) (Xu et al., 2019).

Trajectories and pseudotime analyses

Monocle 2 (version 2.6.1) (Qiu *et al.*, 2017) was used to estimate a pseudo temporal path of B cells differentiation. The "differentialGeneTest" function in Monocle2 was adopted to identify the "ordering_genes" following with default following with default. The DDRTree algorithm was used for dimensionality reduction.

Statistical analysis

Statistical analyses were performed using the R software package (version 3.6.0) and involved Shapiro-Wilk test, F-test, Levene's test, one-way ANOVA, two-sided unpaired Wilcoxon test and Spearman's correlation test. Data normality was determined by the Shapiro-Wilk test. The F-test was used to examine homogeneity of variance for two groups comparisons. The Levene's test was used to examine homogeneity of variance for one-way ANOVA. We performed one-way ANOVA on cell subset proportions, gene expression and cell-cell interaction among ten age groups. When comparing node degree and betweenness centrality of immune checkpoint receptors in CCI network, GSVA enrichment score and two-sided unpaired Wilcoxon tests were used. Association between age and variables including cell subset proportions, gene expression and diversity of TCR was tested using a Spearman's correlation test. The p value significance thresholds used for each result are described in the corresponding methods and figure legends.

Data availability

The raw sequence data reported in this study will been deposited in the Genome Sequence Archive of the Beijing Institute of Genomics (BIG) Data Center, BIG, Chinese Academy of Sciences.

REFERENCES

Ahmed, H., Howton, T.C., Sun, Y., Weinberger, N., Belkhadir, Y., and Mukhtar, M.S. (2018). Network biology discovers pathogen contact points in host protein-protein interactomes. Nature communications *9*, 2312. 10.1038/s41467-018-04632-8.

Ahmed, R., Omidian, Z., Giwa, A., Cornwell, B., Majety, N., Bell, D.R., Lee, S., Zhang, H., Michels, A., Desiderio, S., et al. (2019). A Public BCR Present in a Unique Dual-Receptor-Expressing Lymphocyte from Type 1 Diabetes Patients Encodes a Potent T Cell Autoantigen. Cell *177*, 1583-1599.e1516. 10.1016/j.cell.2019.05.007.

Aibar, S., Gonzalez-Blas, C.B., Moerman, T., Huynh-Thu, V.A., Imrichova, H., Hulselmans, G., Rambow, F., Marine, J.C., Geurts, P., Aerts, J., et al. (2017). SCENIC: single-cell regulatory network inference and clustering. Nature methods *14*, 1083-1086. 10.1038/nmeth.4463.

Alpert, A., Pickman, Y., Leipold, M., Rosenberg-Hasson, Y., Ji, X., Gaujoux, R., Rabani, H., Starosvetsky, E., Kveler, K., Schaffert, S., et al. (2019). A clinically meaningful metric of immune age derived from high-dimensional longitudinal monitoring. Nat Med *25*, 487-495. 10.1038/s41591-019-0381-y.

Arabpour, M., Rasolmali, R., Talei, A., Mehdipour, F., and Ghaderi, A. (2019). Granzyme B production by activated B cells derived from breast cancer-draining lymph nodes. Molecular immunology *114*, 172-178. 10.1016/j.molimm.2019.07.019.

Asselin-Labat, M.L., David, M., Biola-Vidamment, A., Lecoeuche, D., Zennaro, M.C., Bertoglio, J., and Pallardy, M. (2004). GILZ, a new target for the transcription factor FoxO3, protects T lymphocytes from interleukin-2 withdrawal-induced apoptosis. Blood *104*, 215-223. 10.1182/blood-2003-12-4295.

Bagchi, S., Yuan, R., and Engleman, E.G. (2021). Immune Checkpoint Inhibitors for the Treatment of Cancer: Clinical Impact and Mechanisms of Response and Resistance. Annual review of pathology *16*, 223-249. 10.1146/annurev-pathol-042020-042741.

Barata, J.T., Durum, S.K., and Seddon, B. (2019). Flip the coin: IL-7 and IL-7R in health and disease. Nat Immunol *20*, 1584-1593. 10.1038/s41590-019-0479-x.

Bender, C., Rodriguez-Calvo, T., Amirian, N., Coppieters, K.T., and von Herrath, M.G. (2020). The healthy exocrine pancreas contains preproinsulin-specific CD8 T cells that attack islets in type 1 diabetes. Science advances *6*. 10.1126/sciadv.abc5586.

Bernard, K., Cambiaggi, A., Guia, S., Bertucci, F., Granjeaud, S., Tagett, R., N'Guyen, C., Jordan, B.R., and Vivier, E. (1999). Engagement of natural cytotoxicity programs regulates AP-1 expression in the NKL human NK cell line. J Immunol *162*, 4062-

4068.
Boldrini, V.O., Marques, A.M., Quintiliano, R.P.S., Moraes, A.S., Stella, C., Longhini, A.L.F., Santos, I., Andrade, M., Ferrari, B., Damasceno, A., et al. (2022). Cytotoxic B Cells in Relapsing-Remitting Multiple Sclerosis Patients. Frontiers in immunology *13*, 750660. 10.3389/fimmu.2022.750660.
Britanova, O.V., Putintseva, E.V., Shugay, M., Merzlyak, E.M., Turchaninova, M.A., Staroverov, D.B., Bolotin, D.A., Lukyanov, S., Bogdanova, E.A., Mamedov, I.Z., et al. (2014). Age-related decrease in TCR repertoire diversity measured with deep and normalized sequence profiling. J Immunol *192*, 2689-2698. 10.4049/jimmunol.1302064.
Britanova, O.V., Shugay, M., Merzlyak, E.M., Staroverov, D.B., Putintseva, E.V., Turchaninova, M.A., Mamedov, I.Z., Pogorelyy, M.V., Bolotin, D.A., Izraelson, M., et al. (2016). Dynamics of Individual T Cell Repertoires: From Cord Blood to Centenarians. J Immunol *196*, 5005-5013. 10.4049/jimmunol.1600005.
Brodin, P., and Davis, M.M. (2017). Human immune system variation. Nature reviews. Immunology *17*, 21-29. 10.1038/nri.2016.125.
Buchan, S.L., Rogel, A., and Al-Shamkhani, A. (2018). The immunobiology of CD27 and OX40 and their potential as targets for cancer immunotherapy. Blood *131*, 39-48. 10.1182/blood-2017-07-741025.
Buchbinder, E.I., and Desai, A. (2016). CTLA-4 and PD-1 Pathways: Similarities, Differences, and Implications of Their Inhibition. American journal of clinical oncology *39*, 98-106. 10.1097/coc.0000000000000239.
Cattoretti, G., Angelin-Duclos, C., Shaknovich, R., Zhou, H., Wang, D., and Alobeid, B. (2005). PRDM1/Blimp-1 is expressed in human B-lymphocytes committed to the plasma cell lineage. The Journal of pathology *206*, 76-86. 10.1002/path.1752.
Chen, Z., Bozec, A., Ramming, A., and Schett, G. (2019). Anti-inflammatory and immune-regulatory cytokines in rheumatoid arthritis. Nature reviews. Rheumatology *15*, 9-17. 10.1038/s41584-018-0109-2.
Corkum, C.P., Ings, D.P., Burgess, C., Karwowska, S., Kroll, W., and Michalak, T.I. (2015). Immune cell subsets and their gene expression profiles from human PBMC isolated by Vacutainer Cell Preparation Tube (CPT) and standard density gradient. BMC Immunol *16*, 48. 10.1186/s12865-015-0113-0.
Crow, M., Paul, A., Ballouz, S., Huang, Z.J., and Gillis, J. (2018). Characterizing the replicability of cell types defined by single cell RNA-sequencing data using MetaNeighbor. Nature communications *9*, 884. 10.1038/s41467-018-03282-0.
Dias, J., Hengst, J., Parrot, T., Leeansyah, E., Lunemann, S., Malone, D.F.G., Hardtke,

S., Strauss, O., Zimmer, C.L., Berglin, L., et al. (2019). Chronic hepatitis delta virus infection leads to functional impairment and severe loss of MAIT cells. Journal of hepatology *71*, 301-312. 10.1016/j.jhep.2019.04.009.

Dong, C., Juedes, A.E., Temann, U.A., Shresta, S., Allison, J.P., Ruddle, N.H., and Flavell, R.A. (2001). ICOS co-stimulatory receptor is essential for T-cell activation and function. Nature *409*, 97-101. 10.1038/35051100.

