

# Reading the immune clock: a machine learning model predicts mouse immune age from cellular patterns

Received: 11 June 2025

Accepted: 28 November 2025

Published online: 10 December 2025

 Check for updates

Hyun Bo Sim<sup>1,8</sup>, Ji-Hun Jang<sup>2,8</sup>, Seul-Ki Mun<sup>1,3</sup>, Sonny C. Ramos<sup>1</sup>, Dae-Han Park<sup>1</sup>, Yu-Jeong Choi<sup>1</sup>, Ji Yeon Han<sup>1</sup>, Ju-Bin Lee<sup>1</sup>, Ho Seok Chung<sup>4</sup>, Kyung-Bok Lee<sup>5</sup>, Dong-Jo Chang<sup>6</sup>✉, Seung-Hyun Jeong<sup>6,7</sup>✉ & Jong-Jin Kim<sup>1</sup>✉

Aging is accompanied by profound alterations in the immune system; yet, an accurate prediction of immunological age remains challenging. While transcriptomic approaches have yielded insights into immune aging, protein-level profiling and machine learning-based prediction tools remain underdeveloped. Here, we employ mass cytometry to analyse murine splenic CD45<sup>+</sup> immune cells across various age groups, profiling the expression of 30 protein markers and monitoring age-related immune changes. By analysing six major immune subsets (CD8<sup>+</sup> T cells, CD4<sup>+</sup> T cells, B cells, conventional type 1 and type 2 dendritic cells, and macrophages), we extract 103 molecular features and train a machine learning model using support vector regression (SVR) to predict immunological age. The model demonstrates robust generalizability by accurately predicting age in independent, test samples that were not used during model training. Furthermore, we confirm the robustness of our model using an obese mouse model, which exhibits metabolic dysfunction-associated immune senescence. Thus, our findings establish a robust framework for predicting immune-aging based on multidimensional protein expression data and machine learning. This tool enables quantitative assessment of immune aging and demonstrates strong translational potential for identifying obesity- and disease-related immune senescence.

Modern advancements in medical care and living standards have contributed to a dramatic increase in life expectancy, positioning aging as a central concern in public health and biomedical research<sup>1</sup>. By the mid-21st century, 16% of the global population is projected to be over the age of 65, equaling the proportion of children under the age of 12<sup>2</sup>. However, the biological manifestation of aging varies substantially among individuals, with an increasing divergence

between chronological age and biological function over time<sup>3–5</sup>. Thus, assessing biological age, rather than relying solely on chronological age, is essential for accurately evaluating health status and disease risk. Aging is characterized by nine canonical hallmarks, including genomic instability, telomere attrition, epigenetic alterations, mitochondrial dysfunction, and a decline in intercellular communication, which collectively form the molecular framework of

<sup>1</sup>Department of Biomedical Science, Suncheon National University, Suncheon, Republic of Korea. <sup>2</sup>College of Pharmacy, Chonnam National University, Gwangju, Republic of Korea. <sup>3</sup>R&D Team, Suncheon Research Center for Bio Health Care, Suncheon, Republic of Korea. <sup>4</sup>Department of Urology, Chonnam National University Medical School, Gwangju, Republic of Korea. <sup>5</sup>Center for Bio-imaging & Translational Research, Korea Basic Science Institute (KBSI), Cheongju, Republic of Korea. <sup>6</sup>College of Pharmacy and Research Institute of Life and Pharmaceutical Sciences, Suncheon National University, Suncheon-si, Republic of Korea. <sup>7</sup>College of Pharmacy, Suncheon National University, Suncheon-si, Jeollanam-do, Republic of Korea. <sup>8</sup>These authors contributed equally: Hyun Bo Sim, Ji-Hun Jang. ✉e-mail: [djchang@scnu.ac.kr](mailto:djchang@scnu.ac.kr); [jeongsh@scnu.ac.kr](mailto:jeongsh@scnu.ac.kr); [kimjj@scnu.ac.kr](mailto:kimjj@scnu.ac.kr)

biological aging<sup>6,7</sup>. Traditionally, biological age has been assessed by detecting age-associated changes in blood components such as nucleic acids, cytokines, growth factors, and extracellular vesicles<sup>8</sup>. Recent approaches have even employed routine blood parameters (e.g., hemoglobin, lymphocyte count, albumin, creatinine) to predict biological age through the integration of multi-system functional data<sup>4</sup>.

As aging progresses, the incidence of age-related diseases, such as cancer, cardiovascular disease, and neurodegeneration, rises sharply<sup>1</sup>. A key contributor to this decline in physiological resilience is inflammation, a state of low-grade chronic inflammation driven by proinflammatory cytokines (e.g., IL-1 $\beta$ , IL-6, TNF, IL-8)<sup>9</sup>. This phenomenon arises partly from the accumulation of senescent cells that evade immune clearance and persistently secrete SASP (senescence-associated secretory phenotype) factors<sup>10,11</sup>. Although the precise mechanism of senescent cell accumulation remains unclear<sup>12</sup>, it is evident that immune cell function itself is altered with age. Numerous studies have documented age-related phenotypic and functional changes in immune subsets across both mice and humans<sup>13–17</sup>. For instance, aging B cells are prone to autoantibody production, perpetuating a feedback loop that reinforces immune senescence<sup>18,19</sup>. Aging T cells undergo replicative senescence due to telomere shortening, resulting in reduced cell numbers and impaired functional capacity<sup>20</sup>. In particular, CD8<sup>+</sup> T cells upregulate exhaustion markers such as TOX, contributing to chronic inflammation through a hyperactivated, dysregulated immune state<sup>21,22</sup>.

Recent single-cell transcriptomic analyses have revealed that aging alters immune cell landscapes in complex and coordinated ways<sup>14,17,23–27</sup>. Notably, specific immune subsets, such as age-associated B cells (ABC) and CD8<sup>+</sup> Taa (TOX<sup>+</sup>PD-1<sup>+</sup>GZMK<sup>+</sup>), accumulate with age. ABCs, characterized by CD11b<sup>+</sup>, CD172a<sup>+</sup>, and Zbtb32<sup>+</sup> expression, are thought to originate from migrating B-1 cells and exhibit impaired BCR and CD40 responses, while promoting Th17 differentiation<sup>28–30</sup>. CD8<sup>+</sup> Taa cells, although resembling exhausted T cells, uniquely secrete GZMK, a granzyme implicated in amplifying SASP responses and accelerating tissue senescence in synergy with IFN $\gamma$ <sup>28,31–35</sup>.

Despite these insights, a comprehensive understanding of immune aging remains elusive due to its complex, multidimensional nature. The rapid accumulation of high-dimensional immunological datasets has prompted the integration of machine learning (ML) approaches to extract interpretable patterns. ML offers strong predictive power and is increasingly applied in genomics, systems biology, and aging research<sup>36</sup>. Supervised and unsupervised deep learning methods are now capable of modeling nonlinear relationships and uncovering latent patterns across large datasets<sup>36–38</sup>. In aging biology, ML has been used to construct age predictors based on DNA methylation patterns<sup>7,39–42</sup> and proteomic signatures<sup>43</sup>. Immune aging scores derived from cytokine and chemokine profiles can forecast mortality and age-related disease risk<sup>13</sup>. Meanwhile, single-cell ML approaches have begun to reveal interindividual variability in immune aging trajectories<sup>37</sup>. Nevertheless, few studies have applied ML to deeply profile immune aging using protein-level expression patterns in defined immune cell subsets, particularly in relation to age-accelerating conditions.

In this study, we address this gap by applying high-dimensional mass cytometry (CyTOF) to comprehensively profile murine splenic CD45<sup>+</sup> immune cells across five life stages—corresponding to adolescence, early adulthood, middle age, and advanced age in humans. Unlike conventional flow cytometry, CyTOF enables simultaneous quantification of dozens of protein markers with minimal spectral overlap and high sensitivity. It is therefore ideally suited for characterizing immune heterogeneity and identifying rare cell populations at the single-cell level.

Using this approach, we systematically map age-related shifts in immune cell composition and molecular signatures and establish a

predictive machine learning model that infers immune age from expression profiles. The model is validated both internally and through simulated external datasets and demonstrates robustness when applied to an obesity model, capturing immune age acceleration under metabolic stress. These findings establish integrative cytometry and computational modeling as effective approaches for quantifying immune aging and underscore their potential applications in monitoring age-related and metabolic diseases.

