Multi-modal profiling of peripheral blood cells across the human lifespan reveals distinct immune cell signatures of aging and longevity

Tanya T. Karagiannis,a,n†, Todd W. Dowrey,b,n Carlos Villacorta-Martin,b Monty Montano,c,d Eric Reed,e Anna C. Belkina,f,g Stacy L. Andersen,h Thomas T. Perls,b Stefano Monti,i,j,k,n George J. Murphy,b,l,n and Paola Sebastiania,m,n

aInstitute for Clinical Research and Health Policy Studies, Tufts Medical Center, Boston, MA, USA
bCenter for Regenerative Medicine (CReM), Boston University and Boston Medical Center, Boston, MA, USA
cDepartment of Medicine, Brigham and Women’s Hospital, Harvard Medical School, Boston, MA, USA
dBoston Pepper Center, Brigham and Women’s Hospital, Harvard Medical School, Boston, MA, USA
eData Intensive Study Center, Tufts University, Boston, MA, USA
fDepartment of Pathology and Laboratory Medicine, Boston University School of Medicine, Boston, MA, USA
gFlow Cytometry Core Facility, Boston University School of Medicine, Boston, MA, USA
hDepartment of Medicine, Geriatrics Section, Boston University School of Medicine, Boston, MA, USA
iBioinformatics Program, Boston University, Boston, MA, USA
jDivision of Computational Biomedicine, Boston University School of Medicine, Boston, MA, USA
kDepartment of Biostatistics, Boston University School of Public Health, Boston, MA, USA
lSection of Hematology and Medical Oncology, Boston University School of Medicine, Boston, MA, USA
mDepartment of Medicine, Tufts University, Boston, MA, USA

Summary

Background Age-related changes in immune cell composition and functionality are associated with multimorbidity and mortality. However, many centenarians delay the onset of aging-related disease suggesting the presence of elite immunity that remains highly functional at extreme old age.

Methods To identify immune-specific patterns of aging and extreme human longevity, we analyzed novel single cell profiles from the peripheral blood mononuclear cells (PBMCs) of a random sample of 7 centenarians (mean age 106) and publicly available single cell RNA-sequencing (scRNA-seq) datasets that included an additional 7 centenarians as well as 52 people at younger ages (20–89 years).

Findings The analysis confirmed known shifts in the ratio of lymphocytes to myeloid cells, and noncytotoxic to cytotoxic cell distributions with aging, but also identified significant shifts from CD4+ T cell to B cell populations in centenarians suggesting a history of exposure to natural and environmental immunogens. We validated several of these findings using flow cytometry analysis of the same samples. Our transcriptional analysis identified cell type signatures specific to exceptional longevity that included genes with age-related changes (e.g., increased expression of STK17A, a gene known to be involved in DNA damage response) as well as genes expressed uniquely in centenarians’ PBMCs (e.g., S100A4, part of the S100 protein family studied in age-related disease and connected to longevity and metabolic regulation).

Interpretation Collectively, these data suggest that centenarians harbor unique, highly functional immune systems that have successfully adapted to a history of insults allowing for the achievement of exceptional longevity.

Funding TK, SM, PS, GM, SA, TP are supported by NIH-NIA UH2AG064704 and U19AG023122. MM and PS are supported by NIH NIA Pepper center: P30 AG031679-10. This project is supported by the Flow Cytometry Core Facility at BUSM. FCCF is funded by the NIH Instrumentation grant: S10 OD021587.

Copyright © 2023 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Keywords: Aging; Extreme human longevity; Elite immunity; Single cell RNA sequencing
Research in context

Evidence before this study
Aging is associated with multi-morbidity and mortality. Transcriptional studies of blood have previously identified global immune cell dysfunction with changes in the composition and gene expression profiles of lymphocytes and myeloid cells with age (Geiger et al., 2013, Petersone et al., 2015, Alpert et al., 2019). Centenarians, a rare population of individuals that reach 100 years, experience delays in aging-related diseases and mortality suggesting their immune systems remain functional at extreme old age. A recent study of peripheral blood mononuclear cells (PBMCs) from Japanese centenarians observed changes in composition of immune cells and increased cytotoxic lymphocytes compared to younger ages (Hashimoto et al., 2019). However, the generalizability of these results to other ethnicities and most importantly the transcriptional changes that occur in peripheral immune cell types of centenarians compared to younger individuals are still unclear. Further investigation to characterize the repertoire of immune cells of individuals who reach Extreme Longevity (EL) may indicate important mechanisms that promote EL.

Added value of this study
We performed CITE-seq on PBMCs from 7 centenarians to provide transcriptome-wide expression data in addition to a 10 cell-surface protein marker panel to increase the specificity of the identification of major immune cell types. We then integrated this dataset with two publicly available scRNA-seq datasets of PBMCs to investigate compositional and transcriptional changes in circulating immune profiles across the human lifespan and extreme old age. We found three patterns of changes in the immune cell type composition and gene expression profiles across the human lifespan that include aging related changes and changes unique to centenarians. We also validated these patterns using an orthogonal methodology (flow cytometry). The unique transcriptional changes we observed in centenarians compared to younger ages point to changes in metabolic regulation. In addition, we observed compositional changes within myeloid and lymphocyte lineages that reflect a greater exposure to infections and a shift in the immune resiliency strategy in EL compared to younger ages, which could be used for further investigation into immune resilience of EL and healthy aging.