Efremova, M., Vento-Tormo, M., Teichmann, S.A., and Vento-Tormo, R. (2020). CellPhoneDB: inferring cell-cell communication from combined expression of multi-subunit ligand-receptor complexes. Nature protocols *15*, 1484-1506. 10.1038/s41596-020-0292-x.

Fagnoni, F.F., Vescovini, R., Passeri, G., Bologna, G., Pedrazzoni, M., Lavagetto, G., Casti, A., Franceschi, C., Passeri, M., and Sansoni, P. (2000). Shortage of circulating naive CD8(+) T cells provides new insights on immunodeficiency in aging. Blood *95*, 2860-2868.

Gong, W., and Wu, X. (2021). Differential Diagnosis of Latent Tuberculosis Infection and Active Tuberculosis: A Key to a Successful Tuberculosis Control Strategy. Frontiers in microbiology *12*, 745592. 10.3389/fmicb.2021.745592.

Goronzy, J.J., Gustafson, C.E., and Weyand, C.M. (2019). 38 - Immune Deficiencies at the Extremes of Age. In Clinical Immunology (Fifth Edition), R.R. Rich, T.A. Fleisher, W.T. Shearer, H.W. Schroeder, A.J. Frew, and C.M. Weyand, eds. (Elsevier), pp. 535-543.e531. https://doi.org/10.1016/B978-0-7020-6896-6.00038-7.

Goronzy, J.J., and Weyand, C.M. (2019). Mechanisms underlying T cell ageing. Nature reviews. Immunology *19*, 573-583. 10.1038/s41577-019-0180-1.

Guggino, G., Di Liberto, D., Lo Pizzo, M., Saieva, L., Alessandro, R., Dieli, F., Triolo, G., and Ciccia, F. (2017). IL-17 polarization of MAIT cells is derived from the activation of two different pathways. Eur J Immunol *47*, 2002-2003. 10.1002/eji.201747140.

Hagn, M., Ebel, V., Sontheimer, K., Schwesinger, E., Lunov, O., Beyer, T., Fabricius, D., Barth, T., Viardot, A., Stilgenbauer, S., et al. (2010). CD5+ B cells from individuals with systemic lupus erythematosus express granzyme B. European journal of immunology *40*, 2060-2069. 10.1002/eji.200940113.

Hagn, M., Schwesinger, E., Ebel, V., Sontheimer, K., Maier, J., Beyer, T., Syrovets, T., Laumonnier, Y., Fabricius, D., Simmet, T., and Jahrsdörfer, B. (2009). Human B cells secrete granzyme B when recognizing viral antigens in the context of the acute phase cytokine IL-21. J Immunol *183*, 1838-1845. 10.4049/jimmunol.0901066.

Han, X., Zhou, Z., Fei, L., Sun, H., Wang, R., Chen, Y., Chen, H., Wang, J., Tang, H.,

Ge, W., et al. (2020). Construction of a human cell landscape at single-cell level. Nature *581*, 303-309. 10.1038/s41586-020-2157-4.

Hänzelmann, S., Castelo, R., and Guinney, J. (2013). GSVA: gene set variation analysis for microarray and RNA-seq data. BMC bioinformatics *14*, 7. 10.1186/1471-2105-14-7.

Hao, S., Yan, K.K., Ding, L., Qian, C., Chi, H., and Yu, J. (2020). Network Approaches for Dissecting the Immune System. iScience *23*, 101354. 10.1016/j.isci.2020.101354.

Hashimoto, K., Kouno, T., Ikawa, T., Hayatsu, N., Miyajima, Y., Yabukami, H., Terooatea, T., Sasaki, T., Suzuki, T., Valentine, M., et al. (2019). Single-cell transcriptomics reveals expansion of cytotoxic CD4 T cells in supercentenarians. Proc Natl Acad Sci U S A *116*, 24242-24251. 10.1073/pnas.1907883116.

He, S., and Sharpless, N.E. (2017). Senescence in Health and Disease. Cell *169*, 1000-1011. 10.1016/j.cell.2017.05.015.

Hill, D.L., Carr, E.J., Rutishauser, T., Moncunill, G., Campo, J.J., Innocentin, S., Mpina, M., Nhabomba, A., Tumbo, A., Jairoce, C., et al. (2020). Immune system development varies according to age, location, and anemia in African children. Science translational medicine *12*. 10.1126/scitranslmed.aaw9522.

Ivanov, II, McKenzie, B.S., Zhou, L., Tadokoro, C.E., Lepelley, A., Lafaille, J.J., Cua, D.J., and Littman, D.R. (2006). The orphan nuclear receptor RORgammat directs the differentiation program of proinflammatory IL-17+ T helper cells. Cell *126*, 1121-1133. 10.1016/j.cell.2006.07.035.

Jin, S., Guerrero-Juarez, C.F., Zhang, L., Chang, I., Ramos, R., Kuan, C.H., Myung, P., Plikus, M.V., and Nie, Q. (2021). Inference and analysis of cell-cell communication using CellChat. Nature communications *12*, 1088. 10.1038/s41467-021-21246-9.

Kollmann, T.R., Kampmann, B., Mazmanian, S.K., Marchant, A., and Levy, O. (2017). Protecting the Newborn and Young Infant from Infectious Diseases: Lessons from Immune Ontogeny. Immunity *46*, 350-363. 10.1016/j.immuni.2017.03.009.

Ksienzyk, A., Neumann, B., Nandakumar, R., Finsterbusch, K., Grashoff, M., Zawatzky, R., Bernhardt, G., Hauser, H., and Kröger, A. (2011). IRF-1 expression is essential for natural killer cells to suppress metastasis. Cancer Res *71*, 6410-6418. 10.1158/0008-5472.can-11-1565.

Kundu, P., Blacher, E., Elinav, E., and Pettersson, S. (2017). Our Gut Microbiome: The Evolving Inner Self. Cell *171*, 1481-1493. 10.1016/j.cell.2017.11.024.

Legoux, F., Bellet, D., Daviaud, C., El Morr, Y., Darbois, A., Niort, K., Procopio, E., Salou, M., Gilet, J., Ryffel, B., et al. (2019). Microbial metabolites control the thymic

development of mucosal-associated invariant T cells. Science *366*, 494-499. 10.1126/science.aaw2719.

Legoux, F., Salou, M., and Lantz, O. (2020). MAIT Cell Development and Functions: the Microbial Connection. Immunity *53*, 710-723. 10.1016/j.immuni.2020.09.009.

Leuti, A., Fazio, D., Fava, M., Piccoli, A., Oddi, S., and Maccarrone, M. (2020). Bioactive lipids, inflammation and chronic diseases. Advanced drug delivery reviews *159*, 133-169. 10.1016/j.addr.2020.06.028.

Li, H., Horns, F., Wu, B., Xie, Q., Li, J., Li, T., Luginbuhl, D.J., Quake, S.R., and Luo, L. (2017). Classifying Drosophila Olfactory Projection Neuron Subtypes by Single-Cell RNA Sequencing. Cell *171*, 1206-1220.e1222. 10.1016/j.cell.2017.10.019.

Li, J., Liang, X., Jiang, J., Yang, L., Xin, J., Shi, D., Lu, Y., Li, J., Ren, K., Hassan, H.M., et al. (2022). PBMC transcriptomics identifies immune-metabolism disorder during the development of HBV-ACLF. Gut *71*, 163-175. 10.1136/gutjnl-2020-323395.

Liberzon, A., Subramanian, A., Pinchback, R., Thorvaldsdóttir, H., Tamayo, P., and Mesirov, J.P. (2011). Molecular signatures database (MSigDB) 3.0. Bioinformatics (Oxford, England) *27*, 1739-1740. 10.1093/bioinformatics/btr260.

López-Otín, C., and Kroemer, G. (2021). Hallmarks of Health. Cell *184*, 33-63. 10.1016/j.cell.2020.11.034.

Ma, S., Sun, S., Geng, L., Song, M., Wang, W., Ye, Y., Ji, Q., Zou, Z., Wang, S., He, X., et al. (2020). Caloric Restriction Reprograms the Single-Cell Transcriptional Landscape of Rattus Norvegicus Aging. Cell *180*, 984-1001.e1022. 10.1016/j.cell.2020.02.008.