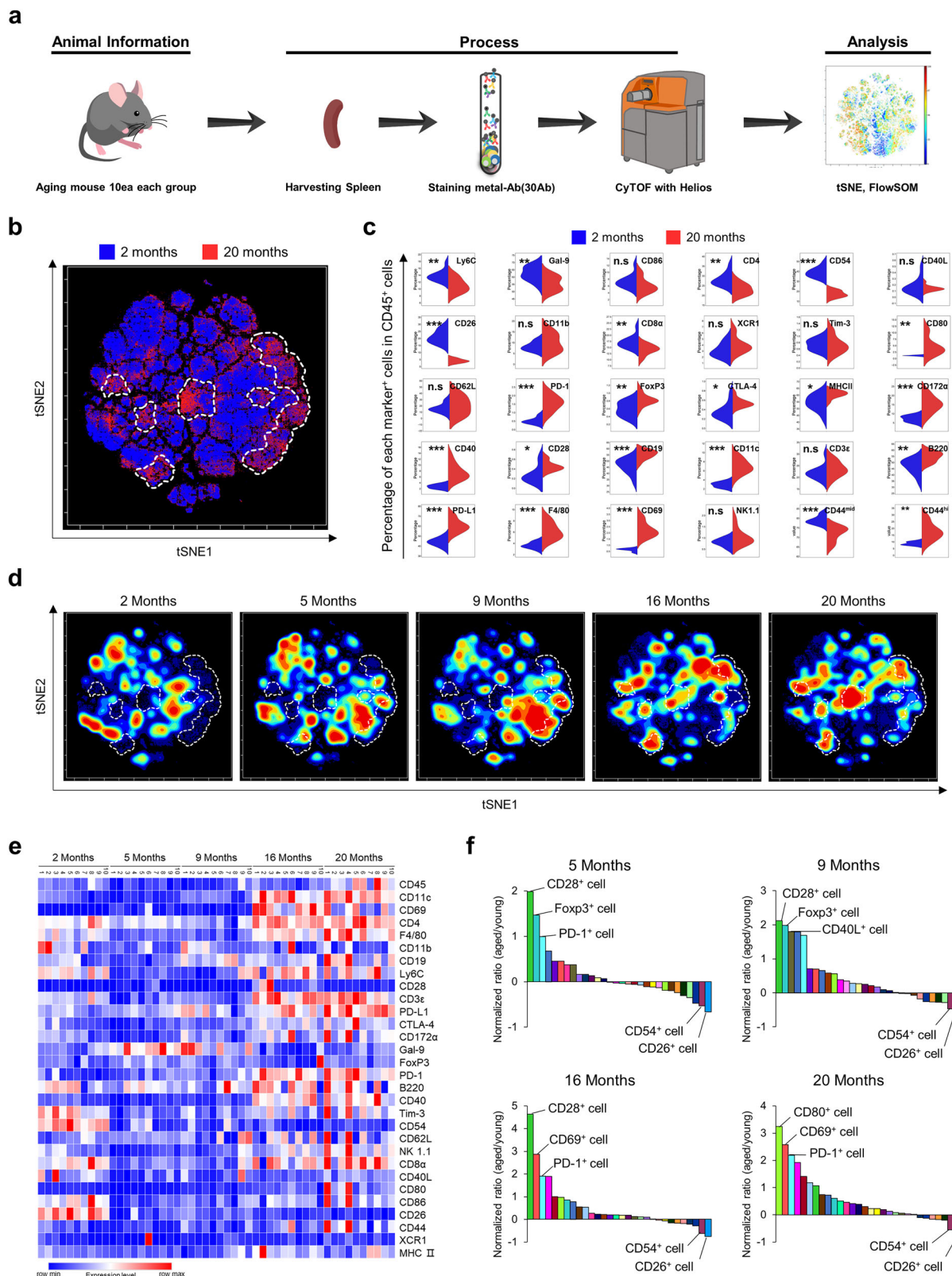
## Results

### Age-associated remodeling of splenic immune cell phenotypes and marker expression

To investigate systemic immune alterations during aging, we performed high-dimensional analysis using CyTOF on spleen-derived CD45<sup>+</sup> immune cells stained with a panel of 30 antibodies (Fig. 1a). t-SNE maps were generated using pooled data from 2-month-old ( $n=10$ ) and 20-month-old ( $n=10$ ) mice, each comprising 50,000 CD45<sup>+</sup> events. The map from young mice (blue) was overlaid on that from aged mice (red), and areas enriched in the 20-month group were marked with white dotted lines (Fig. 1b). Next, we compared the proportion of 30 marker<sup>+</sup> cells in the two age groups. Fifteen immune populations (CD28<sup>+</sup>, CD69<sup>+</sup>, CD40<sup>+</sup>, PD-1<sup>+</sup>, CD11c<sup>+</sup>, CD80<sup>+</sup>, CTLA-4<sup>+</sup>, F4/80<sup>+</sup>, CD172a<sup>+</sup>, PD-L1<sup>+</sup>, CD44<sup>hi</sup>, CD19<sup>+</sup>, B220<sup>+</sup>, MHC II<sup>+</sup>, FoxP3<sup>+</sup>) were significantly increased in aged mice, while seven populations (Ly-6C<sup>+</sup>, CD4<sup>+</sup>, Gal-9<sup>+</sup>, CD54<sup>+</sup>, CD26<sup>+</sup>, CD8 $\alpha$ <sup>+</sup>, CD44<sup>mid</sup>) were significantly decreased. The remaining eight populations (CD86<sup>+</sup>, CD40L<sup>+</sup>, CD11b<sup>+</sup>, XCR1<sup>+</sup>, Tim-3<sup>+</sup>, CD62L<sup>+</sup>, CD3e<sup>+</sup>, NK1.1<sup>+</sup>) did not show significant differences (Fig. 1c). Among them, CD80<sup>+</sup> cells showed the most prominent increase (4.10-fold), whereas CD26<sup>+</sup> cells exhibited the greatest reduction (5.11-fold). CD45<sup>+</sup> immune cell (Supplementary Fig. 1) distributions were visualized at five time points (2, 5, 9, 16, and 20 months) and mapped using density t-SNE plots (Fig. 1d). We observed different patterns in the number of CD45<sup>+</sup> cells at each time point. Notably, the white-dotted area in Fig. 1b was tracked across all ages, revealing age-associated remodeling of the splenic immune landscape. We further analyzed the expression profiles of the 30 markers across all individual samples using a heatmap, which clearly revealed age-dependent shifts in marker expression (Fig. 1e and Supplementary Figs. 2–6). Finally, we quantified the relative frequency of marker-positive cells over time by setting the 2-month level as a reference (Fig. 1f). CD28<sup>+</sup> cells consistently showed a 2.34–3.07-fold increase from 5 to 16 months. PD-1<sup>+</sup>, FoxP3<sup>+</sup>, and CD69<sup>+</sup> cells also displayed age-specific expansions, whereas CD26<sup>+</sup> and CD54<sup>+</sup> cells exhibited sustained reductions across all time points. Collectively, these findings demonstrate that splenic immune cells undergo substantial phenotypic remodeling with age, characterized by dynamic changes in both marker expression and population frequency.

### Subset-level remodeling of splenic immune populations during aging

To further investigate the age-related remodeling of splenic immune cell composition, we applied FlowSOM clustering to the t-SNE map of CD45<sup>+</sup> cells. FlowSOM is an unsupervised learning approach that clusters individual cells and merges similar clusters to define cell populations. We classified the splenic immune populations into six major metaclusters: CD8<sup>+</sup> T cells, CD4<sup>+</sup> T cells, B cells, conventional dendritic cells 1 (cDC1), cDC2, and macrophages (Fig. 2a). These subsets were validated by the expression of lineage-specific markers including CD3e, CD4, CD8 $\alpha$ , CD19, B220, Ly-6C, F4/80, CD11b, CD11c, XCR1, and CD172a (Fig. 2b). To identify the T cell lineage, we confirmed the CD3e expression in CD8<sup>+</sup> and CD4<sup>+</sup> T cells (Supplementary Fig. 7a) and validated it by comparing with other lineages, B cells (Supplementary Fig. 7b). Quantification of each population over time showed three distinct trends (Fig. 2c). First, the proportions of CD8<sup>+</sup> and CD4<sup>+</sup> T cells significantly declined with age, with a 3.17% and 5.32% reduction, respectively, at



20 months compared to 2 months. Similarly, cDC1 cells exhibited a consistent age-associated decrease. In contrast, cDC2 and macrophage populations remained relatively stable. Notably, B cells showed a progressive increase with age, with a significant 8.74% elevation at 20 months. Interestingly, at 5 months of age, macrophages, cDC1, and cDC2 displayed a transient decrease, a pattern that has also been reported in previous studies<sup>44</sup>. To dissect phenotypic heterogeneity

within the three dominant populations (CD8<sup>+</sup> T, CD4<sup>+</sup> T, and B cells), we selected their respective metaclusters and generated optimized SNE (opt-SNE) maps (Fig. 2d). Age-related expression changes of functional markers within each subset were visualized and quantified. For CD8<sup>+</sup> T cells, density-mapped opt-SNE projections revealed spatial and molecular shifts over time (Figs. 2e and S8a). Subtype analysis showed a sharp decline in CD26<sup>+</sup> cells after 2 months, while CD69<sup>+</sup> and CD44<sup>+</sup>

**Fig. 1 | Age-related molecular expression changes in the spleen.** **a** Schematic diagram of the experiment. Aging analyses were performed using mice aged 2, 5, 9, 16, and 20 months ( $n = 10$  per month). The following artwork was licensed and modified from iStockphoto.com/vector/mus-gm1068307150-285754008 (mouse icon) **b** t-SNE maps of 2- and 20-month-old mice (50,000 CD45<sup>+</sup> cells). White dashed lines indicate areas with increased cell numbers at 20 months. **c** We compared the proportion of indicated 30 cells at 2 and 20 months of age (CD28<sup>+</sup>, CD69<sup>+</sup>, CD40<sup>+</sup>, PD-1<sup>+</sup>, CD11c<sup>+</sup>, CD80<sup>+</sup>, CTLA-4<sup>+</sup>, F4/80<sup>+</sup>, CD172 $\alpha$ <sup>+</sup>, PD-L1<sup>+</sup>, CD44<sup>mid</sup>, CD44<sup>hi</sup>, CD19<sup>+</sup>, B220<sup>+</sup>, MHC II<sup>+</sup>, FoxP3<sup>+</sup>, Ly-6C<sup>+</sup>, CD4<sup>+</sup>, Gal-9<sup>+</sup>, CD54<sup>+</sup>, CD26<sup>+</sup>, CD8 $\alpha$ <sup>+</sup>, CD86<sup>+</sup>, CD40L<sup>+</sup>, CD11b<sup>+</sup>, XCR1<sup>+</sup>, Tim-3<sup>+</sup>, CD62L<sup>+</sup>, CD3e<sup>+</sup>, NK1.1<sup>+</sup> cells). All analyses

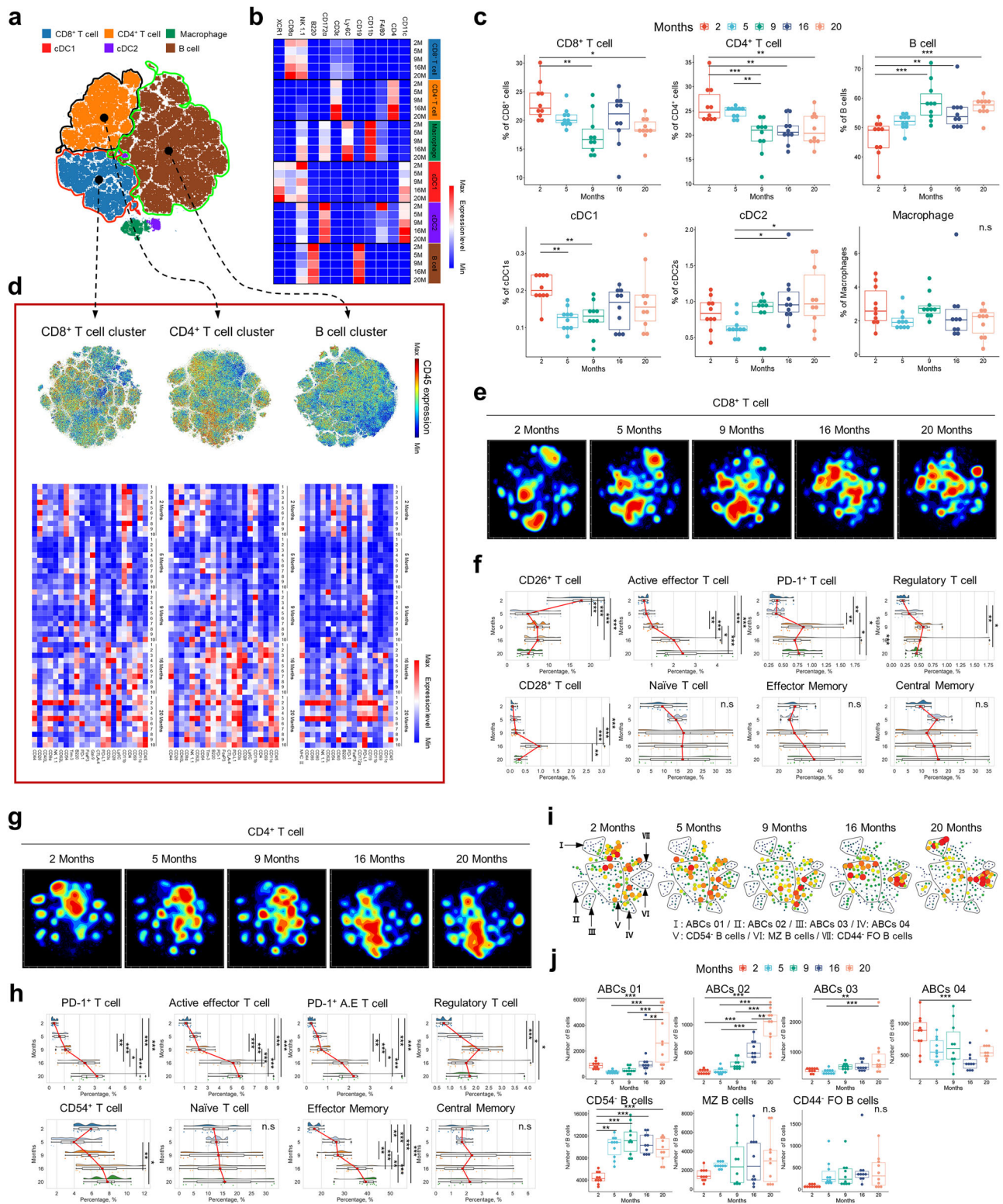
were performed by gating on CD45<sup>+</sup> cells. **d** Changes in the increased area (green dashed lines) at 16 months compared to 2 months were confirmed at 5 months (2, 5, 9, 16, and 20 months). **e** The expression levels of 30 molecules by group were compared using a heatmap. **f** Normalized to the expression level at 2 months, the expression level changes at 5 months, 9 months, 16 months, and 20 months were compared to express increased and decreased molecules. Mass cytometry data were analyzed using Cytobank version 10.6. Cell frequencies were expressed as percentages based on 50,000 CD45<sup>+</sup> cells. Data are shown as mean  $\pm$  S.D. The reported  $p$ -values were obtained from a Student's  $t$ -test. ( $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ , n.s.: not significant).