Implications of all the available evidence
Our single cell analysis of peripheral blood immune cell populations across the human lifespan confirms observations made in previous studies of aging and identifies novel cell type specific compositional and transcriptional changes that are specific to centenarians and reflect immunocompetent profiles. These findings provide a foundation to investigate immune resilience mechanisms of extreme longevity as a target for healthy aging therapeutics.

Introduction
While a decline in cellular, organismic, and overall functionality is an inexorable outcome of aging, the rate and impact of aging is increasingly recognized to be affected by multiple factors including environment, genetics, and immune history.1–4 At a phenotypic level, aging leads to functional abnormalities and alterations in hematopoietic cell populations that prevent a proper immune response and lead to increased susceptibility to infections, cancers, and auto-immune diseases.1–3 Driven by transcriptional changes and alterations in gene expression, the global immune cell dysfunction generally observed with aging results in distinct shifts in the composition of peripheral immune cell types characterized by a loss of naive B and T cells and an accumulation of memory effector T and B cells,5,6 and more recently the age-associated expansion of granzyme K (GZMK)-expressing CD8+ T cells.7 In addition, an increase in inflammation-promoting cell populations, such as Natural Killer and myeloid cells (e.g., monocytes) are observed with aging,5,6 in parallel with gene expression changes in these populations.5,6

At the extreme of the human aging process is extreme longevity (EL), characterized by survival beyond an age reached by less than 1% of a cohort.8 EL is often, but not always, associated with a marked delay of disability and in majority (about 60%), common aging-related diseases.1–13 Changes in immune cells are considered one of the hallmarks of aging, with growing recognition that the loss of immune competence to control inflammation and rebound from immune stressors is central to the progression of age-limiting morbidities.14 Since centenarians—individuals who live to at least 100 years—appear to experience a slower pace of aging,13,15 characterizing the repertoire of immune cells of these elite individuals may point to important mechanisms that promote EL.

A recent study using single cell RNA sequencing (scRNA-seq) of peripheral blood mononuclear cells (PBMCs) displayed changes in the distribution of lymphocytes and myeloid cells, and a significant expansion of cytotoxic CD4+ T cells in individuals who live to at least 105 years compared to younger individuals.16 Hashimoto’s study focused on a cohort of Japanese individuals,16 thus it is not clear whether those results generalize to other ethnicities. Furthermore, the study did not perform an extensive characterization of the transcriptional changes—i.e., changes in expression levels as a function of age and EL—that occur in these cell types.
In this study, we employ a multi-modal approach that combines single-cell transcriptomics with cell-surface protein profiling to characterize both the composition and transcriptional profiles of the peripheral immune system of centenarians. We perform a harmonization of a novel single cell dataset of centenarians with diverse, publicly available peripheral blood scRNA-seq datasets of aging and longevity in an effort to understand the dynamics of circulating immune cell populations throughout the human lifespan, and in particular, in EL.

Methods

Experimental procedure

Centenarians were enrolled in North America in 2019. The study was approved by the Boston Medical Center and Boston University Medical Campus IRB and all participants provided written informed consent. We provide a brief overview of the overall approach, and detailed information and methods for the recruitment for human subjects, blood sample collection and processing, CITE-seq and flow cytometry analysis of PBMCs, and all statistical methods in Supplementary Materials.

Single cell analysis

CITE-seq data of centenarians

Cellular Indexing and epitopes sequencing (CITE-seq) was performed on the 7 centenarians and 2 younger age individuals using a commercial droplet-based platform (10x Chromium). Detailed methods for preprocessing data, filtering, PCA analysis, batch correction, clustering, and cell type identification are provided in the Supplementary Materials.

Harmonization of datasets of aging and EL

We harmonized the CITE-seq data of centenarians from NECS with publicly available scRNA-seq data sets using the Harmony algorithm\(^7\) to perform integrative analyses across four age groups of the human lifespan (Supplementary Table S1). To avoid assumptions on the age patterns (e.g. linearity) we determined four age groups in the integrated datasets by grouping subjects into four approximate, quantile-based groups of age (Supplementary Table S3).

Heterogeneity of the overall cell type distribution

We calculated a normalized entropy-based metric\(^18\) to describe the heterogeneity of the proportions of all cell types of each sample. This metric is higher in samples with more heterogenous (uniform) proportions of cell types and lower in samples that are more enriched in specific cell types.

Effect of age and sex on cell type distribution

We analyzed the effect of age and sex on the proportions of all the 13 cell types at subject level, using a Bayesian multinomial regression model.

Analysis of the hierarchy of peripheral immune compartments

We used K2Taxonomer (v1.0.5)\(^19\) to perform top-down partitioning of cell types based on the relative similarity of their transcriptomic profiles and generate a data-driven hierarchy of peripheral immune compartments. Within each level of the hierarchy, we tested the significance of the differences in the cell type diversity statistics between age groups using ANOVA with p-value threshold of 0.05.

Validation of cell type patterns

We calculated cell type proportions of the NECS samples using a spectral flow cytometry panel that included specific markers for all identified populations. We integrated these new data with a mass cytometry dataset previously published as a part of the immune monitoring study across multiple age groups.\(^5\)

Cell type specific differential gene expression analysis

To investigate the cell type specific differences across age groups, we analyzed the log-normalized gene expression abundances using a Bayesian mixed effects regression model of age adjusted by sex, ethnicity, batches of data. We included a random effect to account for the grouping of cells into subjects. We calculated the FDR for each age group comparison based on the Benjamin and Hochberg correction for multiple testing across all genes tested. We selected significantly different genes based on FDR < 0.05 and fold change (FC) of at least 10 percent.