Malka-Ruimy, C., Ben Youssef, G., Lambert, M., Tourret, M., Ghazarian, L., Faye, A., Caillat-Zucman, S., and Houdouin, V. (2019). Mucosal-Associated Invariant T Cell Levels Are Reduced in the Peripheral Blood and Lungs of Children With Active Pulmonary Tuberculosis. Frontiers in immunology *10*, 206. 10.3389/fimmu.2019.00206.

Mazzola, A., Cianti, R., Bini, L., Armini, A., Eberini, I., Pompella, G., Capecchi, P.L., Natale, M., Abbracchio, M.P., and Laghi-Pasini, F. (2008). Using peripheral blood mononuclear cells to determine proteome profiles in human cardiac failure. European journal of heart failure *10*, 749-757. 10.1016/j.ejheart.2008.06.003.

Mittelbrunn, M., and Kroemer, G. (2021). Hallmarks of T cell aging. Nat Immunol *22*, 687-698. 10.1038/s41590-021-00927-z.

Mogilenko, D.A., Shpynov, O., Andhey, P.S., Arthur, L., Swain, A., Esaulova, E., Brioschi, S., Shchukina, I., Kerndl, M., Bambouskova, M., et al. (2021).

Comprehensive Profiling of an Aging Immune System Reveals Clonal GZMK(+) CD8(+) T Cells as Conserved Hallmark of Inflammaging. Immunity *54*, 99-115.e112. 10.1016/j.immuni.2020.11.005.

Montecino-Rodriguez, E., Berent-Maoz, B., and Dorshkind, K. (2013). Causes, consequences, and reversal of immune system aging. The Journal of clinical investigation *123*, 958-965. 10.1172/JCI64096.

Moore, D.F., Li, H., Jeffries, N., Wright, V., Cooper, R.A., Jr., Elkahloun, A., Gelderman, M.P., Zudaire, E., Blevins, G., Yu, H., et al. (2005). Using peripheral blood mononuclear cells to determine a gene expression profile of acute ischemic stroke: a pilot investigation. Circulation *111*, 212-221. 10.1161/01.cir.0000152105.79665.c6.

Nehar-Belaid, D., Hong, S., Marches, R., Chen, G., Bolisetty, M., Baisch, J., Walters, L., Punaro, M., Rossi, R.J., Chung, C.H., et al. (2020). Mapping systemic lupus erythematosus heterogeneity at the single-cell level. Nat Immunol *21*, 1094-1106. 10.1038/s41590-020-0743-0.

Nikolich-Zugich, J. (2018). The twilight of immunity: emerging concepts in aging of the immune system. Nat Immunol *19*, 10-19. 10.1038/s41590-017-0006-x.

Nikolich-Žugich, J. (2018). The twilight of immunity: emerging concepts in aging of the immune system. Nature immunology *19*, 10-19. 10.1038/s41590-017-0006-x.

Oracki, S.A., Walker, J.A., Hibbs, M.L., Corcoran, L.M., and Tarlinton, D.M. (2010). Plasma cell development and survival. Immunological reviews *237*, 140-159. 10.1111/j.1600-065X.2010.00940.x.

Peters, M.J., Joehanes, R., Pilling, L.C., Schurmann, C., Conneely, K.N., Powell, J., Reinmaa, E., Sutphin, G.L., Zhernakova, A., Schramm, K., et al. (2015). The transcriptional landscape of age in human peripheral blood. Nature communications *6*, 8570. 10.1038/ncomms9570.

Pomaznoy, M., Kuan, R., Lindvall, M., Burel, J.G., Seumois, G., Vijayanand, P., Taplitz, R., Gilman, R.H., Saito, M., Lewinsohn, D.M., et al. (2020). Quantitative and Qualitative Perturbations of CD8(+) MAITs in Healthy Mycobacterium tuberculosis-Infected Individuals. ImmunoHorizons *4*, 292-307. 10.4049/immunohorizons.2000031.

Provine, N.M., and Klenerman, P. (2020). MAIT Cells in Health and Disease. Annual review of immunology *38*, 203-228. 10.1146/annurev-immunol-080719-015428.

Qi, Q., Liu, Y., Cheng, Y., Glanville, J., Zhang, D., Lee, J.Y., Olshen, R.A., Weyand, C.M., Boyd, S.D., and Goronzy, J.J. (2014). Diversity and clonal selection in the human T-cell repertoire. Proc Natl Acad Sci U S A *111*, 13139-13144.

10.1073/pnas.1409155111.
Qin, L., Jing, X., Qiu, Z., Cao, W., Jiao, Y., Routy, J., and Li, T. (2016). Aging of immune system: Immune signature from peripheral blood lymphocyte subsets in 1068 healthy adults. Aging *8*, 848-859. 10.18632/aging.100894.
Qiu, X., Mao, Q., Tang, Y., Wang, L., Chawla, R., Pliner, H.A., and Trapnell, C. (2017). Reversed graph embedding resolves complex single-cell trajectories. Nature methods *14*, 979-982. 10.1038/nmeth.4402.
Ren, X., Wen, W., Fan, X., Hou, W., Su, B., Cai, P., Li, J., Liu, Y., Tang, F., Zhang, F., et al. (2021). COVID-19 immune features revealed by a large-scale single-cell transcriptome atlas. Cell *184*, 1895-1913.e1819. 10.1016/j.cell.2021.01.053.
Rieckmann, J.C., Geiger, R., Hornburg, D., Wolf, T., Kveler, K., Jarrossay, D., Sallusto, F., Shen-Orr, S.S., Lanzavecchia, A., Mann, M., and Meissner, F. (2017). Social network architecture of human immune cells unveiled by quantitative proteomics. Nat Immunol *18*, 583-593. 10.1038/ni.3693.
Ruibal, P., Voogd, L., Joosten, S.A., and Ottenhoff, T.H.M. (2021). The role of donor-unrestricted T-cells, innate lymphoid cells, and NK cells in anti-mycobacterial immunity. Immunological reviews *301*, 30-47. 10.1111/imr.12948.
Saso, A., and Kampmann, B. (2017). Vaccine responses in newborns. Seminars in immunopathology *39*, 627-642. 10.1007/s00281-017-0654-9.
Sayed, N., Huang, Y., Nguyen, K., Krejciova-Rajaniemi, Z., Grawe, A.P., Gao, T., Tibshirani, R., Hastie, T., Alpert, A., Cui, L., et al. (2021). An inflammatory aging clock (iAge) based on deep learning tracks multimorbidity, immunosenescence, frailty and cardiovascular aging. Nature Aging *1*, 598-615. 10.1038/s43587-021-00082-y.
See, P., Dutertre, C.A., Chen, J., Günther, P., McGovern, N., Irac, S.E., Gunawan, M., Beyer, M., Händler, K., Duan, K., et al. (2017). Mapping the human DC lineage through the integration of high-dimensional techniques. Science *356*. 10.1126/science.aag3009.
Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome research *13*, 2498-2504. 10.1101/gr.1239303.
Shao, X., Liao, J., Li, C., Lu, X., Cheng, J., and Fan, X. (2021). CellTalkDB: a manually curated database of ligand-receptor interactions in humans and mice. Briefings in bioinformatics *22*. 10.1093/bib/bbaa269.
Simon, A.K., Hollander, G.A., and McMichael, A. (2015). Evolution of the immune system in humans from infancy to old age. Proceedings. Biological sciences *282*,

