nature communications

Article

https://doi.org/10.1038/s41467-024-52522-z

Quantifiable blood TCR repertoire components associate with immune aging

Received: 13 June 2024

Accepted: 11 September 2024

Published online: 17 September 2024

Check for updates

Jing Hu[1,2], Mingyao Pan[1], Brett Reid[3], Shelley Tworoger[3,4] & Bo Li[1,2]

T cell senescence alters the homeostasis of distinct T cell populations and results in decayed adaptive immune protection in older individuals, but a link between aging and dynamic T cell clone changes has not been made. Here, using a newly developed computational framework, Repertoire Functional Units (RFU), we investigate over 6500 publicly available TCR repertoire sequencing samples from multiple human cohorts and identify age-associated RFUs consistently across different cohorts. Quantification of RFU reduction with aging reveals accelerated loss under immunosuppressive conditions. Systematic analysis of age-associated RFUs in clinical samples manifests a potential link between these RFUs and improved clinical outcomes, such as lower ICU admission and reduced risk of complications, during acute viral infections. Finally, patients receiving bone marrow transplantation show a secondary expansion of the age-associated clones upon stem cell transfer from younger donors. Together, our results suggest the existence of a 'TCR clock' that could reflect the immune functions in aging populations.

Deterioration of the immune system is one of the key features of aging[1]. Previous work has uncovered systematic changes in the innate and adaptive cellular compartments in older individuals[2,3]. Many studies have further focused on T cell senescence given their critical role in host defense against external pathogens and cancer[4,5]. Phenotypically, exhaustion-like T cells emerge and accumulate in aged tissues, with an altered, pro-inflammatory cytokine production program[6]. Quantitatively, there is continuous shrinkage of naïve T cell populations due to thymic atrophy[7,8], which is usually accompanied by the expansion of memory clones in older individuals[9]. In addition, decreased abundance of innate-like T cell compartments was also observed, such as $\gamma\delta$ T cells[10] and mucosal-associated invariant T cells[11]. Consistently, studies using the immunogenomics profiling of T cell receptors (TCR) confirmed a robust decrease in repertoire diversity with aging[9,12–14]. Concomitant with the reduction of repertoire diversity, certain clones become more dominant while others decline or disappear, resulting in imbalanced naïve-memory ratio and increased clonality[13,15,16].

Compared to cross-sectional studies, research based on longitudinal cohorts could estimate age-associated changes more precisely given the inter-individual heterogeneity of TCR repertoires[17]. Yoshida et al. profiled TCRβ repertoires of CD4+ and CD8 + T cells collected from six healthy donors over three visits, and reported significantly decreased CDR3β diversity and increased frequencies of clonal populations with age for CD8 + T cells[18]. Another longitudinal cohort with more donors and larger age spans further confirmed the reduction of richness in the naïve CD4+ and CD8+ subsets, but not in memory T cells[19]. Contrarily, retention of the same TCRs is found in different timepoints of the same donors, with more prominence in CD8+ memory cells than in CD4+ or naïve cells[19].

While these studies provided critical insights on the overall decay of adaptive immune functions in aging populations, there has been no report of the age impact on specific T cell clones defined by the T cell receptor sequences. Such an investigation is challenging given the vast diversity of the human TCR repertoire[20], yet it holds the promise to uncover shared 'aging antigens' that could be targeted in the older population for improved immune protection[6]. To address this challenge, we have developed a new TCR analysis method, Repertoire Functional Unit (RFU)[21], which dissects the hypervariable TCR

[1]Department of Pathology and Laboratory Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA. [2]Department of Pathology and Laboratory Medicine, Children's Hospital of Philadelphia, Philadelphia, PA, USA. [3]Department of Cancer Epidemiology, Moffitt Cancer Center, Tampa, FL, USA. [4]Knight Cancer Institute and Division of Oncological Sciences, Oregon Health and Science University, Portland, OR, USA.
e-mail: lib3@chop.edu

repertoire into quantifiable segments, or RFUs, that can be compared across different individuals.