Bulk level differential gene expression analysis

We performed differential gene expression analysis at the bulk level between age groups using DESeq\(^20\) of genes expressed in at least 50% of the smallest cell type population. Significant genes at the bulk level were evaluated according to a fold change of at least 50% and FDR < 0.05.

Role of funding source

No funding sources had any role in the study design, data collection, data analyses, interpretation, writing of report, and decision to submit the paper for publication.

Results

The peripheral immune landscape of centenarians at single cell resolution

Using the 10X Genomics platform for droplet-based Cellular Indexing of Transcriptomes and Epitopes sequencing (CITE-seq),\(^7\) we simultaneously profiled the transcriptome-wide expression and the 10 marker cell surface-level protein expression of 16,082 PBMCs from seven centenarians (100–119 years of age) and two younger individuals with no known history of familial longevity (20–59 years of age). All nine subjects were of
European descent and from the New England Centenarian Study (NECS), with an average sample capture of 1833 cells (Table 1). We accounted for technical differences and integrated multiple samples using Harmony.\(^1\) (Supplementary Figure S1). We then performed Louvain graph-based clustering to group cells into populations of similar expression profiles, and used Uniform Manifold Approximation and Projection (UMAP)\(^2\) of cell expression profiles to visualize the single cells in a 2-dimensional space.

We leveraged the 10 immune cell-surface protein expression markers to identify the major lymphocyte and myeloid cell types (Supplementary Figure S2, Supplementary Table S1), with subtypes subsequently characterized by transcriptional immune cell signatures previously characterized in human peripheral blood\(^3\) and fetal liver\(^4\) (Supplementary Figures S3–S7, Supplementary Table S2). This approach identified 11 immune cell types that included major lymphocyte populations: CD4\(^+\) T cells (CD4TC) with noncytotoxic naive and memory subtypes (nCD4TC, mCD4TC) and cytotoxic subtype (cCD4TC), CD8\(^+\) T cells (CD8TC) with cytotoxic subtype (cCD8TC), B cells (BC) with naive and memory subtypes (nBC and mBC), and Natural Killer cells (NK) (Supplementary Figures S2 and S3). In addition, we identified major myeloid populations: monocytes with CD14\(^+\) and CD16\(^+\) subtypes (M14 and M16) and dendritic cells (DC) with myeloid and plasmacytoid subtypes (mDC and pDC) (Supplementary Figures S2 and S3).

**Centenarians display alterations in immune cell repertoire in comparison to younger age groups**

To characterize peripheral immune cell type composition and gene expression profiles across the human lifespan, we integrated our data with two publicly available PBMC datasets of aging and longevity that include subjects of European and Japanese descent\(^6,25\) (Fig. 1A and B, Supplementary Figure S8). The integration of these datasets with our novel NECs dataset produced a total of 102,284 cells from 66 individuals across four age groups: 12 subjects of younger age (20–39 years), 26 subjects of middle age (40–59 years), 14 subjects of older age (60–89 years), and 14 EL subjects (100–119 years) (Supplementary Tables S3 and S4). Technical differences between datasets were accounted for using Harmony.\(^1\) This analysis identified two additional cell types: plasma B cells (pBC) and gamma delta T cells (gdTC) for a total of 13 cell types across the datasets (Fig. 1B).

Fig. 2A displays the observed proportions of the 13 immune cell types in the 66 subjects, stratified by age. Notably, the EL group was characterized by an increase in the proportion of myeloid cells and a reduction in the proportion of lymphocytes: The ratio of myeloid cells to lymphocytes was approximately 13.8/86.2% across the three younger age groups (20–89 years) and shifted to 25.2/74.8% in the EL group (Fig. 2A, Supplementary Table S5). This shift in myeloid cells and lymphocytes is an expected trend in aging.\(^5\) The barplot in Fig. 2A suggests that the distribution of proportions of the 13 immune cell types becomes more uniform in the EL group. To formalize this observation, we next calculated the cell type diversity statistic of each sample.\(^8\) This statistic is essentially an entropy-based score that we introduced to summarize the vector of proportions of cell types in a sample. The score is normalized between −1 and 0, with −1 corresponding to the case of a single cell type being present and 0 corresponding to the case when all cell types are present in the exact same proportion and therefore more uniform. The analysis of the 13 immune populations showed a trend towards an increase of the cell type diversity statistic in EL compared to younger age groups, although this difference was not statistically significant (F-test, p-value = 0.7231) (Supplementary Figure S9). When nCD4TC and mCD4TC were combined as noncytotoxic CD4\(^+\) T cells, the increase of the cell type diversity statistic in EL compared to younger ages was statistically significant (F-test, p-value = 0.0001875) (Fig. 2B, Supplementary Table S6). This analysis formalized the observation that the composition of PBMCs in EL subjects becomes more heterogeneous.