20143085. 10.1098/rspb.2014.3085.
Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., and Mesirov, J.P. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A *102*, 15545-15550. 10.1073/pnas.0506580102.
Supek, F., Bošnjak, M., Škunca, N., and Šmuc, T. (2011). REVIGO summarizes and visualizes long lists of gene ontology terms. PloS one *6*, e21800. 10.1371/journal.pone.0021800.
Taylor, M.K., and Cohen, J.J. (1992). Cell-mediated cytotoxicity. Curr Opin Immunol *4*, 338-343. 10.1016/0952-7915(92)90086-t.
Toubal, A., Nel, I., Lotersztajn, S., and Lehuen, A. (2019). Mucosal-associated invariant T cells and disease. Nature reviews. Immunology *19*, 643-657. 10.1038/s41577-019-0191-y.
Vallejo, A.N. (2011). Immunological hurdles of ageing: indispensable research of the human model. Ageing research reviews *10*, 315-318. 10.1016/j.arr.2011.01.005.
van den Broek, T., Borghans, J.A.M., and van Wijk, F. (2018). The full spectrum of human naive T cells. Nature reviews. Immunology *18*, 363-373. 10.1038/s41577-018-0001-y.
van Gisbergen, K.P., Klarenbeek, P.L., Kragten, N.A., Unger, P.P., Nieuwenhuis, M.B., Wensveen, F.M., ten Brinke, A., Tak, P.P., Eldering, E., Nolte, M.A., and van Lier, R.A. (2011). The costimulatory molecule CD27 maintains clonally diverse CD8(+) T cell responses of low antigen affinity to protect against viral variants. Immunity *35*, 97-108. 10.1016/j.immuni.2011.04.020.
van Wilgenburg, B., Loh, L., Chen, Z., Pediongco, T.J., Wang, H., Shi, M., Zhao, Z., Koutsakos, M., Nüssing, S., Sant, S., et al. (2018). MAIT cells contribute to protection against lethal influenza infection in vivo. Nature communications *9*, 4706-4706. 10.1038/s41467-018-07207-9.
Vijg, J., and Calder, R.B. (2004). Transcripts of aging. Trends Genet *20*, 221-224. 10.1016/j.tig.2004.04.007.
Villani, A.C., Satija, R., Reynolds, G., Sarkizova, S., Shekhar, K., Fletcher, J., Griesbeck, M., Butler, A., Zheng, S., Lazo, S., et al. (2017). Single-cell RNA-seq reveals new types of human blood dendritic cells, monocytes, and progenitors. Science *356*. 10.1126/science.aah4573.
Wang, P., Ning, S., Zhang, Y., Li, R., Ye, J., Zhao, Z., Zhi, H., Wang, T., Guo, Z., and Li, X. (2015). Identification of lncRNA-associated competing triplets reveals global

patterns and prognostic markers for cancer. Nucleic acids research *43*, 3478-3489. 10.1093/nar/gkv233.

Wang, Z., Xie, L., Ding, G., Song, S., Chen, L., Li, G., Xia, M., Han, D., Zheng, Y., Liu, J., et al. (2021). Single-cell RNA sequencing of peripheral blood mononuclear cells from acute Kawasaki disease patients. Nature communications *12*, 5444. 10.1038/s41467-021-25771-5.

Wei, S.C., Levine, J.H., Cogdill, A.P., Zhao, Y., Anang, N.A.S., Andrews, M.C., Sharma, P., Wang, J., Wargo, J.A., Pe'er, D., and Allison, J.P. (2017). Distinct Cellular Mechanisms Underlie Anti-CTLA-4 and Anti-PD-1 Checkpoint Blockade. Cell *170*, 1120-1133.e1117. 10.1016/j.cell.2017.07.024.

Wolock, S.L., Lopez, R., and Klein, A.M. (2019). Scrublet: Computational Identification of Cell Doublets in Single-Cell Transcriptomic Data. Cell systems *8*, 281-291.e289. 10.1016/j.cels.2018.11.005.

World Medical, A. (2013). World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. JAMA *310*, 2191-2194. 10.1001/jama.2013.281053.

Xu, L., Su, Y., Zhao, Y., Sheng, X., Tong, R., Ying, X., Gao, L., Ji, Q., Gao, Y., Yan, Y., et al. (2019). Melatonin differentially regulates pathological and physiological cardiac hypertrophy: Crucial role of circadian nuclear receptor RORalpha signaling. J Pineal Res *67*, e12579. 10.1111/jpi.12579.

Xu, W., Narayanan, P., Kang, N., Clayton, S., Ohne, Y., Shi, P., Herve, M.C., Balderas, R., Picard, C., Casanova, J.L., et al. (2014). Human plasma cells express granzyme B. Eur J Immunol *44*, 275-284. 10.1002/eji.201343711.

Yousefzadeh, M., Flores, R., Zhu, Y., Schmiechen, Z., Brooks, R., Trussoni, C., Cui, Y., Angelini, L., Lee, K., McGowan, S., et al. (2021). An aged immune system drives senescence and ageing of solid organs. Nature. 10.1038/s41586-021-03547-7.

Yu, G., Wang, L.G., Han, Y., and He, Q.Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. Omics : a journal of integrative biology *16*, 284-287. 10.1089/omi.2011.0118.

Zhang, L., Dong, X., Lee, M., Maslov, A.Y., Wang, T., and Vijg, J. (2019). Single-cell whole-genome sequencing reveals the functional landscape of somatic mutations in B lymphocytes across the human lifespan. Proc Natl Acad Sci U S A *116*, 9014-9019. 10.1073/pnas.1902510116.

Zhang, L., Yu, X., Zheng, L., Zhang, Y., Li, Y., Fang, Q., Gao, R., Kang, B., Zhang, Q., Huang, J.Y., et al. (2018). Lineage tracking reveals dynamic relationships of T cells in colorectal cancer. Nature *564*, 268-272. 10.1038/s41586-018-0694-x.

Zhang, Y., Yao, Y., Chen, P., Liu, Y., Zhang, H., Liu, H., Liu, Y., Xu, H., Tian, X., Wang, Z., et al. (2020). Checkpoint therapeutic target database (CKTTD): the first comprehensive database for checkpoint targets and their modulators in cancer immunotherapy. Journal for immunotherapy of cancer *8*. 10.1136/jitc-2020-001247.

Zhu, L., Yang, P., Zhao, Y., Zhuang, Z., Wang, Z., Song, R., Zhang, J., Liu, C., Gao, Q., Xu, Q., et al. (2020). Single-Cell Sequencing of Peripheral Mononuclear Cells Reveals Distinct Immune Response Landscapes of COVID-19 and Influenza Patients. Immunity *53*, 685-696.e683. 10.1016/j.immuni.2020.07.009.