Here in this work, we apply the RFU method to investigate a large number of repertoire samples deeply sequenced for the β chain complementarity determining region 3 (CDR3) of TCRs and systematically study the age associations of individual RFUs. Our analysis reveals 13 RFUs showing dynamic changes with age and confirms this observation in multiple independent cohorts. These clones carry the features of mucosal-associated invariant T cells (MAIT) that recognize antigens derived from bacteria. Through investigation of the clinical impact of these RFUs, we confirm their prognostic value during acute viral infections and quantify their loss under normal or pathological conditions. Our work provides novel insights over immune aging that can open up potential opportunities to maintain a 'younger' T cell repertoire in the older population.

Results

Robustly age-associated TCR signatures in multiple human cohorts

We first implemented Repertoire Functional Unit, or RFU, our recent computational method to analyze the TCR repertoire samples[21]. In brief, each TCR sequence was projected to a 500-dimensional Euclidean space, with shorter distance between a pair of TCRs representing higher similarity. The embedding space was then divided into 5000 conservative TCR neighborhoods, with the centroid of each neighborhood defined as an RFU. A new TCR-seq sample can be converted into a 5000-length numeric vector of RFU counts, which can be compared across multiple samples in a large cohort.

We then investigated the age associations of RFUs using three TCR-seq blood sample cohorts, including Nolan cohort[22] of 1414 adults with COVID-19, Emerson cohort of 666 samples mainly from healthy adults[23] and Mitchell cohort of 359 samples from children[24]. RFU-by-sample matrix (hereby denoted as RFU matrix) of each cohort was calculated using the above method. We compared the age associations of the two adult cohorts and observed consistency of a subset of negatively correlated RFUs in both cohorts (Fig. 1a, Supplementary Data 1). We defined the 13 RFUs with Spearman's $\rho \le -0.2$ in the Emerson cohort and $\rho \le -0.3$ in the Nolan cohort as 'essential RFU', or eRFU (Fig. 1b and Supplementary Fig. 1a). The T cell clones of these eRFUs demonstrated an approximately 50% reduction by age 80 or above compared to early adulthood (Fig. 1b). This trend was reversed when we investigated the Mitchell cohort of younger individuals: all the adult age-related eRFUs were positively associated with age between birth to early 20 s (Fig. 1c, d). Combining adult and children cohorts, we obtained the full spectrum of non-linear age association (Fig. 1e). All the 13 eRFUs exhibited dynamic changes with chronological age, experiencing rapid expansion after birth. They reached peaks around age 27 and then gradually decreased, ultimately returning to the birth level as participants aged.

We next explored the nature of eRFU T cells. Public TCRs, defined as TCRs shared across multiple individuals, appear to be an important part of the TCR repertoire[25]. To explore the relationship between eRFU and public TCRs, we repeated the previous work[12] to define public TCRs. We found that eRFUs were not enriched for public TCRs compared to a randomly selected RFU set (Fig. 1f) and the public TCRs were not significantly associated with age (Fig. 1g). Next, to understand how eRFUs are shaped by genetic factors, we took advantage of a TCR-seq sample cohort of identical twins[26]. Interestingly, while the TCR repertoire is overall more similar within than between twins (Supplementary Fig. 1b), this heritability was not observed for the eRFUs (Fig. 1h). Further, intra-twin over inter-twin similarity calculated as Spearman's ρ between a pair of individuals using eRFUs were not higher when compared to those using randomly selected RFU sets (Fig. 1i). Finally, we sought to determine whether eRFUs had a specific T cell phenotypic subset. We analyzed two sample cohorts with memory CD4 and CD8 T cell repertoires separately sequenced[27,28] and discovered that eRFUs were consistently enriched for the CD8+ population (Fig. 1j–l), suggesting that eRFUs clones are potentially cytotoxic. However, examination of their associations to the common MHC-I alleles did not identify any significant associations (Supplementary Fig. 2a). These observations suggested that eRFUs are non-public and MHC-independent CD8 T cells.