effector T cells significantly expanded after 16 months. PD-1<sup>+</sup> and FoxP3<sup>+</sup> regulatory T cells (Treg) increased starting at 5 months, though Tregs gradually declined after peaking. CD28<sup>+</sup> cells rose until 16 months and plateaued thereafter. Naïve (CD62L<sup>+</sup>CD44<sup>-</sup>), Effector Memory (EM, CD62L<sup>-</sup>CD44<sup>+</sup>), and Central Memory (CM, CD62L<sup>+</sup>CD44<sup>+</sup>) T cells exhibited considerable inter-individual variability, with no consistent trends (Fig. 2f). When comparing changes in CD8<sup>+</sup> T cells by absolute cell number rather than percentage, the overall pattern was consistent with the relative proportion (Supplementary Fig. 8b). Among the CD8<sup>+</sup> T-cell subsets, Active Effector, CD26<sup>+</sup>, and CD28<sup>+</sup> T cells exhibited distinct age-associated differences, whereas PD-1<sup>+</sup>, Central Memory, Naïve, and Effector Memory T cells showed patterns without significant statistical differences (Supplementary Fig. 8c). CD4<sup>+</sup> T cell dynamics differed in several respects (Figs. 2g and S9a). PD-1<sup>+</sup> CD4<sup>+</sup> T cells notably increased at 16 months, while effector (CD69<sup>+</sup>, CD44<sup>+</sup>), PD-1<sup>+</sup> A.E. (CD69<sup>+</sup>CD44<sup>+</sup>PD-1<sup>+</sup>), and FoxP3<sup>+</sup> Tregs followed similar patterns to their CD8<sup>+</sup> counterparts. CD54<sup>+</sup> CD4<sup>+</sup> T cells initially decreased at 5 months but subsequently increased with age. Among CD4<sup>+</sup> T memory subsets, EM cells showed a significant age-associated increase, unlike naïve or CM cells (Fig. 2h). When comparing CD4<sup>+</sup> T-cell changes by absolute number rather than percentage, the total CD4<sup>+</sup> T-cell counts exhibited patterns consistent with the relative proportion (Supplementary Fig. 9b). Among CD4<sup>+</sup> T-cell subsets, more diverse patterns were observed compared with the total CD4<sup>+</sup> T-cell population (Supplementary Fig. 9c). Notably, while the total CD4<sup>+</sup> T-cell population showed an age-associated decline, several subsets exhibited relative increases, suggesting compositional remodeling within the CD4<sup>+</sup> compartment. The subsets driving the overall pattern of absolute CD4<sup>+</sup> T-cell numbers were identified as Ly6C<sup>+</sup> and CD26<sup>+</sup> CD4<sup>+</sup> T cells, which showed statistically significant decreases with age (Supplementary Fig. 9d, e). Given the limited age-related changes in the B cell opt-SNE distribution, we employed Spanning Tree Progression analysis of Density Normalized Events (SPADE) analysis to resolve fine substructures (Fig. 2i). To clearly visualize the detailed population of B cells, we employed SPADE analysis, which constructs a density-based tree structure of cell populations, making it particularly useful for distinguishing rare populations and visualizing differences in marker expression (Supplementary Fig. 10). Seven subpopulations were defined: four subsets of Age-associated B cells (ABCs: CD11b<sup>+</sup>, CD172 $\alpha$ <sup>+</sup>, CD80<sup>+</sup>, CD44<sup>+</sup>), CD54<sup>-</sup> B cells, Marginal Aone (MZ, CD62L<sup>+</sup>CD44<sup>+</sup>) B cells, and Follicular B cells (CD44<sup>-</sup>FO, CD62L<sup>-</sup>CD44<sup>-</sup>). Quantitative analysis revealed distinct aging trajectories: ABCs 01–03 progressively increased after 16 or 20 months, whereas ABCs 04 declined. CD54<sup>-</sup> B cells expanded significantly by 2.26-fold after 5 months and remained elevated. In contrast, MZ and FO B cells showed no significant temporal change (Figs. 2j and S10). Collectively, these findings demonstrate that aging induces distinct, cell-type-specific remodeling of the splenic immune landscape, involving shifts not only in major lineage abundance but also in phenotypic heterogeneity and functional subset composition.

### Age-related changes in immune cell clustering and composition ratio using PCA and machine learning

To confirm age-related compositional patterns among the six major immune cell clusters, we applied ML algorithms to model these

patterns (Fig. 3a). This approach enabled the development of a model capable of quantitatively analyzing and predicting changes in immune cell composition across age. Prior to model training, 103 molecules extracted from six immune cell populations (CD8<sup>+</sup> T cells: 22 molecules, CD4<sup>+</sup> T cells: 19, B cells: 18, cDC1: 13, cDC2: 14, and macrophages: 17, Supplementary Table 1). These molecules were then grouped by cluster based on expression profiles, and compositional shifts by age were analyzed using principal component analysis (PCA) (Fig. 3b). PCA was conducted in three dimensions to visualize high-dimensional data in a reduced space and to explore variance differences across age groups. The results revealed significant variance separation between age groups along one or more principal components (PCs), indicating that molecular expression patterns in immune cell clusters change in an age-dependent manner and can be effectively captured by specific PCs. Notably, over 50% of the total variance across all age groups was explained by the first three PCs. PCA loading vector analysis identified the major contributing molecules for each cell type (Supplementary Tables 2–7). In CD8<sup>+</sup> T cells, CD54 and CD26 were the strongest contributors to PC2, whereas CD4 was the most influential marker for CD4<sup>+</sup> T cells in PC1. For B cells, CD54 was the key contributor in PC2; for cDC1 cells, both CD54 and CD26 were again prominent in PC2. In cDC2 cells, MHC II contributed most to PC3, while CD172 $\alpha$  was identified as the primary contributor to PC1 in macrophages. The PCA results confirmed that the data structure was suitable for pattern learning and predictive modeling using ML. Support vector regression (SVR), a supervised learning algorithm that maps input data into high-dimensional feature spaces to perform regression through the use of kernel functions<sup>45,46</sup>, was identified as the optimal algorithm. Kernel type and hyperparameters were optimized for each immune subset. In most cases, radial basis function (RBF) kernels provided the best performance by capturing non-linear age-expression relationships, whereas in some subsets sigmoid kernels were selected, likely reflecting saturating or monotonic patterns of molecular expression with age. The model's performance was evaluated using quantitative metrics, including mean squared error (MSE), mean absolute error (MAE), and  $R^2$  score, alongside qualitative assessments via bootstrapping and 95% confidence interval analysis (Supplementary Tables 8 and 9). The model accurately predicted age-associated changes in immune cell composition, both in the test set and when using the full dataset (Fig. 3c, d). Notably, the inset plots in Fig. 3c, d for cDC1, cDC2, and macrophages were included to resolve age-associated changes in cell types with low absolute frequencies. These innate immune subsets displayed broader confidence intervals, reflecting higher variability and biological heterogeneity across samples. The cross-validation MSE scores for each immune subset were  $\leq 2.25$  (Fig. 3e), corresponding to prediction errors of less than  $\sim 2.5$  months relative to the full 18-month biological age span tested. This level of error is consistent with or less than prior murine immune or epigenetic clock models, where mean errors of 2–4 months have been reported<sup>47–49</sup>. Residual analysis demonstrated that predicted values were symmetrically distributed around a mean zero and closely matched actual values. Further diagnostics, including residual normality tests, Q-Q plot, kernel density estimation (KDE), and learning curves, confirmed the model's robustness and predictive reliability. Overall,



our ML framework effectively learns and predicts age-associated compositional changes across six major immune cell clusters, supporting its efficacy for modeling immune aging.

**Machine learning of age-related expression patterns within clustered immune cell populations**

Previously, we successfully applied machine learning to model age-related compositional changes across six major immune cell clusters. In the present analysis, we extended this approach to examine

molecular expression patterns within each immune cell population under the same experimental conditions. The results demonstrated that the developed model could accurately predict age-related expression changes in immune molecules across clusters. This was confirmed through the 95% confidence intervals (Cis) of the predicted values estimated through bootstrapping ( $n = 1000$ ) using the test set (red dots; Fig. 3f, h, j, l, n, p), as well as through trend analysis using the full datasets (blue dots; Fig. 3g, i, k, m, o, q). The model effectively captured the average age-related trends in molecular expression and

**Fig. 2 | Age-related changes in CD8<sup>+</sup> T cell, CD4<sup>+</sup> T cell, and B cell subsets.** **a** FlowSOM clustering identified six metaclusters in the t-SNE map. **b** Each metacluster was named using 12 signature markers for cell identification (CD8<sup>+</sup> T cell: CD3e<sup>+</sup>, CD8a<sup>+</sup>; CD4<sup>+</sup> T cell: CD3e<sup>+</sup>, CD4<sup>+</sup>; B cell: CD19<sup>+</sup>, B220<sup>+</sup>; cDC1: CD11c<sup>+</sup>, XCR1<sup>+</sup>, CD8a<sup>+</sup>; cDC2: CD11c<sup>+</sup>, CD11b<sup>+</sup>, CD172a<sup>+</sup>; Macrophage: Ly-6C<sup>+</sup>, F4/80<sup>+</sup>, CD11b<sup>+</sup>) The expression level of each marker was represented as a heatmap. **c** Age-related percentage change for each identified cell. Data were presented as box plots ( $n = 10$  per month). **d** For CD8<sup>+</sup> T cell, CD4<sup>+</sup> T cell, and B cell subtype analysis, metaclusters of each cell were extracted from FlowSOM to create new opt-SNE maps. The molecular expression levels in each cell population were expressed as heatmaps. **e** CD8<sup>+</sup> T cell opt-SNE is expressed as density dots. The regions of molecules strongly expressed at locations where cell counts change over months are indicated by individual color lines. **f** CD8<sup>+</sup> cell subtypes were analyzed. Data were presented as raincloud plots ( $n = 10$  per month). **g** CD4<sup>+</sup> T cell opt-SNE is

expressed as density dots. The regions of molecules strongly expressed at locations where cell counts change over months are indicated by individual color lines.

**h** CD4<sup>+</sup> cell subtypes were analyzed. Data were presented as raincloud plots ( $n = 10$  per month). **i** Based on the SPADE analysis, we examined the proportion of sub-population bubbles (I–VII) within each B cell population. **j** B cell subtypes were analyzed. Data were presented as box plots ( $n = 10$  per month). Mass cytometry data were analyzed using Cytobank version 10.6 Cell frequencies were expressed as percentages based on 50,000 CD45<sup>+</sup> cells. Data are shown as mean  $\pm$  S.D ( $n = 10$  per month). In box plots, the centre line represents the median, box limits represent the first and third quartiles, and whiskers extend to data points within 1.5 $\times$  the interquartile range. The reported  $p$ -values were obtained from a one-way analysis of variance (ANOVA) and correspond to two-sided statistical tests. ( $p < 0.05$ ,  $^{**}p < 0.01$ ,  $^{***}p < 0.001$ , n.s.: not significant).

explained key features of data distribution across individual clusters. In addition, model performance was supported by residual diagnostics, including symmetry and normality assessments, Q-Q plots, KDE, and learning curves, all of which indicated a robust and well-fitted model. In conclusion, the ML model developed in this study was capable of accurately learning and predicting not only the compositional shifts of immune cell clusters but also the dynamic molecular expression patterns within each cell type. Specifically, the model performed well across CD8<sup>+</sup> T cells (Fig. 3f, g), CD4<sup>+</sup> T cells (Fig. 3h, i), B cells (Fig. 3j, k), cDC1 (Fig. 3l, m), cDC2 (Fig. 3n, o), and macrophages (Fig. 3p, q) throughout the aging process.

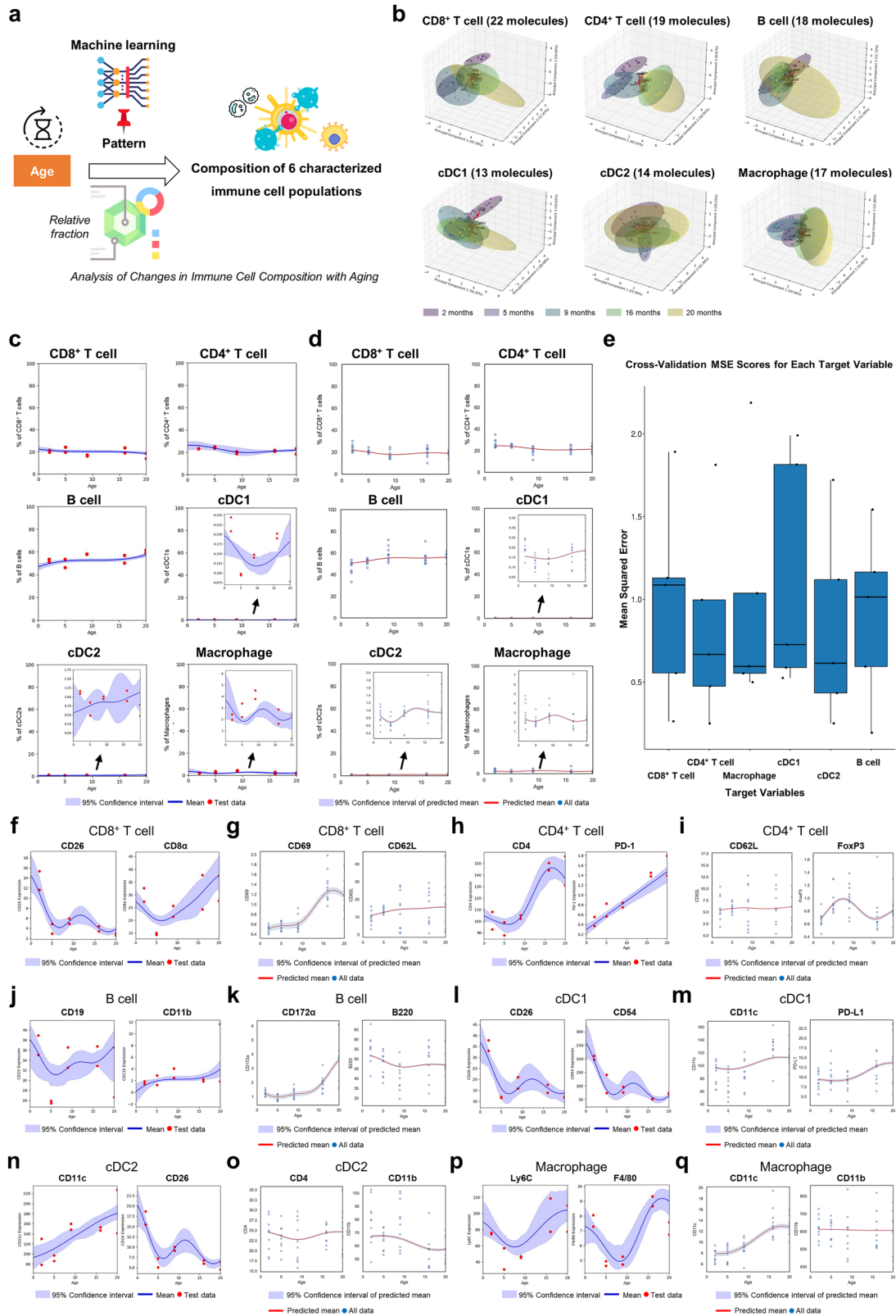
### Development of an application tool for predicting and diagnosing immune aging