Figures and figure legends

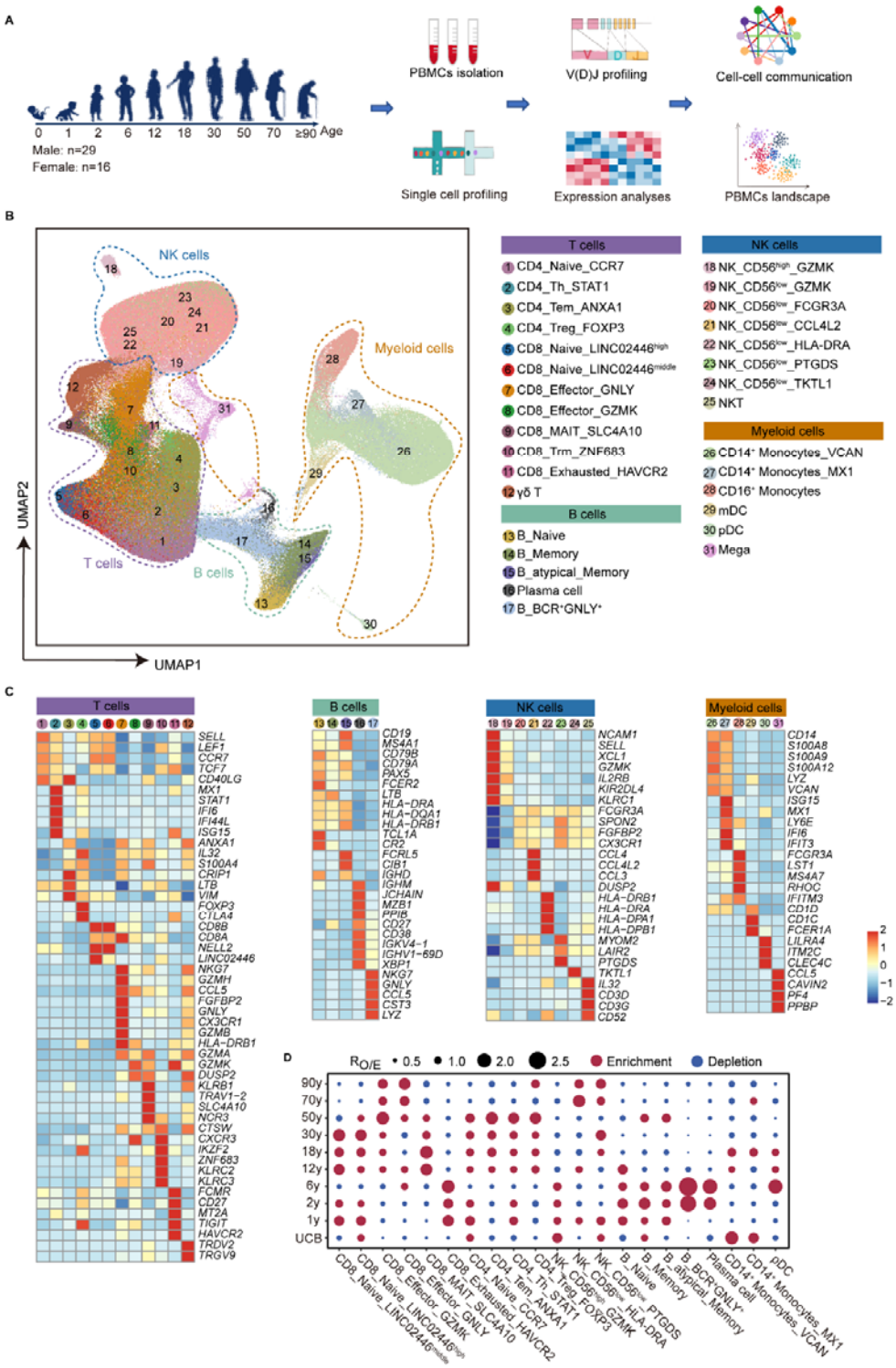


Figure 1. Overview of the single-cell transcriptome of peripheral immune cells across the entire lifespan

(A) Workflow of the overall experimental design of this study.

(B) UMAP charts of more than 380,000 single immune cells from 45 human individuals. In the UMAP, 31 immune cell subsets are labelled with different colors.

(C) Heatmap showing the expression of signature genes in each cell subset. Red, high expression; blue, low expression.

(D) Age group preference of each cell subset measured by the ratio of observed to randomly expected cell numbers ($R_{O/E}$).

mDC, myeloid dendritic cells; pDC, plasmacytoid dendritic cells; Mega, megakaryocytes.

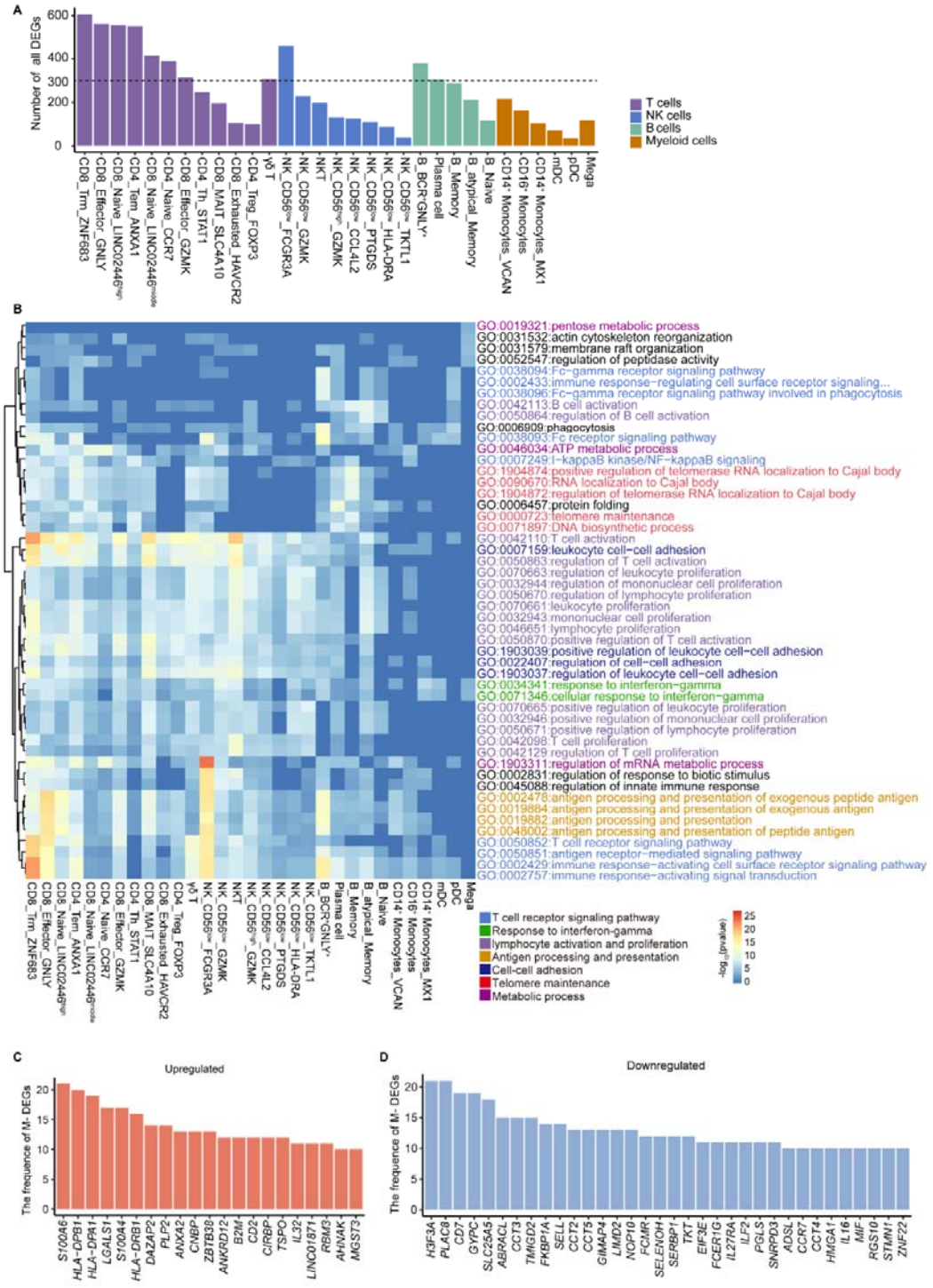


Figure 2. Global transcriptomic dynamics of peripheral immune cells throughout lifespan

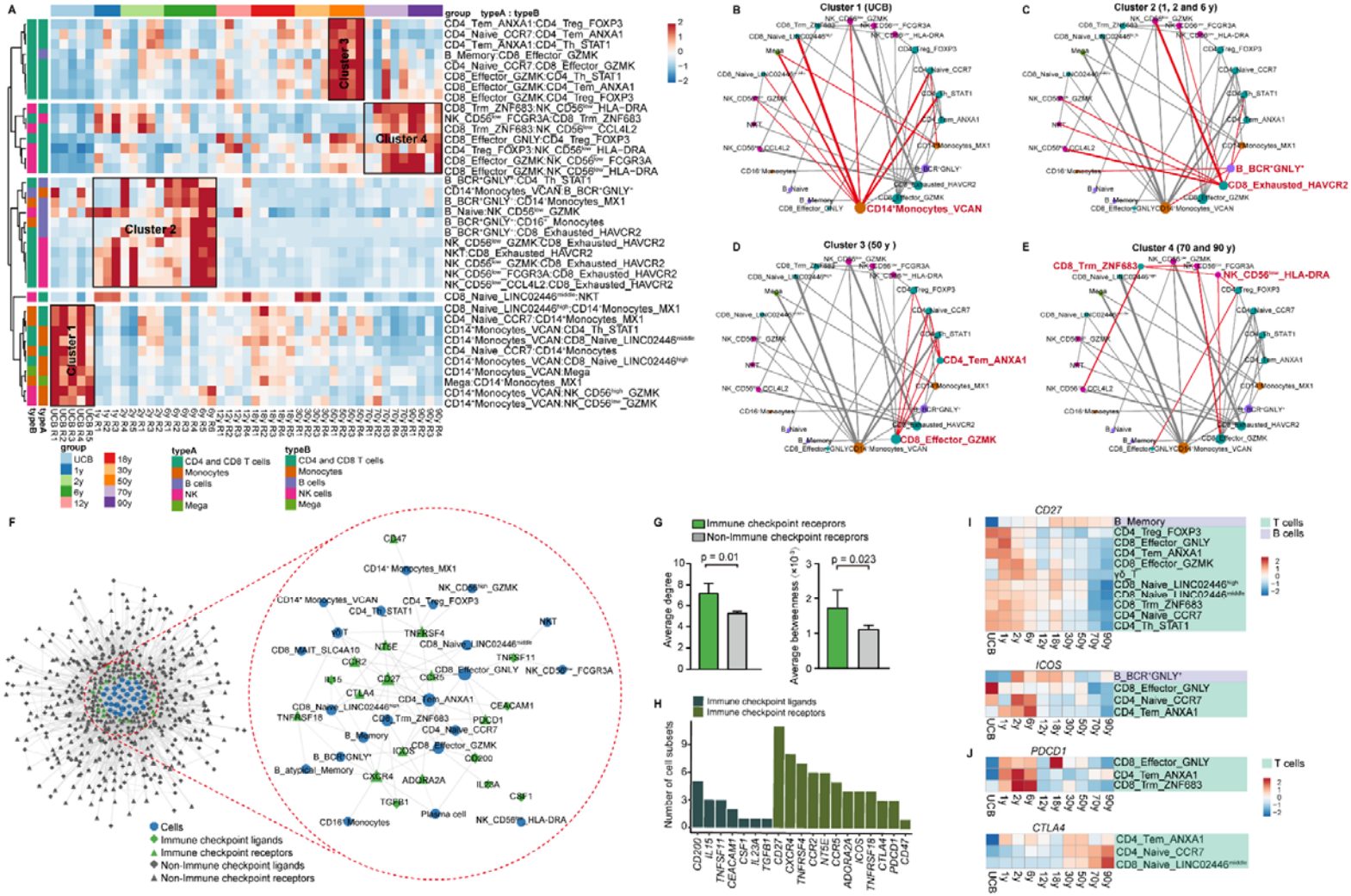
(A) Bar plots showing the number of DEGs in all cell subsets.