To estimate the proportions of cell types by age and sex, we next analyzed the observed cell type proportions from all 66 subjects using a Bayesian multinomial regression (See Methods, Supplementary Tables S7–S10). This analysis produced age and sex specific estimates of each of the 13 cell type proportions across the four age groups (Fig. 2C and D). The estimates suggest that there are three main groups of immune cells based on their distributions at different ages and EL: 1) cell types whose proportions increase or decreases monotonically with age and EL (Aging-Related), 2) cell types whose proportions increase or decrease only in the EL group (EL-Specific), 3) cell types whose proportions increase or decrease with age, but these changes do not continue in the EL group (Aging-Specific).

<table>
<thead>
<tr>
<th>Sample ID</th>
<th>Age range</th>
<th>Sex</th>
<th>Age</th>
</tr>
</thead>
<tbody>
<tr>
<td>EL1</td>
<td>100-105</td>
<td>Male</td>
<td>57</td>
</tr>
<tr>
<td>EL2</td>
<td>105+</td>
<td>Female</td>
<td>97</td>
</tr>
<tr>
<td>EL3</td>
<td>105+</td>
<td>Female</td>
<td>103</td>
</tr>
<tr>
<td>EL4</td>
<td>100-105</td>
<td>Female</td>
<td>100</td>
</tr>
<tr>
<td>EL5</td>
<td>105+</td>
<td>Male</td>
<td>102</td>
</tr>
<tr>
<td>EL6</td>
<td>105+</td>
<td>Male</td>
<td>100</td>
</tr>
<tr>
<td>EL7</td>
<td>100-105</td>
<td>Female</td>
<td>100</td>
</tr>
<tr>
<td>YA1</td>
<td>105</td>
<td>Male</td>
<td>100</td>
</tr>
<tr>
<td>YA2</td>
<td>100-105</td>
<td>Female</td>
<td>100</td>
</tr>
</tbody>
</table>

Table 1: Table of demographic characteristics of the 7 centenarians included in this study.
We observed Aging-Related changes (i.e., change in both aging and EL) in cCD4TC, gdTC, and M14 populations. The estimated proportion of cCD4TC increased steadily with increasing age, representing 7.5% of PBMCs in males with EL and 6.0% of PBMCs in females with EL compared to less than 1% of the PBMCs in the younger age groups (Fig. 2C and D). This significant change in cCD4TC in centenarians compared to younger individuals is consistent with previous findings.16 We observed a similar aging-related change in the estimated proportions of gdTC and M14 (Fig. 2C and D).

Five lymphocyte and myeloid populations were observed to have EL-Specific changes (i.e., change only in EL) in nCD4TC, mCD4TC, pBC, mDC, and pDC. The lower frequency of nCD4TC and mCD4TC in EL is known.16 However, the estimated lower proportion of mDC that decreased to 0.9% of PBMCs in males with EL and 0.7% of PBMCs in females with EL compared to 1.3% in males and 1.1% in females in the younger age group has not been reported (Fig. 2C and D). Furthermore, we observed Aging-Specific changes (i.e., change in aging but not EL) in nBC, mBC, and M16 populations. The estimated proportion of nBC significantly increased to 6.2% of PBMCs in males with EL and 0.7% of PBMCs in females with EL compared to 1.3% in males and 1.1% in females in the younger age group compared to younger individuals is consistent with previous findings.16 We observed a similar aging-related change in the estimated proportions of gdTC and M14 (Fig. 2C and D).

Fig. 2: The immune landscape of peripheral blood cells from subjects with extreme longevity at single cell resolution. A. UMAP embeddings of PBMCs from all 66 subjects representative of the human lifespan from the integrated scRNA-seq datasets (NECS, PNAS, and NATGEN, labelled by the identified immune cell subtypes. B. UMAP embeddings of PBMCs from all 66 subjects representative of the human lifespan from the integrated scRNA-seq datasets (NECS, PNAS, and NATGEN, labelled by the four age groups.

We observed Aging-Related changes (i.e., change in both aging and EL) in cCD4TC, gdTC, and M14 populations. The estimated proportion of cCD4TC increased steadily with increasing age, representing 7.5% of PBMCs in males with EL and 6.0% of PBMCs in females with EL compared to less than 1% of the PBMCs in the younger age groups (Fig. 2C and D). This significant change in cCD4TC in centenarians compared to younger individuals is consistent with previous findings.16 We observed a similar aging-related change in the estimated proportions of gdTC and M14 (Fig. 2C and D).

Extreme longevity displays a shift in immune resilience strategy within lymphocyte and myeloid populations compared to younger age groups

To investigate whether the makeup within immune compartments also changed with age and EL and with possible effects on their biological functions and immune resiliency strategies, we next examined the make-up of various immune cell types within their myeloid and lymphocyte lineages. To this end, we generated a hierarchy of peripheral immune compartments, partitioned cell types, based on the gene expression profiles of the immune cell types using K2Taxonomer,17 and then calculated the average proportions of cell types within each level of the hierarchy (see Methods) (Supplementary Tables S12 and S13). Fig. 3 displays the results of this analysis and reveals specific changes of the composition of myeloid and lymphocyte compartments in the EL group as compared to younger age groups. In addition, we generated independent estimates of a large portion of the same cell counts in the NECS subjects using flow cytometry, and integrated the
data with publicly available mass cytometry data from Alpert et al.\textsuperscript{5} that included 39 younger individuals (mean age = 29) and 50 older individuals (mean age = 80). Supplementary Figure S11 contrasts the age trend of cell types based on the cytometry dataset to the original analysis based on scRNA-seq data. We discuss selected results next, with the complete analysis available in Supplementary Tables S14–S26.