Based on the expression data of 103 molecules extracted from six immune cell populations (CD8<sup>+</sup> T cells: 22 molecules, CD4<sup>+</sup> T cells: 19, B cells: 18, cDC1: 13, cDC2: 14, and macrophages: 17; Supplementary Table 1), predicted immune ages were compared with actual age values from an externally generated dataset. A strong correlation was observed between the model-predicted age and actual age ( $R = 0.93$ ,  $R^2 = 0.87$ ,  $p < 0.05$ ) in the test dataset (Fig. 4a).  $K$ -fold cross-validation further confirmed the model's high predictive performance, yielding an average  $R^2$  value greater than 0.9 (Fig. 4b). Comparison of actual and predicted ages showed that the model could estimate immune age within  $\sim 80\%$  accuracy (Fig. 4c). For additional validation, an external dataset ( $n = 100$ ) was randomly generated to simulate molecular expression values corresponding to ages within the 0–20-month range (Fig. 4d). This dataset exhibited a normal distribution centered around an average age of 7–8 months, supporting the reliability of data generation, which reflects the central age distribution of the training dataset (Fig. 4e). The molecular expression values from this external dataset were input into the model to predict immune age. The predicted values closely matched the expected age distribution, demonstrating the model's strong generalization capability. Despite heterogeneity between the simulated dataset and the original training data, the model achieved an explanatory power of 93.43% (Fig. 4f), indicating high adaptability and robustness. These results confirm that the ML-based model developed in this study is not only effective within the training context but also applicable to external datasets. Overall, the model demonstrates strong potential as a quantitative tool for predicting and diagnosing immune aging (Fig. 4g). This finding underscores its practical utility in assessing immune health and age-related immunological decline.

### Evaluation of immune aging acceleration in an obesity model using predictive machine learning

To examine whether aging-associated immune alterations are exacerbated in disease states, we applied our previously trained ML model to a murine model of obesity. For the analysis of the obesity model, we employed the same strategy, utilizing 103 molecules extracted from

six immune cell populations (CD8<sup>+</sup> T cells: 22, CD4<sup>+</sup> T cells: 19, B cells: 18, cDC1: 13, cDC2: 14, and macrophages: 17). Mice at 5 and 17 months of age were fed either a normal-fat diet (NFD) or a high-fat diet (HFD), yielding four experimental groups: 5N, 5H, 17N, and 17H. Body weight analysis confirmed that obesity induction was successful (Supplementary Fig. 11a), and the rate of weight gain was particularly greater in the younger generation (Supplementary Fig. 11b). Splenic CD45<sup>+</sup> immune cells were profiled using CyTOF, and 50,000 events per sample were projected into t-SNE space to visualize obesity-induced compositional shifts (Fig. 5a). Using FlowSOM clustering and 12 lineage-defining markers, six immune subsets were identified: CD8<sup>+</sup> T cells, CD4<sup>+</sup> T cells, B cells, cDC1, cDC2, and macrophages (Fig. S11c, d). Significant changes due to obesity were observed in CD4<sup>+</sup> T cells and macrophages in the 17-month group (Fig. 5b). CD4<sup>+</sup> T cells decreased 2.1% with obesity, while macrophages increased 2.48-fold. Other immune cell types, including CD8<sup>+</sup> T cells, B cells, cDC1, and cDC2, exhibited age-related patterns but showed no significant changes associated with obesity. To quantitatively assess obesity-driven immune aging, we extracted the expression levels of 103 molecules from the six immune populations and applied our established ML model. Immunological age was predicted using a PCA-based framework, adjusted post-hoc at the group-level to account for variance beyond the training set (Fig. 5c). Chronological age predictions for control groups (5N and 17N) aligned closely with actual age (5 and 17 months, respectively), while the HFD groups exhibited signs of accelerated immune aging. In the HFD, predicted immune age showed modest, non-significant elevations in both 5-month-old (+2 months,  $p > 0.05$ ) and 17-month-old (+1.5 months,  $p > 0.05$ ) mice (ANOVA with Tukey's HSD). These results suggest that the model is sensitive to the inflammatory burden associated with obesity. A sample-level heatmap of predicted immune ages showed high consistency within groups, with greater heterogeneity observed in the 17H group (Fig. 5d). PCA mapping of molecular profiles revealed that 5N and 5H clustered with younger reference groups (2M–9M), whereas 17N and 17H aligned with older groups (16M–20M). Notably, the 17-month HFD group had individuals distributed widely, even reaching the 5-month group area, visually demonstrating the high heterogeneity within the group as previously predicted and suggesting that obesity in old age increases the heterogeneity of the immune system (Fig. 5e). By contrast, variation along PC1 primarily reflected batch effects rather than age, consistent with the observation that reference groups spanning 2–20 months did not distribute progressively along PC1 but instead separated more clearly along PC2. Finally, to explore the underlying molecular drivers, we analyzed the top contributors to PC2 across all groups (2M–20M and 5N/5H/17N/17H) and visualized pairwise correlations among the top 10 molecules (Fig. 5f, and Supplementary Fig. 12). These drivers therefore represent features jointly influenced by chronological aging and diet-induced acceleration, highlighting molecular pathways commonly contributing to immune age shifts under both conditions. Strong positive correlations ( $r > 0.8$ ) were



observed among immune markers functionally related to T cell exhaustion (e.g., PD-L1, CD44), T cell activation (CD4, CD3ε), and B cell maturation (CD172α, CD11c, CD44), suggesting the presence of coordinated immune regulatory modules driving the observed age-associated changes. We confirmed the high correlation and significance of 10 molecules and further investigated how their expression patterns change with aging. Both PD-L1 and CD44 on CD8<sup>+</sup> T cells

increased consistently and significantly with age (Supplementary Fig. 13a). PD-L1 and CD44 on CD4<sup>+</sup> T cells also increased consistently and significantly, like CD8<sup>+</sup> T cells, while CD4 and CD3ε increased dramatically at 16 months (Supplementary Fig. 13b). Finally, CD40 and CD11c on B cells also increased sharply at 16 months, and CD44 and CD172α significantly increased at 20 months (Supplementary Fig. 13c). All 10 molecules showed significant increases with aging, each with a

**Fig. 3 | Post-machine learning validation using the proportions and expressed molecules of 6 cell populations.** **a** Summary of the learning process using ML algorithms. Created using icons sourced from Flaticon ([www.flaticon.com](http://www.flaticon.com)), distributed under Flaticon's Free License (<https://www.flaticon.com/legal>). **b** Using 3D PCA, we evaluated whether molecular expression levels in each cell showed age-related heterogeneity. Each "X" represents one mouse, and the colored surfaces indicate clustering by age group. Red arrows represent the molecular loadings driving PC separation. **c, d** Validation of ML results using cell pattern changes. Red dots represent test samples not used during training. Inset plots for cDC1, cDC2, and macrophages display zoomed-in views due to their lower abundance and narrower expression ranges, allowing clearer visualization of predicted trends and confidence intervals. **e** Cross-validation results using 5-fold cross-validation for performance evaluation of the multivariate support vector regression ML model. The *x*- and *y*-axes of the graph represent mean squared error and the types of immune-related target molecules, respectively ( $n = 5$  per cell type).

**c, f, h, j, l, n, p** Estimation of confidence intervals of predicted values through bootstrapping of an established multivariate SVR ML model and overlapping results of observations in the test set. The *x*- and *y*-axes of the graph represent age and the composition ratio of identifiable immune cell groups within the population, respectively. To simultaneously compare the overall compositional proportions within a population, the *y*-axes of all graphs of cell composition proportions were fixed to the same scale. **d, g, i, k, m, o, q** Analysis of the change pattern of predicted values using the established multivariate support vector regression ML model and overlapping results of observations in all data sets. The *x*- and *y*-axes of the graph represent age and the composition ratio of identifiable immune cell groups within the population, respectively. The red solid lines in the graph show the mean predicted patterns of the trained model. Cell frequencies were expressed as percentages based on 50,000 CD45<sup>+</sup> cells. In box plots, the centre line represents the median, box limits represent the first and third quartiles, and whiskers extend to data points within 1.5× the interquartile range.

different pattern. Next, we evaluated the potential for model simplification using 10 molecules previously identified as correlated with aging patterns. The compact 10-marker SVR model achieved comparable predictive performance to the full 103-marker model ( $R^2 = 0.89 \pm 0.06$ ; RMSE = 2.25), maintaining over 95% concordance within the  $\pm 1.96$  SD limits in Bland–Altman analysis (Supplementary Fig. 13d, e). Group-wise predictions showed consistent immune-age elevation in HFD cohorts compared to normal diet controls (Supplementary Fig. 13f). Together, these results validate our ML-based model as a robust platform for immune age estimation and highlight its translational potential for identifying immune aging acceleration in pathological conditions such as obesity.