(B) Functional enrichment analysis of DEGs in all cell subsets.

(C) Bar plots showing the frequencies of upregulated M-DEGs observed in all cell

subsets and the frequency is greater than or equal to 10.

(D) Bar plots showing the frequencies of downregulated M-DEGs observed in all cell subsets and the frequency is greater than or equal to 10.


A
group typeA : typeB
Cluster 1
Cluster 2
Cluster 3
Cluster 4
group
UCB
1y
2y
6y
12y
18y
30y
50y
70y
90y
typeA
CD4 and CD8 T cells
Monocytes
B cells
NK
Mega
typeB
CD4 and CD8 T cells
Monocytes
B cells
NK cells
Mega
B
Cluster 1 (UCB)
C
Cluster 2 (1, 2 and 6 y)
D
Cluster 3 (50 y)
E
Cluster 4 (70 and 90 y)
F
Cells
Immune checkpoint ligands
Immune checkpoint receptors
Non-Immune checkpoint ligands
Non-Immune checkpoint receptors
G
Immune checkpoint receprors
Non-Immune checkpoint receprors
p = 0.01
p = 0.023
Average degree
Average betweenness
H
Immune checkpoint ligands
Immune checkpoint receptors
Number of cell subsets
CD200
IL15
TNFSF11
CEACAM1
CSF1
IL23A
TGFB1
CD27
CXCR4
TNFRSF4
CCR2
N75E
CCR5
ADORA2A
ICOS
TNFRSF18
CTLA4
PDCD1
CD47
I
CD27
B_Memory
CD4_Treg_FOXP3
CD8_Effector_GNLY
CD4_Tem_ANXA1
CD8_Effector_GZMK
γδ_T
CD8_Naive_LINC02446
CD8_Trm_ZNF683
CD4_Naive_CCR7
CD4_Th_STAT1
T cells
B cells
ICOS
B_BCR+GNLY+
CD8_Effector_GNLY
CD4_Naive_CCR7
CD4_Tem_ANXA1
J
PDCD1
CD8_Effector_GNLY
CD4_Tem_ANXA1
CD8_Trm_ZNF683
T cells
CTLA4
CD4_Tem_ANXA1
CD4_Naive_CCR7
CD8_Naive_LINC02446

1

Figure 3. Rewiring of cell-cell interaction among peripheral immune cells across the whole lifespan

(A) Heatmap showing the probabilities of cell-cell interactions varied significantly across the whole lifespan. These cell-cell interactions were clustered into five clusters based on hierarchical clustering, four of which were enriched in specific age groups. The varied significantly CCIs among different age groups were determined by comparing the CCI probabilities of all cell subset pairs in each sample using ANOVA analysis. A p value was adjusted with Bonferroni and p-value < 0.05 was considered statistically significant.

(B)-(E) Cell-cell interactions in (A) were visualized by cytoscape and the cell-cell interactions in each cluster were marked with red edges. The width of line represents the average interaction strength of all age groups. The size of nodes represents the degree of cell subset. Hubs of cluster were highlighted in red.

(F) Cell-cell interaction network mediated by differentially expressed ligands and receptors. The sub-network circled in red showing the immune checkpoints and the corresponding affected cell subsets.

(G) Boxplot showing the degree and betweenness centrality of immune checkpoint receptors and non-immune checkpoint receptors in (F). P value determined by two-sided unpaired Wilcoxon test. A p-value < 0.05 was considered statistically significant.

(H) Bar plot showing the immune checkpoints differentially expressed across the human lifespan and the numbers of cell subsets affected by them.

(I) Heatmap showing the expression of co-stimulatory checkpoint *CD27* and *ICOS* in the corresponding affected cell subsets.

(J) Heatmap showing the expression of co-inhibitory checkpoint *PDCD1* and *CTLA4* in the corresponding affected cell subsets.

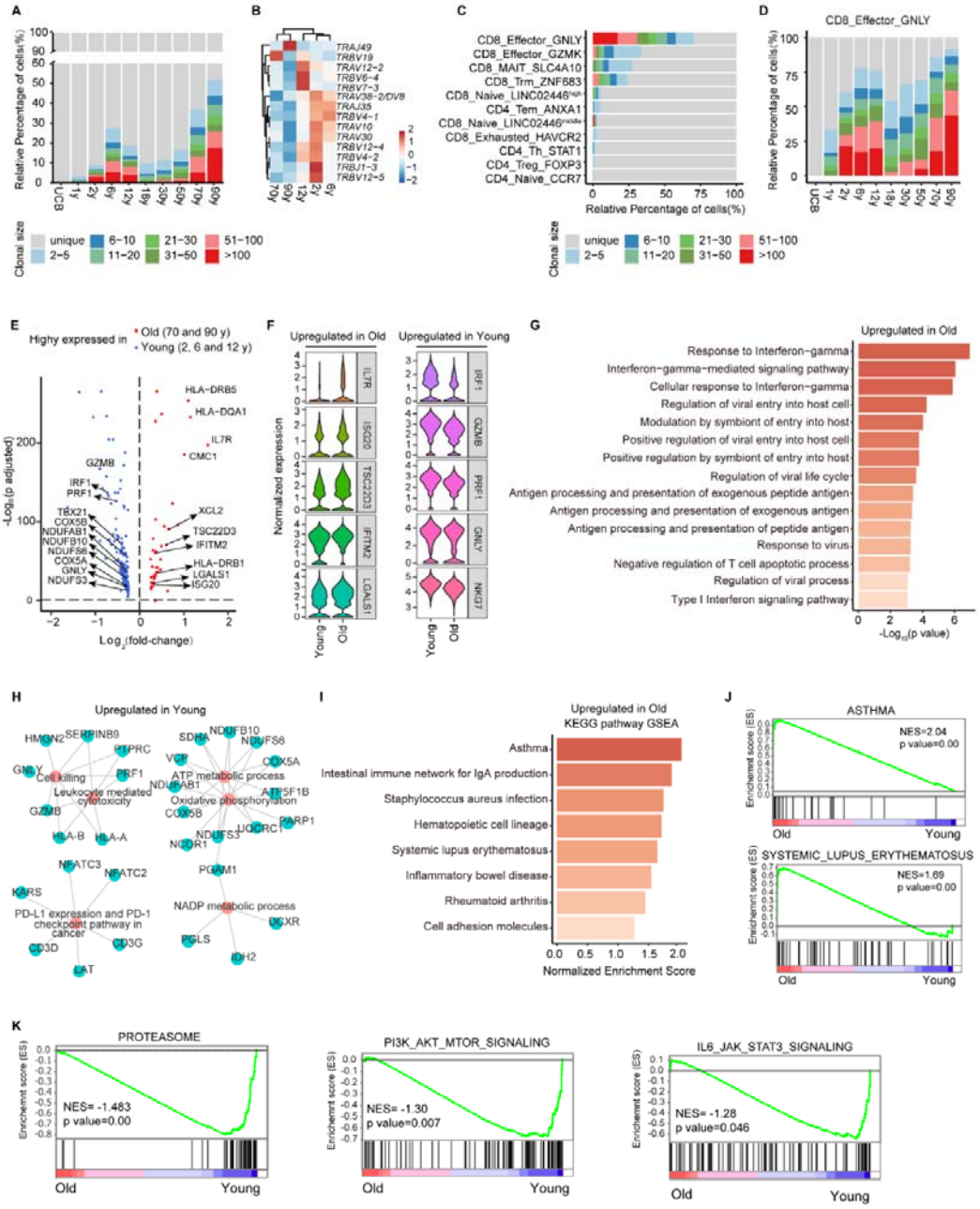


Figure 4. Clonal expansion of CD8 effector T cells in young and old individuals

(A) Bar plots showing the distribution of T cell clone sizes across ten age groups.