The top of the hierarchy recapitulates the significantly larger proportion of myeloid cells (Myeloid) and smaller proportion of lymphocytes (Lymph) observed in centenarians’ PBMCs discussed earlier. The reduction of lymphocytes was validated by flow cytometry analysis of 7 centenarians and 3 younger age individuals from NECS, and 39 younger age and 50 older age individuals from Alpert et al.\textsuperscript{5} (Supplementary Figure S11, Supplementary Table S26). The analysis of the Myeloid compartment (left branch of the tree in Fig. 3) showed that centenarians’ Myeloid cells were mainly monocytes (Mono) rather than dendritic cells (DC) (EL Mono to DC ratio 94.58%/5.42%) compared to a lower fraction of Mono and higher fraction of DC observed in all other age groups (Mono to DC ratio 85.59%/14.41%). The difference in the two distributions was statistically significant, (F-test of cell type diversity statistic, p-value = 0.0001315) (Fig. 3, Supplementary Figure S10). We observed the increase in Mono compared to DC in EL based on the proportions measured by cytometric of EL and younger individuals from NECS (Supplementary Figure S11, Supplementary Table S26). This change in composition of DC has not been reported before.\textsuperscript{6,27} Although no additional changes were detected within Mono and DC, we observed similar trends in cell proportions measured by flow cytometry across age within Mono and DC (Supplementary Figure S11, Supplementary Table S26).

The analysis of the lymphocyte compartments (right branch of the tree in Fig. 3) showed that more than 70% of the lymphocyte and myeloid immune cell subtypes for each sample across the integrated scRNA-seq datasets: Younger Age, Middle age, Older age, and EL. B. Boxplot of subject specific cell type diversity statistic of 12 immune cell subtypes, grouped by age groups (F-test p-value = 0.0001875). C. Bar chart of the estimated proportions of the lymphocyte and myeloid immune cell subtypes in each age group, grouped separately for males and females. The proportions were estimated using a Bayesian multinomial regression model, and the average 1833 cells per sample allowed us to estimate the smallest proportion 0.3% with 95% confidence. D. Heatmap of the age coefficient comparing Middle, Older, and EL age groups to the Younger age group (right) and heatmap of the sex coefficient for each cell type comparing Females compared to Males. We calculated the Z-statistic and p-value of significance for each coefficient, represented with: * p < 0.05, ** p < 0.01, *** p < 0.001.
of centenarians’ lymphocytes were Cytotoxic (Cyt: 70.94% and NonCyt: 29.06%) compared to the younger age group (Cyt: 30.30% and NonCyt: 69.70%). This difference in distributions was only borderline statistically significant in our analysis (F-test of cell type diversity statistic, p-value = 0.05032) but consistent with results reported by other investigators. In addition, the hierarchy showed further sub-types of Cyt and NonCyt that had unique composition in centenarians. For example, centenarians’ NonCyt had a significant lower proportion of CD4TC (81.55%) and larger proportion of BC (18.45%) compared to all younger age groups (CD4TC: 90.86% and BC: 9.41%, F-test, p-value = 0.003313) (Fig. 3, Supplementary Figure S10). We observed a similar shift between CD4TC and BC in EL compared to younger individuals from NECS based on proportions measured by flow cytometry, but was not observed in the younger and older individuals from Alpert et al. (Supplementary Figure S11, Supplementary Table S26). In previous studies, both noncytotoxic CD4TC and BC have been reported to decrease in PBMCs of long-lived individuals, but this shift between the two cell types has not been previously observed.

The composition of CD4TC in centenarians was characterized by an expansion of mCD4TC (65.72%) and a reduction of nCD4TC (34.28%) compared to all younger age groups (mCD4TC: 39.46% and nCD4TC: 60.54%, F-test, p-value = 0.02852) (Fig. 3, Supplementary Figure S10). The composition of BC in the EL group had a similar shift from naive (nBC: 41.06%) to memory cells (mBC: 58.94%) but the change did not reach statistical significance (F-test, p-value = 0.7168) (Fig. 3, Supplementary Figure S10). In addition, we observed similar shifts between nBC and mBC between EL and younger individuals of NECS based on flow cytometry analysis (Supplementary Figure S11, Supplementary Table S26).

In summary, using the hierarchy of peripheral immune compartments, we identified major shifts in the makeup of centenarian PBMCs compared to those of younger ages within the myeloid and lymphocyte lineages that were obscured globally in terms of all 13 cell types together.

Centenarians display unique transcriptional profiles associated with extreme longevity

The previous two analyses characterized the makeup of centenarian PBMCs globally and within their myeloid and lymphocyte lineages in terms of proportions of the various cell types. We next examined their expression profiles relative to younger age groups. For each cell type, we performed an analysis to discover genes with differential expression as a function of age and/or EL, which identified 151 genes with age- or EL-associated differential expression in at least one cell type (Fig. 4, Supplementary Figure S10). The number of significantly differentially expressed genes varied by cell types and comparison groups (Fig. 4A): on average, the comparison of expression profiles of cell types in the middle vs. the younger age group produced a smaller number of differentially expressed genes (i.e., fewer differences) than the comparisons of the older vs. younger age and the EL vs. younger age group. Fig. 4B shows clear differential gene expression patterns for all cell type-
specific signatures across each age group comparison. Of these cell type-specific signatures, we identified 20 differential genes that have been previously identified to change with age in transcriptional studies of aging in PBMCs, including leucine-rich repeat neuronal protein 3 (LRRN3), lymphoid enhancer-binding factor-1 (LEF1), and Cathepsin H (CTSH).