MAIT signatures of T cells with eRFU receptors

Previous studies indicated that mucosal associated invariant T (MAIT) cells possessed the TCR motifs observed in the eRFUs with Aspartic or Glutamic acid on the 5th position[29]. Consistent with our observation, MAIT cells were not MHC-I restricted as they recognize antigens bound by MR1[30]. Interestingly, we identified one putative MAIT TCR, M33.64 that carries the sequence signatures of eRFU 841 (Supplementary Fig. 2b). M33.64 binds to 5-(2-oxopropylideneamino)−6-d-ribitylaminouracil, or 5-OP-RU (Supplementary Fig. 2c), a potent MAIT activator derived from vitamin B metabolism in bacteria[31]. Further, although MAIT cells are known for their invariant α chain[32], conserved patterns in the β chain have also been observed[29], consistent with our findings. We therefore investigated a single cell RNA-seq dataset paired with TCR sequences from human T cells of diverse phenotypes[33]. T cells assigned to eRFUs were enriched in cluster 7 (Supplementary Fig. 2d), which overexpressed putative markers of MAIT cells (Supplementary Fig. 2e). We further selected all the 2619 annotated MAIT cell TCRs and analyzed each of the eRFUs. 7 out of the top 40 RFUs enriched in the MAIT TCRs (O/E ≥ 50) were eRFUs (Supplementary Fig. 2f). These results indicated that most eRFU T cells belong to a subset of CD8 + MAIT cells.

To further investigate the nature of eRFUs, we analyzed another recent scRNA-seq dataset[34], which contained paired scTCR-seq samples from flow-sorted MAIT cells and conventional memory T cells (Tmem). We confirmed that eRFUs were dominantly expressed by MAIT cells (Fig. 2a), with an enrichment in the TRAV1-2+ subset (Fig. 2b). Unbiased differential gene expression analysis revealed more upregulated genes in the eRFU cells compared to the non-eRFU MAIT cells (Fig. 2c), suggesting a globally different transcriptional program between the two groups. Particularly, putative T cell stemness markers *TCF7*, *BACH2* and *KLF2*[35–37] were significantly down-regulated in the eRFU group (Fig. 2d). Unbiased analysis using signature gene sets (up-or down-regulated in CD8 stem T cells)[38] confirmed a low stemness state of eRFU MAITs (Fig. 2e). Finally, pseudotime analysis[39] with direction determined by CytoTRACE[40] revealed that eRFU cells are significantly enriched at the most-differentiated end of the trajectory (Fig. 2f, g). These results collectively indicated that eRFU-expressing cells belong to a more differentiated group of CD8 + MAIT cells.

The same dataset contained individuals aged from 27 to 65, allowing for the investigation of how stemness of eRFU MAITs trends with age. First, T cell stemness markers (*TCF7*, *BACH2*, *KLF2*)[35–37] consistently decrease with age (Supplementary Fig. 3a). The activity of transcription factors BACH2 and KLF2, measured by the expression levels of their regulating targets by SCENIC[41], showed a similar trend (Supplementary Fig. 3b). Signature genes upregulated CD8 stem cells decrease with age in eRFU MAITs, while an opposite trend was observed for the down-regulated genes (Supplementary Fig. 3c, d). These results potentially indicated that the faster shrinkage of eRFU clones than the other MAIT cells is likely due to their relatively lower and continuously decaying stemness with aging.