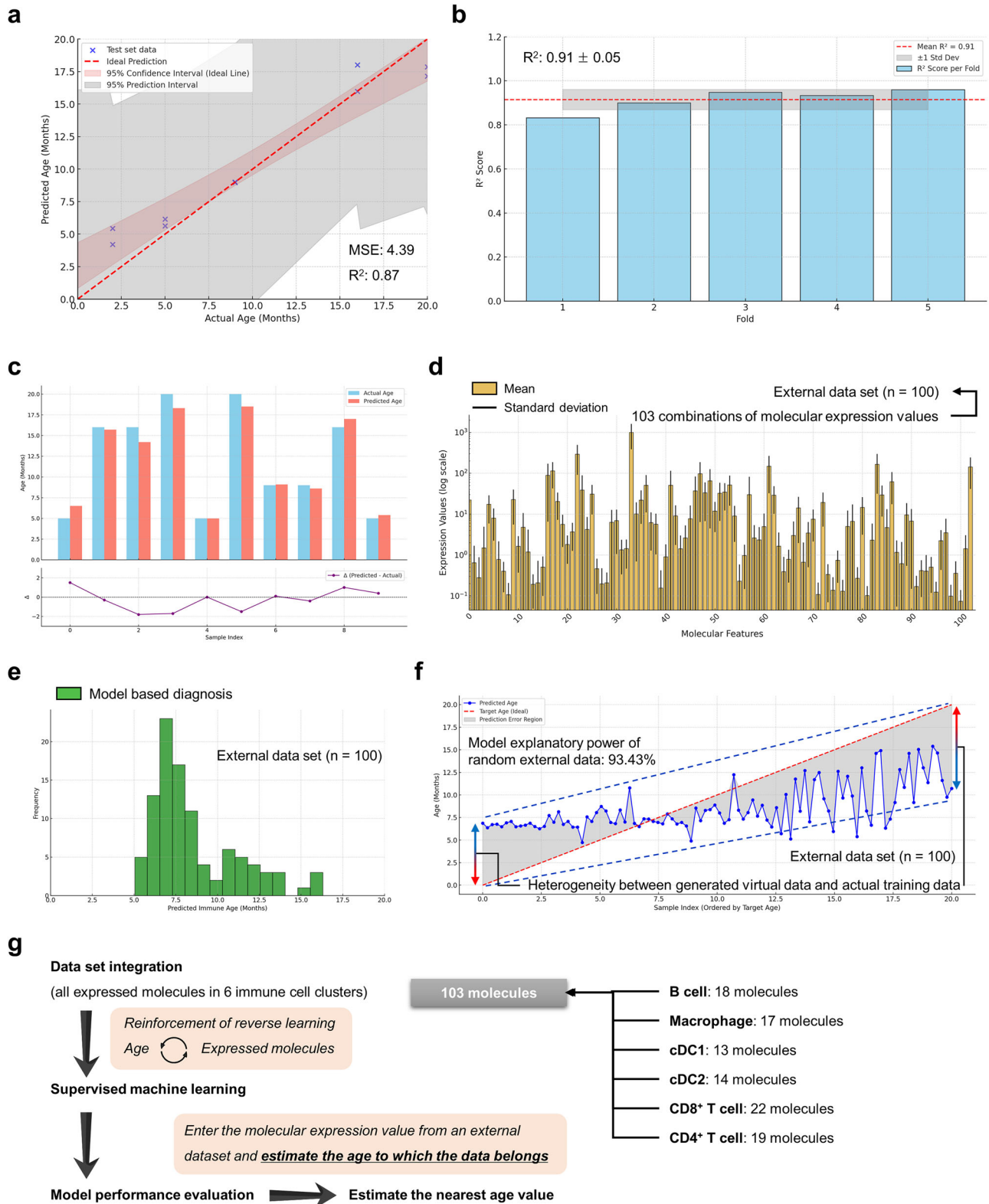
## Discussion

Although it is well established that the immune system undergoes significant changes with age, protein-level tracking of these changes, especially across immune subsets, has remained limited. Most previous studies have relied on transcriptomic data (e.g., scRNA-seq), which, while informative, do not fully capture the proteomic heterogeneity that defines immune cell function. In this study, we used mass cytometry (CyTOF) to resolve protein-level shifts in immune markers and cell populations during aging. Our findings revealed characteristic expansions of immunosuppressive or exhaustion-associated populations (CD28<sup>+</sup>, PD-1<sup>+</sup>, FoxP3<sup>+</sup>), alongside reductions in populations involved in activation and trafficking (CD54<sup>+</sup>, CD26<sup>+</sup>)<sup>50–54</sup>. These age-related shifts suggest that the immune system undergoes both functional exhaustion and migratory impairment with age. Recognizing that immune cell functionality is governed by intercellular communication rather than isolated marker expression, we further analyzed dominant splenic immune subsets, CD8<sup>+</sup> T cells, CD4<sup>+</sup> T cells, and B cells. In CD8<sup>+</sup> T cells, the early and persistent decline in CD26<sup>+</sup> cells, known to counteract T cell exhaustion<sup>52,54</sup>, was accompanied by an expansion of PD-1<sup>+</sup> and FoxP3<sup>+</sup> Tregs, consistent with prior reports linking PD-1<sup>+</sup>CD8<sup>+</sup> T cells to immunosenescence<sup>55</sup>. The increase in CD69<sup>+</sup> cells can be seen as a phenomenon that occurs because cells with CD69, such as Treg and PD-1<sup>+</sup> T cells, increase during aging<sup>56,57</sup>.

In CD4<sup>+</sup> T cells, CD54 expression initially decreased but rebounded in older mice. In CD4<sup>+</sup> T cells, CD54 was known to function as a costimulatory molecule<sup>58,59</sup>. CD4<sup>+</sup> T cells harboring CD54, in addition to TCR stimulation, experience co-stimulation through CD54, which activates PI3K, increases IL-2 promoter transcription, and increases Th1 (T helper type 1) cell proliferation and cytokine secretion<sup>58,59</sup>. Age-related increases in PD-1<sup>+</sup>CD4<sup>+</sup> and FoxP3<sup>+</sup>CD4<sup>+</sup> T cells further support the notion of an immunosuppressive shift within the helper T cell compartment<sup>56,57,60,61</sup>. While T cell subsets (naïve, EM, CM) exhibited high interindividual variability, it is worth noting that the classification markers CD62L and CD44 alone may be insufficient to resolve these phenotypes with high fidelity<sup>62–64</sup>. Analysis of the B cell compartment revealed substantial age-related heterogeneity. SPADE analysis

identified seven distinct subtypes, including four age-associated B cell (ABC) populations. Interestingly, ABC04 declined with age, while ABC01–03 followed distinct trajectories, suggesting functional heterogeneity even within the ABC group. ABC01, which expanded markedly only at 20 months, may represent a late-aging-specific subset. Consistent with prior studies, ABCs (particularly CD11b<sup>+</sup>CD172a<sup>+</sup>CD80<sup>+</sup>CD44<sup>+</sup> cells) are known to drive autoantibody production and Th17 polarization, potentially contributing to auto-immune phenotypes associated with aging<sup>29,65,66</sup>. Meanwhile, CD54<sup>+</sup> B cells, thought to represent functionally inert or immature B cells<sup>67</sup>, increased sharply at 5 months, suggesting early-stage dysregulation. Taken together, both T and B cell populations demonstrated features of functional exhaustion and inflammatory feedback, aligning with the increased susceptibility to infection, cancer, and autoimmunity observed in the elderly. Increases in PD-1, CD44, and CD69 expression on T cells aligned with reported aging-associated patterns<sup>28,57</sup>. Similarly, expansion of PD-L1<sup>+</sup> CD4/CD8<sup>+</sup> T cells<sup>57</sup> and elevations in regulatory T cells and ABCs<sup>29,56,57</sup> were consistent with prior findings. These concordant results reinforce the validity of our observations and the robustness of subsequent machine learning analyses.

Despite these insights, defining a precise “immune aging stage” based on chronological time remains challenging due to the non-linear and multifactorial nature of aging. To overcome this limitation, we developed a predictive ML model based on 103 molecules extracted from six immune cell populations (CD8<sup>+</sup> T cells: 22 molecules, CD4<sup>+</sup> T cells: 19, B cells: 18, cDC1: 13, cDC2: 14, and macrophages: 17). Initial PCA analyses confirmed that molecular expression patterns stratify robustly by age. Among various algorithms, SVR was selected for its capacity to model high-dimensional, nonlinear relationships with minimal risk of overfitting<sup>68,69</sup>. SVR's  $\epsilon$ -insensitive loss function further enabled robust predictions despite interindividual variability<sup>70</sup>. The trained model accurately predicted immunological age from expression data in both test and full datasets. When validated using synthetic datasets ( $n = 100$ ), the model exhibited a strong age distribution centered around 7–8 months and maintained high explanatory power ( $R^2 > 0.93$ ). The external dataset mirrored the statistical properties of the training dataset, resulting in a distribution centered at 7–8 months. This reflects the fact that the training cohort itself (at molecular pattern distribution) was densest in this age range, while very young or old ages were underrepresented. As such, while the generated dataset reliably reproduced the central age distribution, future expansions will aim to incorporate broader age ranges for improved generalizability. Importantly, the model generalizes well to data points beyond the training range, underscoring its robustness and potential applicability in real-world settings. Although direct integration with external datasets was limited due to the lack of publicly available mouse datasets containing simultaneous multi-immune lineage expression across developmental ages, we compared our model's predicted age-dependent expression trajectories with



literature-based trends. Notably, CD8<sup>+</sup> and CD4<sup>+</sup> T cell marker expression trends showed age-associated decline, consistent with reported thymic involution and T cell maturation patterns in mice<sup>71,72</sup>. Similarly, cDC1/cDC2 marker dynamics aligned with reports on dendritic cell subset reprogramming with aging<sup>73,74</sup>. These similarities support the biological validity of our model.

To evaluate practical utility, we applied the model to biological samples from normal 5- and 17-month-old mice. Remarkably,

predicted immune ages closely matched actual chronological ages, validating the model's accuracy in capturing immune aging in heterogeneous datasets. To test its clinical relevance, we next applied the model to a murine obesity model, as obesity is a well-known driver of immunosenescence and inflammation<sup>75-79</sup>. t-SNE and FlowSOM analyses revealed that obesity induced reductions in CD8<sup>+</sup> and CD4<sup>+</sup> T cells and expansions in macrophages and B cells, mirroring patterns observed in aged animals. Notably, these shifts were more pronounced

**Fig. 4 | Immune age estimation using external molecular expression data and reinforcement learning.** **a** Evaluate the consistency between the model-predicted age and actual age values through  $R^2$  in the test dataset. **b** Stratified  $K$ -fold cross-validation  $R^2$  score for evaluating the model's predictive performance. The five bars represent each fold in 5-fold cross-validation. The dataset was randomly divided into five subsets, and the model was iteratively trained on four and validated on one.  $R^2$  scores represent performance on the validation fold. The red dashed line shows the mean  $R^2$  score. **c** Comparison of actual and predicted immune ages for each validation sample. These test samples were not included in model training or validation. **d–f** The external data set refers to a synthetic dataset composed of randomly generated expression profiles for 103 molecules extracted from six immune cell populations. This dataset was used to evaluate model generalizability under unseen conditions. **d** Balanced synthetic data: Mean and variability of molecular expression. **e** Balanced predicted age distribution for the generated

external dataset. **f** Target Age vs. Predicted Age: Visualization matching. Evaluating the explanatory power of the built model on external data sets. The explanatory power (93.43%) corresponds to the coefficient of determination ( $R^2$ ), indicating that ~93% of the variance in target age is explained by the model predictions.  $R^2 = 1 - \frac{\sum(y_{\text{true}} - y_{\text{pred}})^2}{\sum(y_{\text{true}} - \bar{y}_{\text{true}})^2}$ , where  $y_{\text{true}}$  is the target age,  $\bar{y}_{\text{true}}$  is the mean target age, and  $y_{\text{pred}}$  is the predicted age. **g** ML algorithm training and tool development scheme based on molecular expression levels. A total of 103 immune molecular features derived from six immune cell clusters were integrated. These features were used in a SVR to model the relationship between molecular expression and chronological age. For external validation, synthetic datasets were input into the trained model to predict the nearest immune age values. This diagram summarizes the overall process from data integration to model evaluation. Data are shown as mean  $\pm$  S.D.

in the aged (17H) than in the young (5H), suggesting that obesity imposes a greater immunological burden in older individuals. Because HFD-related increases did not reach statistical significance, we do not interpret them as immune aging acceleration. Instead, we consider obesity-induced chronic inflammation as a potential driver of the modest elevations observed in both age groups, reflecting the biological coupling between inflammation and immunosenescence. Future studies with larger sample sizes will be required to determine whether such changes represent true age-acceleration or primarily inflammatory load.

In developing our model, we intentionally avoided applying feature-specific weights, a practice common in previous ML-based biological age models. We reasoned that such weighting could introduce bias, particularly in small or heterogeneous datasets. Instead, we employed uniform variable treatment alongside post hoc correction using PCA and bootstrapped confidence intervals. Notably, PCA of obesity samples revealed a distribution distinct from that of aging controls, yet the model maintained predictive fidelity, indicating successful generalization. Furthermore, correlation analysis of PC contributors revealed key co-regulatory networks: PD-L1<sup>+</sup>CD4<sup>+</sup> and PD-L1<sup>+</sup>CD8<sup>+</sup> T cells are unique populations previously known to increase with aging<sup>57</sup> and are known to mediate inhibitory counter-signaling in the tumor environment<sup>80</sup>. Therefore, they may be one of the reasons for the increased cancer incidence with aging. Correlations between CD4<sup>+</sup> T cell and CD44 also suggest memory phenotype shifts, while CD172 $\alpha$ <sup>+</sup> and CD11c<sup>+</sup> B cells likely represent functionally activated ABCs<sup>30,81</sup>. Collectively, these patterns underscore a coordinated decline in adaptive immune function associated with age and metabolic stress.

Finally, this study demonstrates that ML, when rigorously applied to high-dimensional immune data, can generate reliable and interpretable models of immune aging. Previously, numerous immune clocks have been proposed to quantify immunosenescence using cytokine panels, transcriptomic signatures, or DNA methylation patterns<sup>13,28,37,40,43</sup>. However, our model extends current frameworks by (i) leveraging high-dimensional protein-level data across lineage-defined immune subsets using CyTOF, (ii) enabling bidirectional prediction of age and expression profiles within each immune population, and (iii) demonstrating immune age acceleration in a metabolic disease context (obesity), which most existing clocks have not addressed. Our approach advances beyond prior models based solely on cytokine profiles or DNA methylation. Importantly, the model performed well even with moderate sample sizes and demonstrated translational potential in disease contexts. Previous landmark work by Alpert et al.<sup>13</sup> applied machine learning to multi-dimensional immune profiling in large human cohorts, demonstrating predictive power for mortality and disease outcomes. Our study complements this by establishing a murine immune aging model with defined perturbations such as HFD, thereby enabling mechanistic dissection of age-accelerating conditions. In contrast to

the large-scale marker panels used in human studies, we demonstrate that accurate prediction can also be achieved using a restricted panel of 103 immune markers across defined cell subsets. Thus, our work provides a flexible experimental framework that bridges controlled mechanistic studies with translational potential. While this model demonstrates strong predictive performance, its reliance on over 30 immune markers measured by multiparameter mass cytometry may limit its immediate clinical application. To further enhance the model's translational feasibility, future efforts emphasized marker optimization and simplification. The feature-selection analysis performed here indicates that most of the predictive information can be captured by a focused subset of T-cell activation/exhaustion markers (e.g., CD44, PD-L1) and B-cell markers (e.g., CD44, CD172 $\alpha$ ). A 10-marker configuration retained robust group discrimination and reproduced the overall age-related trajectories observed with the full 103-marker model. This demonstrates that even a substantially reduced panel can preserve biological interpretability and practical utility. Accordingly, the compact framework offers a promising foundation for clinically adaptable immune-age assessment tools, and validation in larger, diverse cohorts will further substantiate its generalizability. Integrating these approaches will be crucial for translating the immune aging clock from research to clinical practice. Future applications may extend this framework to human cohorts, enabling population-level immune age mapping and personalized immunological health assessments.