Further examination of the age-specific expression changes identified three main patterns that are summarized in Fig. 4C–E, and closely mirror the changes in cell type composition described above. These patterns include 1) genes whose expression increase or decrease monotonically with age and EL (Aging-Related) with at least a nominal significance, 2) genes whose expression increase or decrease only in the EL group (EL-Specific), 3) genes whose expression increase or decrease with age, but these changes do not continue in the EL group (Aging-Specific).

We identified 35 genes with the Aging-Related pattern (i.e., change in both aging and EL) of differential expression across the various immune cell types.

Fig. 4C shows selected examples of genes in the first group and Supplementary Figures S12–S20 include more examples. The set includes genes involved in DNA damage response such as serine/threonine-protein kinase 17A (STK17A) in nCD4TC (Fig. 4C). STK17A is a positive regulator of apoptosis and a variant near STK17A was previously reported to be associated with longevity in WGS analysis. In addition, STK17A has been previously identified to change in age in T cells of PBMCs. Furthermore, two genes previously reported to increase in expression with age in peripheral blood including in T cells were in group 1: LRRN3 and LEF1 in nCD4TC (Fig. 4C).

We identified 25 genes with an EL-Specific pattern (i.e., change only in EL) in several cell types (Fig. 4D, Supplementary Figures S12–S20). Fig. 4D shows examples of genes that appear to be only expressed in immune cells from the EL group. An example of a gene only expressed in nCD4TC of the EL group is S100 protein-coding gene, S100A4 (Fig. 4D). S100 proteins have been implicated in aging-related diseases such as...
Alzheimer’s disease as well as longevity. In addition, S100A4 has been previously implicated in age-related changes in T cells of PBMCs. Group 2 also included the HLA class II histocompatibility antigen gene HLA-DPA1 in nCD4TC (Fig. 4D) and cCD8TC (Supplementary Figure S14), and interferon-induced transmembrane proteins IFITM3 in NK (Fig. 4D) and in mDC (Supplementary Figure S18), and IFITM2 in cCD8TC (Supplementary Figure S12). These major histocompatibility and interferon-related genes are involved in antigen presentation and activation of immune response pathways.

We identified 26 genes with the Aging-Specific pattern (i.e., change in aging but not in EL), with expression levels that change with age but not in EL across various cell populations. Fig. 4E shows examples of 3 genes and additional examples are in Supplementary Figures S12–S20. This set includes genes that respond to oxidative stress including sestrin 3 (SESN3) in nCD4TC (Fig. 4E) and mDC (Supplementary Figure S13), argonaute RISC catalytic component 3 (AGO3) in nBC, and cytochrome C oxidase assembly factor (COX16) in mBC (Fig. 4E). SESN3 is part of the sestrin family of stress-induced metabolic proteins and is stimulated in response to oxidative stress/damage by FOXO3, a transcription factor associated with longevity, while AGO3 is part of the argonaute (AGO) family of proteins that have been previously implicated in aging and oxidative stress-induced senescence.

In addition, we found 62 significant cell type specific differential genes between older and EL age groups, including S100A4 which is significantly lower in older age compared to EL in nCD4TC (Supplementary Table S28). This is consistent with the significant increase observed in EL compared to younger age (Fig. 4D). Furthermore, we examined the gene expression changes between females and males. We detected 28 significant differential genes between females compared to males across cell types (Fig. 4A, Supplementary Table S27). Of the 28 genes, six genes were differentially expressed across most or all cell types. This set of genes included X-chromosome and Y-chromosome linked genes such as ribosomal protein S4 X-linked (RPS4X), X inactive specific transcript (XIST), and eukaryotic translation initiation factor 1A X-linked (EIF1AX), ribosomal protein S4 Y-linked (RPS4Y1), DEAD-box helicase 3 Y-linked (DDX3Y), and eukaryotic translation initiation factor 1A Y-linked (EIF1AY) (Supplementary Table S27). With the exception of EIF1AX, these genes have been previously reported to have gene expression differences based on sex across immune cell types in PBMCs. In addition, CD99, involved in T cell activation and pro-inflammatory mechanisms, had significantly lower expression in females nCD4TC than males.

In addition to cell type specific expression profiles, we analyzed gene expression aggregated over different cell types (see Methods). This analysis identified a greater number of genes with significantly different expression in the EL v. younger age group comparison (387 genes) compared to middle v. younger age group (0 genes) and older v. younger age group (3 genes) (Supplementary Figure S21, Supplementary Tables S29). Among the 387 significant differential genes identified in EL v. younger age, we found 136 genes over-expressed and 251 under-expressed in the EL group compared to the younger group (Supplementary Figure S21, Supplementary Tables S29). In addition, of the 387 significant genes identified in EL v. younger age, we identified 164 genes previously identified to change in expression with age, including LEF1 and LRRN3 that we also identified at the single cell level as well as CD28 antigen molecule.