Quantification of eRFU loss with age in adults

We next sought to quantify the rate of eRFU decrease in adult cohorts. First, we observed a linear, homoscedastic relationship between age and log-scaled total eRFU summation in adults (Supplementary Fig. 4a). In the Nolan cohort, $\log_{10}$ eRFU sum is estimated to decay at

https://doi.org/10.1038/s41467-024-52522-z

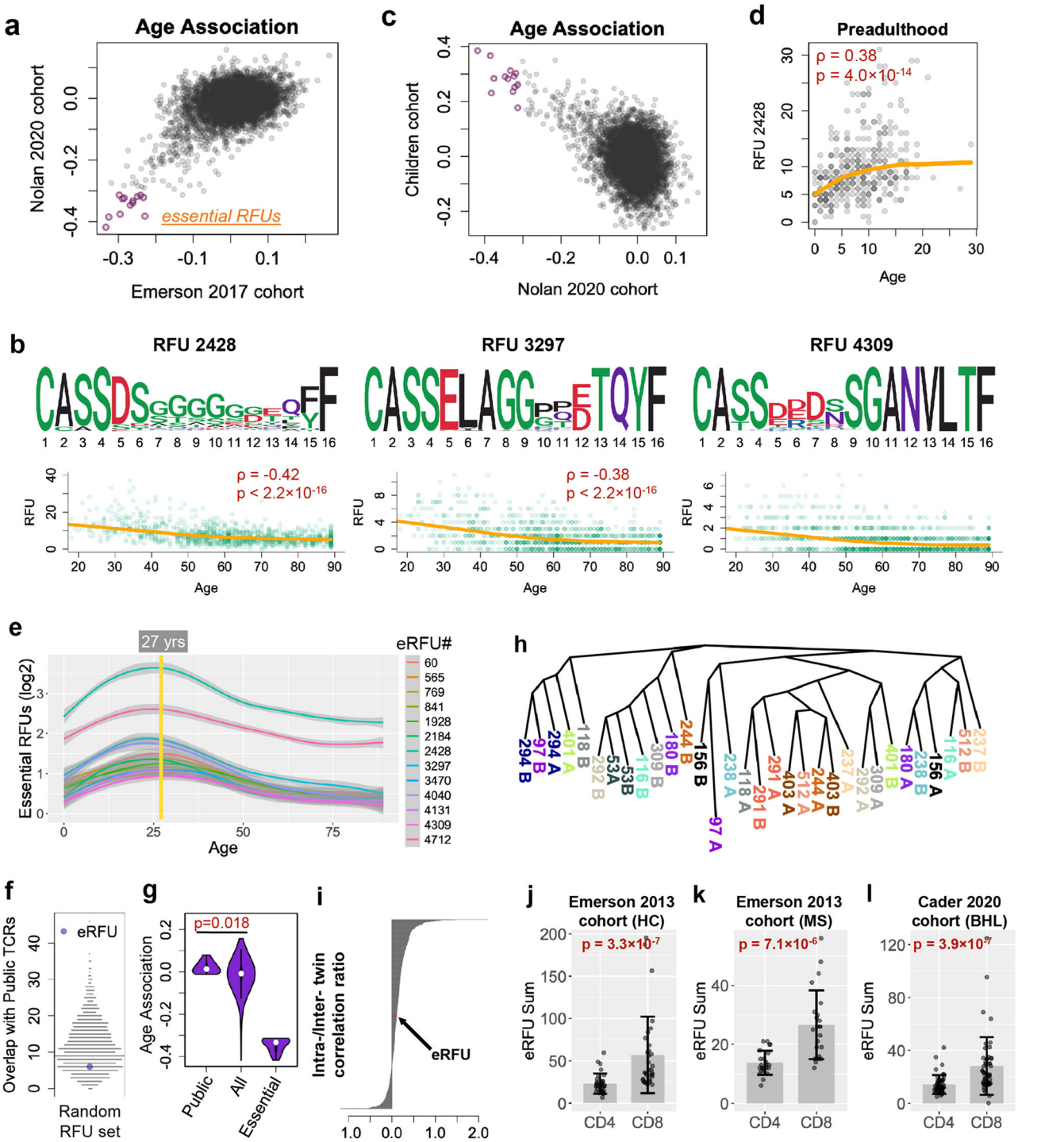
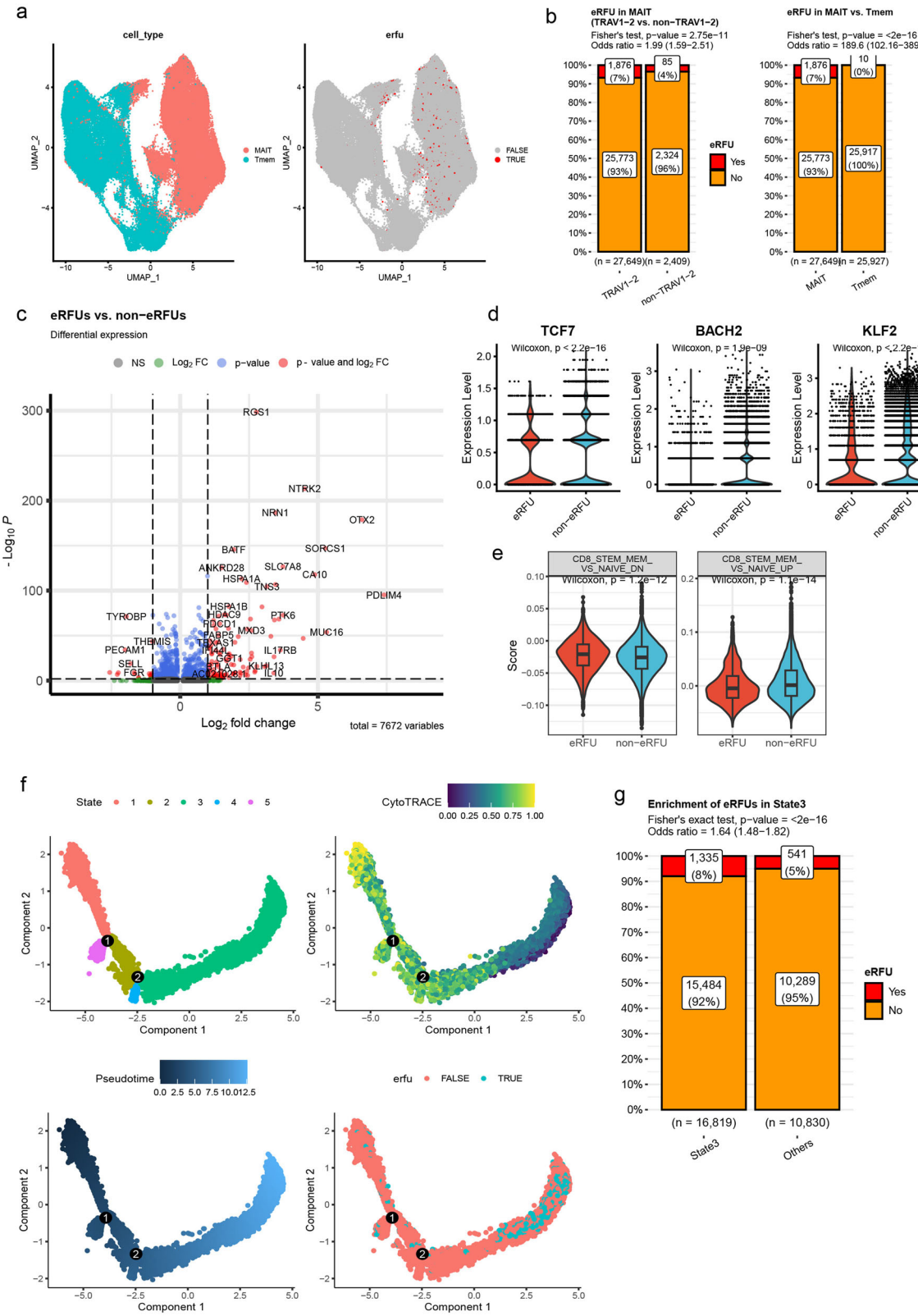