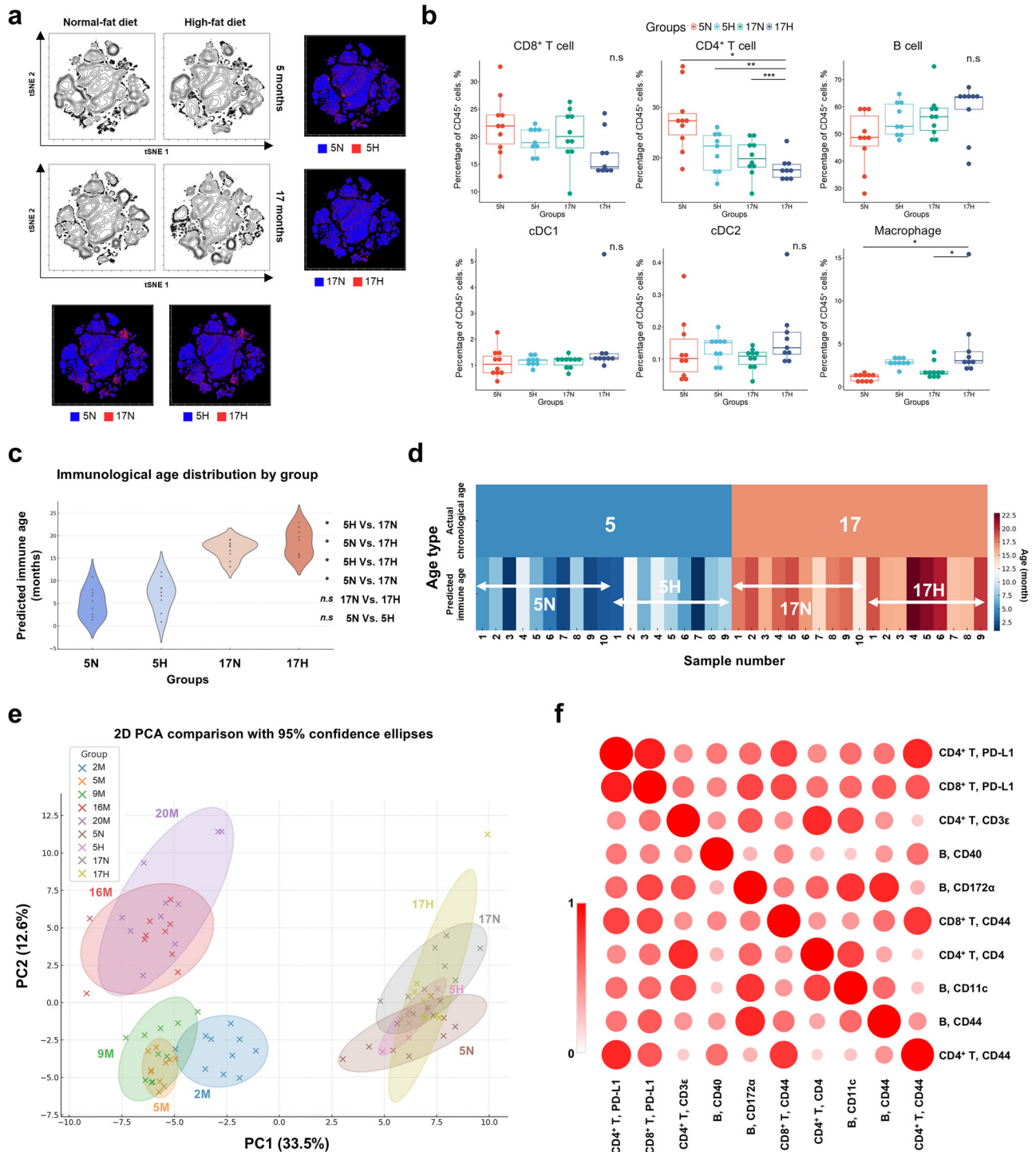
## Methods

### animal

The animal study was performed under the approval of the Suncheon National University Institutional Animal Care and Use Committee (SCNU IACUC-2025-07). C57BL/6J Mice (Male, each age group  $n = 10$ , 2–20 months, NCrI, #00064) were purchased from the Korean Basic Science Institute Animal Facility of Aging Science (KBSI AFAS, Gwangju, Korea). Animals were housed in polycarbonate cages within a controlled environment ( $22 \pm 2$  °C and  $50 \pm 5\%$  humidity, specific pathogen-free/SPF), fed diets with varying compositions for each group, and provided with water. The group for the aging prediction model was fed a standard diet. Experimental/control animals were bred separately. For the obesity model, they were given a normal-fat diet (NFD,  $n = 10$ ) and a high-fat diet (HFD,  $n = 9$ ) for 3 months (Supplementary Table 10)<sup>82</sup>.

### Splenic immune cell isolation

Mice were euthanized using CO<sub>2</sub> gas. The spleen was collected and made into single cells using a cell strainer (100  $\mu$ m; Falcon, NY, USA, #352360). Red blood cells were removed by treating with RBC Lysis Buffer (Invitrogen, MA, USA, #00-4333-57). After that, lipids were removed using a cell strainer (40  $\mu$ m; Falcon, NY, USA, #352340), and cells were counted.



**Fig. 5 | Prediction of accelerated immune-aging due to obesity.** **a** The t-SNE map expresses the number of cells as contour lines. Depending on the aging and obesity conditions, various cell changes appear at the same location on the map (5N, 17N;  $n = 10$ , 5H, 17H;  $n = 9$ ). **b** Population changes of six cell populations (CD8<sup>+</sup> T cell, CD4<sup>+</sup> T cell, B cell, cDC1, cDC2, macrophage) separated using FlowSOM ( $n = 10$  per month). Data were shown as box plots. Although both **c**, **d** represent predicted immune ages, **c** provides a statistical summary of group distributions, whereas **d** highlights heterogeneity at the level of individual samples (that may not be as easily appreciated in **c**). **c** Immunological age distribution by group. **d** Heatmap of actual vs. predicted immune age (per sample). **e** 2D PCA comparison with 95% confidence ellipses. **f** Correlation

matrix of the top 10 PC2-contributing immune markers derived from PCA including all groups (2M–20M and 5N/5H/17N/17H). These markers represent shared molecular drivers of PC2 variation, reflecting combined contributions of aging and obesity, and highlight common pathways underlying immune age shifts. Cell percentages were calculated relative to CD45<sup>+</sup> cells 50,000. Mass cytometry data were analyzed using Cytobank version 10.6. Data are shown as mean  $\pm$  S.D. In box plots, the centre line represents the median, box limits represent the first and third quartiles, and whiskers extend to data points within 1.5 $\times$  the interquartile range. The reported  $p$ -values were obtained from a one-way analysis of variance (ANOVA) and correspond to two-sided statistical tests. ( $p < 0.05$ ,  $^{**}p < 0.01$ ,  $^{***}p < 0.001$ , n.s.: not significant).

### Antibody metal labeling

Metals were conjugated to purified antibodies using Maxpar X8 Antibody Labeling Kits (Standard Biotoools, CA, USA, #201155A, #201156A, #201173A, #201174A, #201176A). Experiments were performed according to the manufacturer's instructions<sup>83</sup>. Of the 30 antibodies used in the experiment, 5 (CD172 $\alpha$ /SIRP $\alpha$ , Galectin-9, CD26, CD44, and XCRI) were conjugated to metals.

### Mass cytometry analysis

Harvested mouse spleens were prepared into single-cell suspensions, with  $1 \times 10^7$  cells for each sample. Cells were stained using a panel of metal-conjugated antibodies (Supplementary Table 11) after cisplatin (Cell-ID Cisplatin, Standard Biotoools, CA, USA, #201064) live/dead staining. DNA was labeled using an Ir-intercalator (Cell-ID Intercalator-Rh, Standard Biotoools, CA, USA, #201103A) to determine whether cells were present. EQ Beads (EQ Four Element Calibration Beads; Standard Biotoools, CA, USA, #201078, EQ Six Element Calibration Beads; Standard Biotoools, CA, USA, #201245) were added to each sample immediately before analysis. Normalization was performed using the CyTOF machine instrument software (Data obtained from Helios was normalized to EQ4 beads, CyTOF XT was normalized to EQ6 beads), and no cell stimulation or barcoding was applied. The study was performed using the Helios (CyTOF2, Standard Biotoools, CA, USA) and CyTOF XT (Standard Biotoools, CA, USA) instruments, and 300,000 events were collected for each sample. All fcs files obtained from analysis on the Helios and CyTOF XT were analyzed using Cytobank (v10.6, Beckman Coulter, <https://premium.cytobank.org/cytobank/experiments>).

### CyTOF data clean-up flow

Data cleanup, as applied to mass cytometry, typically focuses on eliminating dead cells, debris, normalization beads, true aggregates, and coincident ion clouds from raw data. First, the 140Ce signal is removed from the obtained FCS file to remove beads. Second, a process (Residual, Center, Offset, Width, and event length) was performed to secure only pure cell data. Third, only live cells that are 195Pt negative are selected. Finally, single cells (191Ir<sup>+</sup>, 193Ir<sup>+</sup>) were selected, and only immune cells that are 89Y CD45<sup>+</sup> were chosen (Supplementary Fig. 14). This clean-up flow automated nature avoids the subjectivity of other gating methods and results in the unbiased elimination of unwanted events<sup>84</sup>.

### Dimensionality reduction analysis

t-SNE (t-distributed stochastic neighbor embedding) analysis was used to express the data after the clean-up process by dimension reduction as a 2D map. When analyzing entire spleen cells, the t-SNE-CUDA (GPU-accelerated implementation of t-SNE) algorithm was used, 50,000 events were randomly extracted from each connected fcs file, and clustering was performed through automatic settings. When analyzing only individual cells (e.g., B cells, T cells), FlowSOM was run based on the map generated by the previously analyzed t-SNE-CUDA to select each cell population (CD8<sup>+</sup> T cell, CD4<sup>+</sup> T cell, and B cell) and draw a t-SNE map using the opt-SNE (optimized t-SNE) algorithm.

### FlowSOM and SPADE clustering analysis

All clustering analyzes were performed based on the previously obtained t-SNE map. FlowSOM (self-organizing map) analysis was performed for all spleen cells, and 100 clusters were formed using the Hierarchical Consensus clustering method to get six metaclusters. At this time, the seed value was randomly optimized. The six metaclusters were named into six types of cells (CD4<sup>+</sup> T cell, CD8<sup>+</sup> T cell, B cell, cDC1, cDC2, macrophage) through the difference in the expression rate of 12 signature markers (CD4, CD8 $\alpha$ , CD3 $\epsilon$ , NK1.1, CD19, B220, CD11c, CD11b, F4/80, Ly-6C, CD172 $\alpha$ , XCRI). For B cell subtype analysis, a 200-node tree was constructed using SPADE analysis to distinguish changes in cell number. Subtypes were named based on the median intensity of

each molecule. The data obtained through analysis were visualized and graphed using SRplot (A free online platform for data visualization and graphing)<sup>85</sup>.

### Exploring ML algorithms

An optimal ML algorithm search was performed to learn and predict age-related changes in the expression patterns of molecules within each cell cluster. Structurally, it has one independent variable (age) and multiple dependent variables (expressed molecules and composition of immune cells). Since the dependent variables are multi-dimensional and highly correlated, a multivariate regression model analysis approach was attempted. The primary purpose of this study was to learn and predict the simultaneous changes in multiple dependent changes according to the change in independent variables. The longitudinal multiple univariate analysis approach was excluded, and the transverse multivariate analysis approach was selected. The changes in the expression of molecules and the cell composition ratio according to age were mostly not linear, and nonlinear regression algorithms were mainly attempted. After stepwise trials of various nonlinear regression candidate algorithms (neural network, K-nearest neighbors, decision tree, random forest, gradient boosting, support vector, etc.), qualitative and quantitative evaluation finally selected the optimal algorithm that explained all data sets well. The algorithm evaluation mainly applied qualitative and quantitative factors such as model fitting and performance, *R*-squared (*R*<sup>2</sup>) score, MAE, MSE, and cross-validation score.