**Discussion**

**Overview of main results**

Using a multi-modal, single cell approach, we generated cell composition and transcriptional profiles from the PBMCs of 7 centenarians using CITE-seq. We integrated this novel data set with publicly available scRNA-seq datasets of aging and longevity across the human lifespan to characterize cell type composition and gene expression profiles unique to centenarians. We observed substantial changes in the composition of immune cells with age, including novel changes in myeloid cell types: M14, M16, mDC, and pDC. We also conducted a novel analysis of a data-driven hierarchy of peripheral immune compartments, which revealed previously undetected changes in the composition of T cells and B cells in centenarians and we validated several of these changes using orthogonal approaches (flow cytometry). Based on gene expression changes, we identified cell type-specific transcriptional signatures of extreme longevity that include aging-related changes as well as unique gene changes in the immune profiles of centenarians.

**Cell type composition profiles based on the total number of PBMC populations**

The peripheral blood immune cell repertoire of individuals is known to change with age. Previous transcriptional studies have shown decreases in lymphocytes and increases in myeloid cells with age, which we also observed in the peripheral blood of centenarians (Fig. 2). However, in addition to these common changes across aging, our analysis identified patterns of immune cell profiles and compositional alterations that are unique to centenarians. We observed expected shifts in the composition of centenarians’ PBMCs from non-cytotoxic (e.g., nCD4TC and nCD4TC) to cytotoxic lymphocytes (e.g., cCD4TC) that have been observed previously in studies of human longevity. Similarly, the decrease of nBC with aging and longevity has also been reported previously. However, we also discovered novel compositional
patterns of extreme old age including aging-related changes (e.g. a significant increase of M14 in older age that continues in the EL group), EL-specific changes (e.g. mDC and pDC display no significant change among the three younger age groups but a unique, significant decrease occurs in EL), and aging-specific changes independent of EL (e.g. a significant increase of M16 in older age that then decreases in the EL age group) (Fig. 2). The extent to which these unique patterns in centenarians are the drivers of extreme longevity or just the consequence of having reached an extreme old age remains an open question, since not everything we see in centenarians is necessarily important to reach extreme old ages. Additional data are needed to understand the effect of these patterns on human longevity.

Cell type composition profiles within peripheral cell compartments

Utilizing the traditional method for characterizing cell type composition profiles based on the total number of PBMC populations, we identified novel compositional changes of extreme old age (Fig. 2). However, this analysis provides limited insight into immune cell type change within the lymphocyte and myeloid compartments. We used a novel data-driven approach to create a hierarchy of peripheral immune compartments. The analysis summarized in Fig. 3 shows, for example, that the proportion of lymphocytes in the PBMCs of centenarians decreases compared to younger age groups, but a significant change in composition also occurs. Specifically, centenarians' lymphocytes are characterized by an almost 50% decrease of NonCyt, which become enriched for BC that are themselves enriched for mBC. Notably, CD4TC and BC have a significant role in the immune system’s response to infection.47 BC are associated with the antibody-mediated immune response that triggers a quick response against pathogens, while T lymphocytes such as CD4TC are associated with cell-mediated immunity that develops at a slower rate. Studies have also found crosstalk between BC and CD4TC, showing their co-dependence in protective immune response.47 The shift from CD4TC to BC suggests that centenarians develop a more immediate immune response to infections. We also observed a significant shift from naive to memory subtypes within CD4TC and, to an extent, within BC suggesting that centenarians did not escape infection but experienced a greater exposure to infections and were able to develop robust responses to them. Previous studies have also shown associations between the capacity to control inflammation and preserving immunocompetence with longevity as an immune resilience phenotype.48,49 It is possible that the unique make up of immune cells we observed in centenarians may represent an adaptation of their immune system or a compensatory mechanism to the loss of key immune cell types. Interestingly, in almost all compartments displayed in Fig. 3, the EL group was characterized by a lesser skewed distribution of cell types compared to younger age groups. A more heterogeneous distribution of immune cells may be the driver of their immune resiliency. Compositional heterogeneity may reflect immunocompetence as a dynamic balance or homeostasis.49 Reciprocally, immune imbalance is often characteristic of suboptimal responses to infections (e.g., COVID-1950) or compensation by less effective or exhausted cell mediators (e.g., NK cells51) that compromise health span.