(B) Heatmap showing the usage of significantly differentially- expressed V-J genes from TCRα chain and TCR β chain across five age groups. P values were determined by ANOVA and adjusted using FDR. The FDR< 0.05 was considered statistically significant. Red, high expression; blue, low expression.

(C) Bar plots showing the distribution of T cell clone sizes in different T cell subsets.

(D) Bar plots showing the distribution of T cell clone sizes in CD8_Effector_GNLY

cells across ten age groups.

(E) Volcano plot showing the DEGs in highly clonal expanded CD8_Effector_GNLY cells (clone size≥50 cells) in the old high-TCE group compared with those in the young high-TCE group. Example genes are labeled with gene names. Red, upregulated in the old high-TCE group [Log_2 (fold change) ≥ 0.25, adjusted p value < 0.05]; blue, upregulated in the young high-TCE group [Log_2 (fold change) ≤-0.25, adjusted p value < 0.05].

(F) Volcano plot showing the expression of some DEGs in highly clonal expanded CD8_Effector_GNLY cells (clone size≥50 cells) in the old group compared with the young group.

(G) Functional enrichment analysis of upregulated DEGs in highly clonal expanded CD8_Effector_GNLY cells in the old high-TCE group compared with those in the young high-TCE group.

(H) Gene concept network showing the upregulated DEGs in highly clonal expanded CD8_Effector_GNLY cells of young high-TCE group and enriched biological processes and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways.

(I) Bar plots showing the normalized enrichment score of enriched KEGG pathways obtained by GSEA analysis.

(J) GSEA enrichment plots showing two upregulated gene sets in highly clonal expanded CD8_Effector_GNLY cells in old high-TCE group compared with those in the young group. NES, normalized enrichment score. A p-value<□0.05 was statistically significant. NES, normalized enrichment score. A p-value□<□0.05 was considered statistically significant.

(K) GSEA enrichment plots showing three upregulated gene sets in highly clonal expanded CD8_Effector_GNLY cells in young high-TCE group compared with those in the old high-TCE group. NES, normalized enrichment score. A p-value <□0.05 was statistically significant. NES, normalized enrichment score. A p-value□<□0.05 was considered statistically significant.

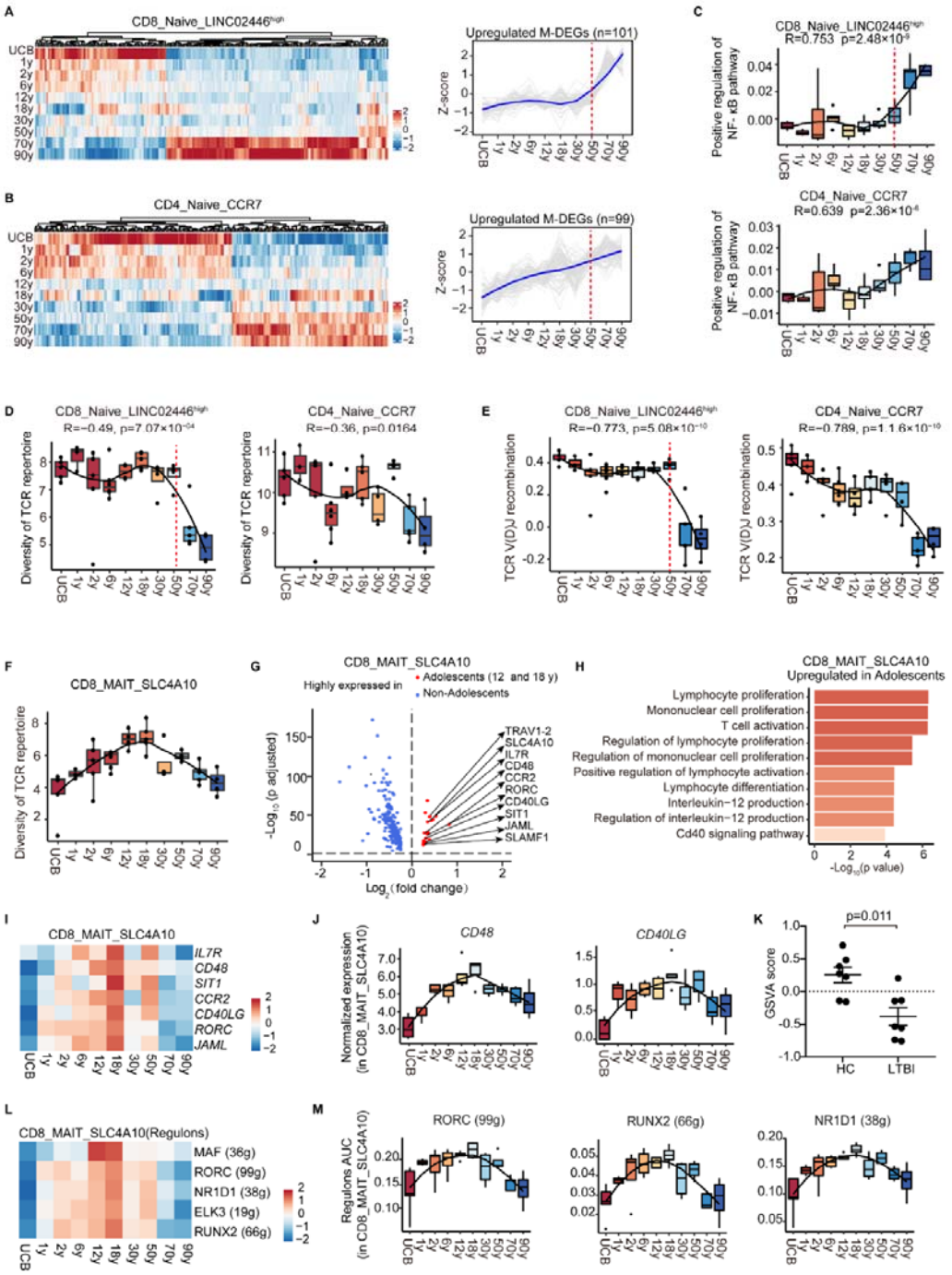


Figure 5. Aging features of naive T cells and CD8^{+} MAIT cells

(A) Heatmap showing the expression of M-DEGs in CD8_Naive_LINC02446^{high} cells across ten age groups (left panel). Expression trajectories of upregulated M-DEGs in CD8_Naive_LINC02446^{high} cells (right panel). Data are shown in terms of z-score of gene expression. Each grey line represents expression trajectory of each gene. Bule

line represents the average expression trajectory.

(B) Heatmap showing the expression of M-DEGs in CD4_Naive_CCR7 cells across ten age groups (left panel). Expression trajectories of upregulated M-DEGs in CD4_Naive_CCR7 cells (right panel). Data are shown in terms of z-score of gene expression. Each grey line represents expression trajectory of each gene. Bule line represents the average expression trajectory.

(C) Box plots showing the expression of positive regulation of NF-κB signaling pathway in CD8_Naive_LINC02446^{high} and CD4_Naive_CCR7 cells across ten age groups. The black line indicates loess regression. Spearman correlation coefficient R and p value are indicated. A p-value < 0.05 was considered statistically significant.

(D) Box plots showing the diversity of TCR repertoire in CD8_Naive_LINC02446^{high} (left) and CD4_Naive_CCR7 (right) cells across ten age groups. All box plots show median, 25th and 75th percentiles, and whiskers extending to maximum and minimum data points, with outliers beyond. The black line indicates loess regression. Spearman correlation coefficient R and p values are indicated. A p-value < 0.05 was considered statistically significant.