### Principal component analysis

Each of the six immune cell populations was clustered by age. Each cluster reflected the similarity between cell populations based on the expression pattern of immune molecules. The changes in the composition ratio of the six significant immune cell populations by age were analyzed in two- and three-dimensional spaces using PCA, and the contribution rate of each PC was defined as the considerable dimension values contributing to the changes in the composition ratio of immune cells, and the interactions between the final significant factors and the data distribution by age group were visualized. In addition, the loading vector values were calculated for each molecule expressed in each immune cell population, and additional identification of variables closely related to the changes in the composition ratio of specific immune cells by age was performed. The F-statistics were calculated to explain the variance difference between age groups of the PCs derived from the PCA results, and a *p*-value of 0.05 was set as the standard to determine whether the variance difference by PC and age was statistically significant.

### Model training and validation

The model was developed and verified through data input, model development, learning and testing, and evaluation. First, independent and dependent variables were set through data input and dataset feature recognition, and all data were processed as continuous variables. Data preprocessing, missing value review, and data normalization to the same scale of variables were performed to secure the stability of model learning. Second, a model was developed by applying an algorithm that reflects the nonlinear characteristics of immune data and appropriately explains changes in immune molecule expression and cell composition ratio according to age. The model was optimized through Kernel functions and hyperparameter settings. Third, model learning and testing were performed through dataset division. The complete dataset (*n* = 50 mice) was randomly divided into a training set (80%, *n* = 40) and a test set (20%, *n* = 10) using stratified random sampling across age groups to preserve class distribution. The training set was used to perform fivefold cross-validation and model optimization, while the test set was exclusively reserved for evaluating final model performance. All random partitions

were conducted with fixed random seeds for reproducibility. The model was optimized to effectively explain data points around the hyperplane while minimizing  $\epsilon$ -insensitive loss, and  $k$ -fold cross-validation ( $k=5$ ) was performed to prevent overfitting during the learning process. Otherwise, to obtain a robust estimate of predictive accuracy, we applied fivefold cross-validation, in which the dataset was randomly divided into five equal subsets. In each iteration, fourfolds were used for training, and onefold was used for validation, with the process repeated until all folds had been used once as the validation set<sup>86,87</sup>. The learned model was evaluated for prediction performance on the test set, and the model's generalization performance and prediction accuracy were confirmed through comparisons between the prediction results and the actual values. If the evaluation results were not satisfactory, the data preprocessing and variable selection steps were reexamined, and the model was improved through model retraining and optimization repetitions. Fourth, various verification tools were applied to evaluate the performance and reliability of the final developed model from multiple angles. Visual fitting analyzes were performed to confirm the model's suitability to compare the predicted values with the actual values, and normality was evaluated through histograms with the residual distributions to check if the residuals were randomly distributed around the mean of 0. In addition, quantile-quantile (Q-Q) plots were generated to visually verify the normality of the residuals, and KDE was used to analyze the continuous distribution of the residuals precisely. To evaluate the change in model performance according to the size of the learning data, a learning curve was created to diagnose overfitting or underfitting, and a five-fold cross-validation was performed for each dependent variable to calculate the MSE score.

### Applying the model

The model was further trained to accurately predict the relationship between immune molecule expression data and age in both directions. After integrating 103 molecules extracted from six immune cell populations (CD8<sup>+</sup> T cells: 22 molecules, CD4<sup>+</sup> T cells: 19, B cells: 18, cDC1: 13, cDC2: 14, and macrophages: 17), supervised machine learning (as multivariate output modeling) was performed to predict immune molecule expression patterns with age as an independent variable or, conversely, to estimate age by inputting immune molecule expression data. Both model directions (age  $\rightarrow$  expression pattern, expression pattern  $\rightarrow$  age) were trained using nonlinear regression algorithms explored during the development process, and generalization performances were confirmed through  $k$ -fold cross-validation ( $k=5$ ). Model training in the "molecular expression pattern  $\rightarrow$  age" direction was performed through reverse mapping reinforcement, which reversely predicts age values from the complex expression patterns of immune molecules. The model's performance was quantitatively analyzed through MSE and  $R^2$  scores, and age estimation was performed based on immune molecule expression patterns in external datasets to confirm the model's reliability and applicability. By applying molecular expression data from external datasets to the developed model, the most similar age values to the overall molecular expression patterns were predicted using the proximity age estimation method. The model's reliability was confirmed by comparing the accuracy of age estimation based on molecular expression with external data. For external validation, we generated a synthetic dataset ( $n=100$ ) by sampling 103 immune marker values from Gaussian distributions parameterized by the mean and standard deviation observed in the training data for each molecule. This procedure was used to evaluate the model's robustness against hypothetical but biologically plausible immune profiles not included in the original dataset.

### Applying aging and obesity data to the model

Finally, to evaluate the independent applicability of the established ML model, spleen molecular expression data from the NFD and HFD

groups of mice at 5 and 17 months were used to train the model. A total of 103 molecules extracted from six immune cell populations (CD8<sup>+</sup> T cells: 22 molecules, CD4<sup>+</sup> T cells: 19, B cells: 18, cDC1: 13, cDC2: 14, and macrophages: 17) were quantitatively measured using the CyTOF, and these were simultaneously input into the previously trained regression model to estimate the "immunological age" of the individuals. In the immunological age estimation, a decision-making method was applied based on the overall expression pattern similarity without specific weighting of each immune-related expression molecule.

To assess the feasibility of a reduced immune-marker panel, we constructed a compact model using the top 10 molecules with the highest loading values in PC2 derived from the 103-marker PCA analysis. These markers primarily represented T- and B-cell activation and differentiation pathways. The reduced dataset was used to retrain a support vector regression (SVR, RBF kernel) model under the same cross-validation conditions applied to the full model. Model performance was evaluated using regression diagnostics ( $R^2$ , RMSE, Bland-Altman analysis). Group-wise predictions were further compared using external diet- and age-based validation cohorts.

### Statistical analysis

Statistical analysis was performed using Student's  $t$ -test for comparisons between two groups or one-way analysis of variance (ANOVA) with Tukey's HSD comparison between more than two groups. Statistical calculations were done using SPSS ver. 27 (SPSS, IL, USA). In Student's  $t$ -test,  $p < 0.05$  was considered statistically significant ( $p < 0.05$ ,  $^{**}p < 0.01$ ,  $^{***}p < 0.001$ ). And after ANOVA, a post-hoc test was conducted only when the significance between groups was  $p < 0.05$ . Symbols ( $p < 0.05$ ,  $^{**}p < 0.01$ ,  $^{***}p < 0.001$ ) indicated subsequent comparisons of significance between groups.

### Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

### Data availability

All data are included in the Supplementary Information or available from the authors, as are unique reagents used in this Article. The raw numbers for charts and graphs are available in the Source Data file whenever possible. Source data are provided with this paper.

### Code availability

We used publicly available software for the analyses. All source code to reproduce the analyses is available on the Python libraries [pandas (<https://github.com/pandas-dev/pandas>), numpy (<https://github.com/numpy/numpy>), matplotlib (<https://github.com/matplotlib/matplotlib>), seaborn (<https://github.com/mwaskom/seaborn>), plotly (<https://github.com/plotly/plotly.py>), sklearn (<https://github.com/scikit-learn/scikit-learn>), PCA (sklearn.decomposition, <https://github.com/scikit-learn/scikit-learn>), scipy.stats (<https://github.com/scipy/scipy>), matplotlib.image (<https://github.com/matplotlib/matplotlib>), OpenCV (<https://github.com/opencv/opencv>)].

### References

1. Niccoli, T. & Partridge, L. Ageing as a risk factor for disease. *Curr. Biol.* **22**, R741-R752 (2012).
2. Fuster, V. Changing demographics: a new approach to global health care due to the aging population. *J. Am. Coll. Cardiol.* **69**, 3002-3005 (2017).
3. Baltes, P. B., Lindenberger, U. & Staudinger, U. M. *Life Span Theory in Developmental Psychology. Handbook of Child Psychology* <https://doi.org/10.1002/9780470147658.chpsy0111>. (2007).

4. Drewelies, J. et al. Using blood test parameters to define biological age among older adults: association with morbidity and mortality independent of chronological age validated in two separate birth cohorts. *GeroScience* **44**, 2685–2699 (2022).
5. Bortz, J. et al. Biological age estimation using circulating blood biomarkers. *Commun. Biol.* **6**, 1089 (2023).
6. Chen, Y. et al. Advances in biomarkers and diagnostic significance of organ aging. *Fundam. Res.* **5**, 683–696 (2023).
7. López-Otín, C., Blasco, M. A., Partridge, L., Serrano, M. & Kroemer, G. The hallmarks of aging. *Cell* **153**, 1194–1217 (2013).
8. Rossi, M. & Abdelmohsen, K. The emergence of senescent surface biomarkers as senotherapeutic targets. *Cells* **10**, 1740 (2021).
9. Salminen, A., Kaarniranta, K. & Kauppinen, A. Inflammaging: disturbed interplay between autophagy and inflammasomes. *Aging* **4**, 166–175 (2012).
10. Ovadya, Y. & Krizhanovsky, V. Strategies targeting cellular senescence. *J. Clin. Invest.* **128**, 1247–1254 (2018).
11. Prata, L. G. P. L., Ovsyannikova, I. G., Tchkonina, T. & Kirkland, J. L. Senescent cell clearance by the immune system: Emerging therapeutic opportunities. *Semin. Immunol.* **40**, 101275 (2018).
12. Song, P., An, J. & Zou, M. H. Immune clearance of senescent cells to combat ageing and chronic diseases. *Cells* **9**, 671 (2020).
13. Alpert, A. et al. A clinically meaningful metric of immune age derived from high-dimensional longitudinal monitoring. *Nat. Med.* **25**, 487–495 (2019).
14. Dulken, B. W. et al. Single-cell analysis reveals T cell infiltration in old neurogenic niches. *Nature* **571**, 205–210 (2019).
15. Goronzy, J. J. & Weyand, C. M. Successful and maladaptive T cell aging. *Immunity* **46**, 364–378 (2017).
16. Hashimoto, K. et al. Single-cell transcriptomics reveals expansion of cytotoxic CD4 T cells in supercentenarians. *Proc. Natl. Acad. Sci. USA* **116**, 24242–24251 (2019).
17. Kimmel, J. C. et al. Murine single-cell RNA-seq reveals cell-identity and tissue-specific trajectories of aging. *Genome Res.* **29**, 2088–2103 (2019).
18. Ma, S., Wang, C., Mao, X. & Hao, Y. R Cells dysfunction associated with aging and autoimmune disease. *Front. Immunol.* **10**, 318 (2019).
19. Rodrigues-Pinto, E., Cardoso, H., Pimenta, S. & Macedo, G. The clinical significance of autoantibodies in hepatitis C patients submitted to interferon treatment. *Port. J. Gastroenterol.* **22**, 24–27 (2015).
20. Alves, A. S. & Bueno, V. Immunosenescence: participation of T lymphocytes and myeloid-derived suppressor cells in aging-related immune response changes. *Einstein* **17**, eRB4733 (2019).
21. Müller, L. & Di Benedetto, S. How immunosenescence and inflammaging may contribute to hyperinflammatory syndrome in COVID-19. *Int. J. Mol. Sci.* **22**, 12539 (2021).
22. Bordon, Y. TOX for tired T cells. *Nat. Rev. Immunol.* **19**, 476 (2019).
23. Konermann, S. A single cell transcriptomic atlas characterizes aging tissues in the mouse. *J. Clin. Sleep Med.* **12**, 263–266 (2016).
24. Angelidis, I. et al. An atlas of the aging lung mapped by single cell transcriptomics and deep tissue proteomics. *Nat. Commun.* **10**, 963 (2019).
25. Martinez-Jimenez, C. P. et al. Aging increases cell-to-cell transcriptional variability upon immune stimulation. *Science* **355**, 1433–1436 (2017).
26. Schaum, N. et al. Ageing hallmarks exhibit organ-specific temporal signatures. *Yearb. Paediatr. Endocrinol.* **583**, 596–602 (2021).
27. Ximerakis, M. et al. Single-cell transcriptomic profiling of the aging mouse brain. *Nat. Neurosci.* **22**, 1696–1708 (2019).
28. Mogilenko, D. A. et al. Comprehensive profiling of an aging immune system reveals clonal GZMK+ CD8+ T cells as conserved hallmark of inflammaging. *Immunity* **54**, 99–115.e12 (2021).
29. Hao, Y., O'Neill, P., Naradikian, M. S., Scholz, J. L. & Cancro, M. P. A B-cell subset uniquely responsive to innate stimuli accumulates in aged mice. *Blood* **118**, 1294–1304 (2011).
30. Franke, K. et al. SIRPα on mouse B1 cells restricts lymphoid tissue migration and natural antibody production. *Front. Immunol.* **11**, 570963 (2020).
31. Alfei, F. et al. TOX reinforces the phenotype and longevity of exhausted T cells in chronic viral infection. *Nature* **571**, 265–269 (2019).
32. Khan, O. et al. TOX transcriptionally and epigenetically programs CD8+ T cell exhaustion. *Nature* **571**, 211–218 (2019).
33. Joeckel, L. T. et al. Mouse granzyme K has pro-inflammatory potential. *Cell Death Differ.* **18**, 1112–1119 (2011).
34. Joeckel, L. T., Allison, C. C., Pellegrini, M., Bird, C. H. & Bird, P. I. Granzyme K-deficient mice show no evidence of impaired antiviral immunity. *Immunol. Cell Biol.* **95**, 676–683 (2017).
35. Faget, D. V., Ren, Q. & Stewart, S. A. Unmasking senescence: context-dependent effects of SASP in cancer. *Nat. Rev. Cancer* **19**, 439–453 (2019).
36. Xu, C. & Jackson, S. A. Machine learning and complex biological data The revolution of biological techniques and demands for new data mining methods. *Genome Biol.* **20**, 76 (2019).
37. Li, W. et al. Single-cell immune aging clocks reveal inter-individual heterogeneity during infection and vaccination. *Nat. Aging* **5**, 607–621 (2025).
38. Greener, J. G. et al. A guide to machine learning for biologists. *Nat. Rev. Mol. Cell Biol.* **23**, 40–55 (2022).
39. Xu, C. J. et al. The emerging landscape of dynamic DNA methylation in early childhood. *BMC Genomics* **18**, 25 (2017).
40. Horvath, S. DNA methylation age of human tissues and cell types. *Genome Biol.* **14**, 3156 (2013).
41. Zhang, Y. et al. DNA methylation signatures in peripheral blood strongly predict all-cause mortality. *Nat. Commun.* **8**, 14617 (2017).
42. Belsky, D. W. et al. Quantification of the pace of biological aging in humans through a blood test, the DunedinPoAm DNA methylation algorithm. *eLife* **9**, e54870 (2020).
43. Sayed, N. et al. Author correction: an inflammatory aging clock (IAge) based on deep learning tracks multimorbidity, immunosenescence. *Frailty Cardiovasc. Aging Nat. Aging* **1**, 598–615 (2021).
44. Pinchuk, L. M. & Filipov, N. M. Differential effects of age on circulating and splenic leukocyte population in C57BL/6 and BALB/c male mice. *Immunity Ageing* **5**, 1 (2008).
45. Smola, A. J. & Schölkopf, B. A tutorial on support vector regression. *Stat. Comput.* **14**, 199–222 (2004).
46. Drucker, H., Burges, C. J. C., Kaufman, L., Smola, A. & Vapnik, V. Support Vector Regression Machines. *Adv. Neural Inf. Process. Syst.* **28**, 779–784 (1997).
47. Perez-Correa, J.-F., Tharmapalan, V., Geiger, H. & Wagner, W. Epigenetic clocks for mice based on age-associated regions that are conserved between mouse strains and human. *Front. Cell Dev. Biol.* **10**, 902857 (2022).
48. Wang, T. et al. Epigenetic aging signatures in mice livers are slowed by dwarfism, calorie restriction and rapamycin treatment. *Genome Biol.* **18**, 57 (2017).
49. Galow, A.-M. & Peleg, S. How to slow down the ticking clock: age-associated epigenetic alterations and related interventions to extend life span. *Cells* **11**, 468 (2022).
50. Wijewarnasuriya, D., Beberitz, C., Lopez, A. V., Rafiq, S. & Brentjens, R. J. Excessive costimulation leads to dysfunction of adoptively transferred T cells Dinali. *Cancer Immunol Res.* **8**, 732–742 (2020).