Gene expression profiles

Our analysis identified three patterns of age-related changes: monotonic changes across the lifespan, age-related changes that are absent in the EL group, and changes that are unique to centenarians. Interestingly, we noticed similar patterns in the serum proteome of centenarians.52 By comparing the serum proteome of centenarians to septuagenarians, we discovered aging-related protein signatures as well as a protein signature that was unique to centenarians. We also discovered an age-specific protein signature that was not extended as expected in centenarians. Without the additional data of blood expression profiling of centenarians when they were of younger age, we cannot determine whether these EL specific patterns are the drivers to extreme human longevity or the effect of extreme old age. However, some of the genes with differential expression in the EL group have been linked to aging and longevity studies. For example, the expression of a variant of STK17A that we found associated with age in nCD4TC (Fig. 4C) was higher in centenarians.53 STK17A is involved in DNA damage response, positive regulation of apoptosis, and mitochondrial and metabolic regulation of reactive oxygen species (ROS). This association is consistent with results from previous studies that correlated DNA damage repair mechanisms to aging and longevity.25,40,41 S100A4, part of the S100 family of calcium-binding proteins, showed an EL specific change in nCD4TC (Fig. 4D). S100 proteins such as S100A13 have been implicated in their role in longevity, including an association with APOE genotypes in centenarians.54 In addition, the S100 family of proteins are associated with inflammatory pathways in the brain connected to aging-related diseases such as Alzheimer’s disease.55 In CD4+ T cells, S100A4 was found to be higher expressed in older mice compared to younger mice, and involved in inflammation and activation.56 A recent study discovered an S100 protein to be a critical regulator of hematopoietic stem cell renewal through mitochondrial metabolic regulation and function.57 In rats, recombinant S100A4 demonstrates an anti-apoptotic function in response to oxidative stress injury.58 Other transcriptional signatures that we identified in our analysis, such as SESN3 and AGO3 (Fig. 4E), are involved in DNA damage response and...
mitochondrial and metabolic regulation activated in response to oxidative stress.\textsuperscript{13,14} Sestrins such as SESN3 are highly conserved stress inducible proteins that protect the immune system in response to DNA damage and oxidative stress.\textsuperscript{14} More specifically, the induction of SESN3 in response to oxidative damage is activated by FOXO3,\textsuperscript{14} a transcription factor associated with longevity.\textsuperscript{12} AGO3 is part of the AGO family of proteins involved in miRNA association and stability that has been implicated in oxidative stress-induced senescence and aging.\textsuperscript{13} In the immune system, mitochondrial regulation plays a role in immune cell transcription and activation\textsuperscript{35,36} and may promote longevity.\textsuperscript{13} In addition, the decline of mitochondrial quality and activity is implicated in oxidative stress-induced senescence and aging.\textsuperscript{35} In the immune system, mitochondrial and metabolic regulation and function should be further investigated. We note that our signatures of aging derived from “pseudo-bulk” data closely match previously published signatures of aging derived from bulk data\textsuperscript{28} and included a much larger number of differentially expressed genes compared to the analysis at single cell level. The difference between the bulk and single-cell signatures could point to the fact that much of the observed differential expression in bulk RNA is driven by differences in cell composition rather than by differences in within-cell type expression.

Caveats and limitations and future directions

This study has several limitations, particularly the cross-sectional nature of the data and the small sample size. We integrated our data with multiple single cell datasets to increase sample size, but we intentionally adopted a conservative approach to identify cell type specific signatures across age groups. Larger studies of centenarians will be needed to detect robust transcriptional changes that characterize EL. Although we identified a small set of transcriptional signatures, we were still able to identify patterns of EL that have been discovered in previous studies including EL specific changes. In addition, the compositional and gene expression changes that we observed in centenarians displayed not only EL specific changes but also age-related changes. How EL differs from regular aging remains unclear, and more investigation and future studies will be required to elucidate this difference and investigate the mechanisms behind the patterns observed in extreme old age. Access to the peripheral blood of centenarian offspring and studying longitudinal changes in PBMC populations may help to better define immunocompetence causal drivers of the beneficial health outcome observed in EL.

Conclusion

Overall, these findings display age-related changes in composition and transcription in both lymphocyte and myeloid cell types that collectively reflect immunocompetent profiles that may in part account for centenarians’ ability to reach extreme ages. The extent to which some of the unique compositional and transcriptional patterns we identified in centenarians’ PBMCs are the drivers or markers of extreme old age remains an open question. To our knowledge, this is the first study to define cell compositional and transcriptional signatures of EL across immune cell types in peripheral blood. This study provides a foundation and resource to explore immune resilience mechanisms engaged in exceptional longevity.

Contributors

PS, SM, and GM conceptualized the study. TD prepared all scRNA-seq libraries. CVM performed processing and generation of scRNA-seq count data. ER performed K2Taxonomer and the generation of the peripheral immune hierarchy. ACB generated flow cytometry data and analyses. TK performed all other computational analyses. TP and GM organized collection of human samples. MM, TP, and SA provided insights and advice. TK, TD, ACB, SM, GM, and PS interpreted the data and wrote the manuscript with the feedback of all the authors. All authors read and approved the final version of the manuscript. TK, ER, ACB, SM, and PS have accessed and verified the underlying data.

Data sharing statement

The data that support these findings are publicly available and were accessed from several repositories. NATGEN single cell expression data and subject level data were publicly available as referenced in\textsuperscript{15}: https://molgenis58.target.rug.nl/scRNA-seq/. PNAS single cell expression data and subject level data was available as referenced in\textsuperscript{15}: http://gerg.gsc.riken.jp/SC2018/. NECS will be available from Synapse (https://adknowledgeportal.synapse.org/Explore/Projects/DetailsPage?Grant%20Number=UH2AG064704). All scripts to reproduce analyses and figures reported in this paper are available on github (https://github.com/Integrative-Longevity-Omics/sc_pbmc_centenarians).

Declaration of interests

TK, TD, MM, SM, PS, GM, SA, and TP report grants from National Institute on Aging, during the conduct of the study. ACB reports grants from NIH Instrumentation grant: S10 OD021587, during the conduct of the study. CVM and ER declare no competing interests.

Acknowledgements

TK, SM, PS, GM, SA, and TP are supported by NIH-NIA UH2AG064704 and U19AG023122. MM and PS are supported by NIH NIA Pepper center: P30 AG031679-10. This project is supported by the Flow Cytometry Core Facility at BUSM. FCCF is funded by the NIH Instrumentation grant: S10 OD021587.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.ebiom.2023.104514.

References