(E) Box plots showing the expression of TCR V(D)J recombination in CD8_Naive_LINC02446^{high} (left) and CD4_Naive_CCR7 (right) cells across ten age groups. All box plots show median, 25th and 75th percentiles, and whiskers extending to maximum and minimum data points, with outliers beyond. The black line indicates loess regression. Spearman correlation coefficient R and p values are indicated. A p-value < 0.05 was considered statistically significant.

(F) Diversity of TCR repertoire in CD8_MAIT_SLC4A10 cells across ten age groups. All box plots show median, 25th and 75th percentiles, and whiskers extending to maximum and minimum data points, with outliers beyond. The black lines indicate loess regression. Spearman correlation coefficient R and p values are indicated.

(G) Volcano plot showing the DEGs in CD8_MAIT_SLC4A10 cells in adolescents compared with those in non-adolescents. Example genes are labeled with gene names. Red, upregulated in adolescents [Log_2 (fold change) ≥ 0.25, adjusted p value < 0.05]; blue, upregulated in in non-adolescents [Log_2 (fold change) ≤-0.25, adjusted p value < 0.05].

(H) Functional enrichment analysis of upregulated DEGs in CD8_MAIT_SLC4A10 cells in adolescents compared with those in non-adolescents.

(I) Heatmap showing the expression of T cell proliferation and activation related

genes in CD8_MAIT_SLC4A10 cells across ten age groups.

(J) Box plots showing the expression of *CD48* and *CD40LG* in CD8_MAIT_SLC4A10 cells across ten age groups. All box plots show median, 25th and 75th percentiles, and whiskers extending to maximum and minimum data points, with outliers beyond. The black line indicates loess regression.

(K) Scatter plots of GSVA score for T cell proliferation and activation related gene set in MAIT cells derived from healthy individuals (HC) and latent tuberculosis infection (LTBI) patients. Scatter dot plots showing mean ± SEM and p value determined by two-sided unpaired Wilcoxon test. A p-value < 0.05 was considered statistically significant.

(L) Heatmap showing the regulon area under the curve of inferred differentially expressed TF regulons in CD8_MAIT_SLC4A10 cells across ten age groups. Numbers between brackets represent the regulon counts for respective transcription factors. Red, high expression; blue, low expression.

(M) Box plots showing the area under curve of RORC (99g), RUNX2(66g) and NR1D1(38g) in CD8_MAIT_SLC4A10 cells across ten age groups. All box plots show median, 25th and 75th percentiles, and whiskers extending to maximum and minimum data points, with outliers beyond. The black line indicates loess regression. Numbers between brackets represent the regulon counts for respective transcription factors.

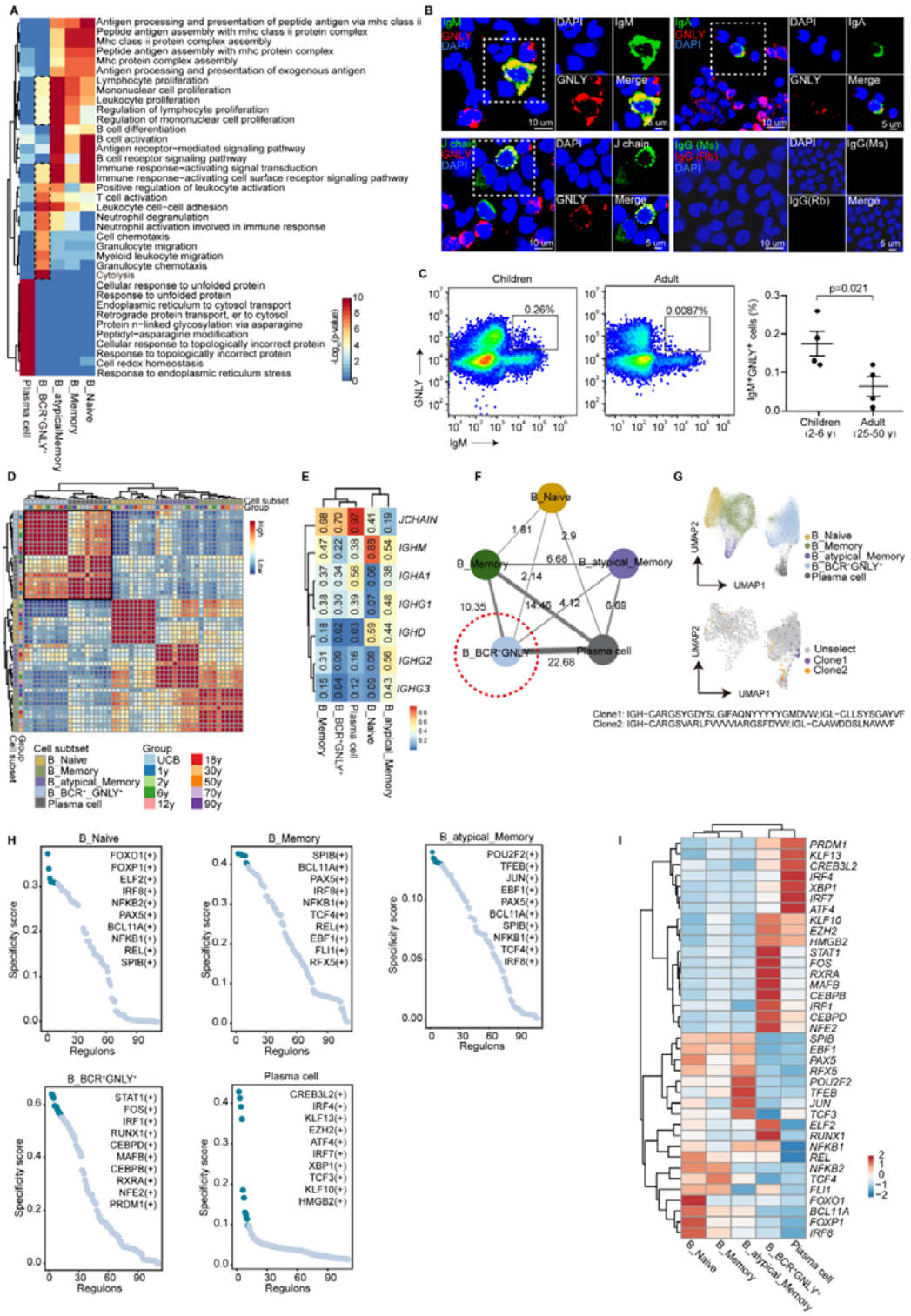


Figure 6. Function of the new-identified cytotoxic B cell subset (B_BCR^{+}GNLY^{+} cells)

(A) Heatmap showing the results of functional enrichment of signature genes in each B cell subset. The top 10 pathways in each cell subset were listed based on the p value.

(B) Immunofluorescence of B cell markers IgM, IgA and J chain, and granulysin (GNLY) in PBMCs collected from children. The green represents IgM, IgA or J chain, red represents granulysin, blue represents the nucleus and bright-yellow indicates double positive cells (IgM and granulysin, IgA and granulysin or J chain and granulysin). The experiment was repeated three times with similar results.

(C) Representative fluorescence activated cell sorting (FACS) scatter plots from 2 independent experiments, showing increased $IgM^{+}GNLY^{+}$ cell abundance in PBMCs in children (left panel). Scatter plots showing the percentage of $IgM^{+}GNLY^{+}$ cells (defined by FACS) in four children (2-6 years old) and four adults (25-50 years old) (right panel). Scatter dot plots showing mean ± SEM and p value determined by two-sided unpaired Wilcoxon test. A p-value < 0.05 was considered statistically significant.

(D) Correlation of gene expression among B cell subsets. Color scale: red indicates a high correlation; blue indicates a low correlation.

(E) Heatmap showing the UMI percentage of *JCHAIN* and immunoglobulin heavy chain genes (*IGHD*, *IGHM*, *IGHA1*, *IGHG1*, *IGHG2* and *IGHG3*) in each B cell subset.

(F) Network showing the ratio of shared BCR clones between two cell subsets. The width of the lines represents the ratio of shared BCR clones.

(G) UMAP plots showing selected shared BCR clones among B cell subsets. The BCR clones are labeled by different colors.

(H) Rank for TF regulons in each B cell subset based on the regulon specificity score (RSS). The top ten TF regulons were listed and marked in blue.

(I) Heatmap showing the expression of top ten TFs in each B cell subset shown in (G). Color scale: red, high expression; blue, low expression.