51. Zumwalde, N. A., Domae, E., Mescher, M. F. & Shimizu, Y. ICAM-1 dependent homotypic aggregates regulate CD8 T cell effector function and differentiation during T cell activation. *J. Immunol.* **191**, 3681–3693 (2013).
52. Morimoto, C. & Schlossman, S. F. The structure and function of CD26 in the T-cell immune response. *Immunol. Rev.* **161**, 55–70 (1998).
53. Vukelic, I. et al. CD26 deficiency controls macrophage polarization markers and signal transducers during colitis development and resolution. *Int. J. Mol. Sci.* **23**, 5506 (2022).
54. Bozorgmehr, N., Hnatiuk, M., Peters, A. C. & Elahi, S. Depletion of polyfunctional CD26highCD8+ T cells repertoire in chronic lymphocytic leukemia. *Exp. Hematol. Oncol.* **12**, 13 (2023).
55. Liston, A. & Aloulou, M. A fresh look at a neglected regulatory lineage: CD8+Foxp3+ regulatory T cells. *Immunol. Lett.* **247**, 22–26 (2022).
56. Soto-Hereder, G. et al. KLRG1 identifies regulatory T cells with mitochondrial alterations that accumulate with aging. *Nat. Aging* **5**, 799–815 (2022).
57. Lages, C. S., Lewkowich, I., Sproles, A., Wills-Karp, M. & Chougnet, C. Partial restoration of T cell function in aged mice by in vitro blockade of the PD-1/PD-L1 pathway. *Aging Cell* **9**, 785–798 (2010).
58. Kohlmeier, J. E., Rumsey, L. M., Chan, M. A. & Benedict, S. H. The outcome of T-cell costimulation through intercellular adhesion molecule-1 differs from costimulation through leucocyte function-associated antigen-1. *Immunology* **108**, 152–157 (2003).
59. Chirathaworn, C. et al. Stimulation through intercellular adhesion molecule-1 provides a second signal for T cell activation. *J. Immunol.* **168**, 5530–5537 (2002).
60. Deng, B. et al. FOXP3+ regulatory T cells and age-related diseases. *FEBS J.* **289**, 319–335 (2022).
61. Garg, S. K. et al. Aging is associated with increased regulatory T-cell function. *Aging Cell* **13**, 441–448 (2014).
62. Jain, A., Sturmlechner, I., Weyand, C. M. & Goronzy, J. J. Heterogeneity of memory T cells in aging. *Front. Immunol.* **14**, 1250916 (2023).
63. Silaeva, Y. Y. et al. Decrease in pool of T lymphocytes with surface phenotypes of effector and central memory cells under influence of TCR transgenic  $\beta$ -chain expression. *Biochemistry* **78**, 549–559 (2013).
64. Benichou, G., Gonzalez, B., Marino, J., Ayasoufi, K. & Valujskikh, A. Role of memory T cells in allograft rejection and tolerance. *Front. Immunol.* **8**, 170 (2017).
65. Collison, J. The ABCs of autoimmune disease. *Nat. Rev. Rheumatol.* **14**, 248 (2018).
66. Zambrano-Zaragoza, J. F., Romo-Martínez, E. J., Durán-Avelar, M. D. J., García-Magallanes, N. & Vibanco-Pérez, N. Th17 cells in autoimmune and infectious diseases. *Int. J. Inflamm.* **2014**, 651503 (2014).
67. Dennig, D., Lacerda, J., Yan, Y., Gasparetto, C. & O'Reilly, R. J. ICAM-1 (CD54) expression on B lymphocytes is associated with their costimulatory function and can be increased by coactivation with IL-1 and IL-7. *Cell. Immunol.* **156**, 414–423 (1994).
68. Basak, D., Pal, S. & Patranabis, D. C. Support vector regression. *Neural Inf. Process.* **11**, 203–224 (2007).
69. Tao, D. et al. Support vector regression model for the prediction of buildings' maximum seismic response based on real monitoring data. *Sci. Rep.* **14**, 29874 (2024).
70. Karal, O.  $\epsilon$ -Insensitive least squares support vector regression with sequential minimal optimization. *Comput. Open* **1**, 2350004 (2023).
71. Liang, Z., Dong, X., Zhang, Z., Zhang, Q. & Zhao, Y. Age-related thymic involution: mechanisms and functional impact. *Aging Cell* **21**, e13671 (2022).
72. Tatari-Calderone, Z. et al. Age-related accumulation of T cells with markers of relatively stronger autoreactivity leads to functional erosion of T cells. *BMC Immunol.* **13**, 8 (2012).
73. Agrawal, A. & Gupta, S. Impact of aging on dendritic cell functions in humans. *Ageing Res. Rev.* **10**, 336–345 (2011).
74. Wong, C. P., Magnusson, K. R. & Ho, E. Aging is associated with altered dendritic cells subset distribution and impaired proinflammatory cytokine production. *Exp. Gerontol.* **45**, 163–169 (2010).
75. Kerr, J. A. et al. Global, regional, and national prevalence of child and adolescent overweight and obesity, 1990–2021, with forecasts to 2050: a forecasting study for the Global Burden of Disease Study 2021. *Lancet* **405**, 785–812 (2025).
76. Villareal, D. T. Obesity and accelerated aging. *J. Nutr. Heal. Aging* **27**, 312–313 (2023).
77. Franceschi, C. & Campisi, J. Chronic inflammation (Inflammaging) and its potential contribution to age-associated diseases. *J. Gerontol. Ser. A Biol. Sci. Med. Sci.* **69**, S4–S9 (2014).
78. Tam, B. T., Morais, J. A. & Santosa, S. Obesity and ageing: two sides of the same coin. *Obes. Rev.* **21**, e12991 (2020).
79. Salvestrini, V., Sell, C. & Lorenzini, A. Obesity may accelerate the aging process. *Front. Endocrinol.* **10**, 266 (2019).
80. Johnson, R. M. G., Wen, T. & Dong, H. Bidirectional signals of PD-L1 in T cells that fraternize with cancer cells. *Nat. Immunol.* **21**, 365–366 (2020).
81. Rothstein, T. L., Griffin, D. O., Holodick, N. E., Quach, T. D. & Kaku, H. Human B-1 cells take the stage. *Ann. N. Y. Acad. Sci.* **1285**, 97–114 (2013).
82. Sim, H. B. et al. Analysis of immune cell remodeling and functional alterations induced by aging and obesity in mice. *Int. Immunopharmacol.* **167**, 115735 (2025).
83. Liu, X., Lv, J., Wang, H., Zheng, Y. & Su, W. Functional analysis of human circulating immune cells based on high-dimensional mass cytometry. *STAR Protoc.* **3**, 101310 (2022).
84. Bagwell, C. B. et al. Automated data cleanup for mass cytometry. *Cytom. Part A* **97**, 184–198 (2020).
85. Tang, D. et al. SRplot: a free online platform for data visualization and graphing. *PLoS ONE* **18**, e0294236 (2023).
86. Kohavi, R. A study of cross-validation and bootstrap for accuracy estimation and model selection. *IJCAI* **14**, 1137–1145 (1995).
87. Arlot, S. & Celisse, A. A survey of cross-validation procedures for model selection. *Statist. Surv.* **4**, 40–79 (2010).

## Acknowledgements

This research was carried out with the support of a National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) (NRF-2021R1C1C1006874 to J.-J.K.). Mass cytometry was performed using Helios™ system (Fluidigm, South San Francisco, CA) according to the manufacturer's instructions at the Korea Basic Science Institute (OC106, Ochang, Republic of Korea).

## Author contributions

D.-J.C., S.-H.J., and J.-J.K. conceptualized this project and supervised the overall experiments. H.B.S., J.-H.J., and S.-H.J. performed the experiments and bioinformatics analysis. D.-H.P., Y.-J.C., J.Y.H., J.-B.L., and H.B.S. contributed to collected the data. H.B.S. S.-K.M., H.S.C. S.-H.J., and J.-J.K. contributed to the analysis and interpretation. K.-B.L. resource provision. H.B.S., J.-H.J., S.C.R., D.-J.C., S.-H.J., and J.-J.K. wrote, reviewed, and edited the manuscript. All of the authors have read the manuscript and approved the final version of the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-025-67393-1>.

**Correspondence** and requests for materials should be addressed to Dong-Jo Chang, Seung-Hyun Jeong or Jong-Jin Kim.

**Peer review information** *Nature Communications* thanks Natalie Riddell, who co-reviewed with Stephanie Watkins, Jessica Lancaster and the other, anonymous, reviewers for their contribution to the peer review of this work. A peer review file is available.

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

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

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

© The Author(s) 2025