Fig. 1 | Identification and characterization of age-associated essential RFUs (eRFUs). a Scatter plot showing the aligned age associations calculated from two large sample cohorts. Selection criteria of eRFUs were described in the main text. RFUs were defined by mapping the top 10,000 most abundant TCR clones in a repertoire to the embedding space. Statistical significance was evaluated using two-sided Spearman's correlation test, with FDR corrected using the Benjamini-Hochberg approach. **b** Sequence logo plot paired with scatter plot showing the direct age association for 3 eRFUs. **c** Scatter plot showing the anti-aligned age associations from one adult and one childhood cohort, with eRFUs labeled with red circles. **d** Positive age association of a selected eRFU in the childhood cohort. Statistical significance of the association in **b** and **d** was evaluated using two-sided Spearman's correlation test, with FDR corrected using the Benjamini-Hochberg approach. **e** Smoothed lines showing dynamic changes with age for all 13 eRFUs. The shaded areas show 95% confidence intervals, which were calculated using 95% standard error estimated by Loess polynomial fitting. **f** Beeswarm plot showing the distribution of overlaps of public TCRs with each of 1000 random RFU set. eRFU was colored in purple. **g**) Violin plots comparing age associations for 3 groups of RFUs. Statistical significance was estimated with one-way ANOVA. **h** Neighbor joining tree showing the relationships between pairs of twins with distance matrix calculated using the 13 eRFUs. **i**) Barplot showing the ratio of intra-/inter- twin associations of random RFU sets or eRFU. Associations were calculated as Spearman's correlation using RFU as markers between a pair of individuals. **j**–**l** Boxplots showing enrichment of eRFUs into CD4 or CD8 subsets in 3 independent adult cohorts. Data are presented as mean values ± SD. Statistical significance was evaluated using two-sided Wilcoxon rank sum test. Source data are provided as a Source Data file.

https://doi.org/10.1038/s41467-024-52522-z



the rate of 0.0092 ± 0.00044 (mean ± s.d., Fig. 3a), which is approximately 2.1% loss per year. The second discovery set, Emerson 2017 cohort, provided a similar estimation of −0.0088, *i.e.* 2.0% annual loss (Fig. 3b). To confirm these results, we analyzed two additional adult patient cohorts: Greenberger 2022 cohort[42] with 480 subjects and Britanova 2016 cohort with 32 individuals[12]. The Britanova cohort was profiled using a different data generation method based on RNA, while all the other cohorts were profiled using genomic DNA. Interestingly, we observed similar rate of decay in both cohorts, with −0.0076 (1.7%) and −0.011 (2.5%) respectively (Fig. 3c, d). We observed similar trend using another large human adult cohort with categorical age information[43], which saw a rate of −0.093 per 10 years on the $\log_{10}$ scale (Fig. 3e). Further analysis of racial groups revealed a potentially faster eRFU loss among Asian and Pacific Islanders (Supplementary

Fig. 2 | MAIT signatures and differentiation status of eRFU cells. a UMAP plot showing the distribution of MAIT or memory T cells in the gene expression space (left) and the distribution of eRFU expressing T cells on the UMAP (right). **b** Percentage plots for eRFU expressing T cells enrichment in TRAV1-2 vs non-TRAV1-2 (left) T cell subsets or in MAIT vs Tmem (right). Statistical significance was evaluated using two-sided Fisher's exact test. **c** Volcano plot showing the differentially expressed genes between eRFU and non-eRFU cells. Statistical significance was evaluated using two-sided Wilcoxon rank sum test, with FDR corrected using the Benjamini-Hochberg approach. Red color marks absolute log2 fold change greater than 1. Blue color indicates statistical significance at FDR = 0.05. **d** Violin plot showing the distributions of three putative T cell stemness markers. **e** Violin plot showing the distribution of two scores measuring the signature genes that either down- (left) or up- (right) regulated in CD8 stem memory cells vs naïve CD8 T cells. Statistical significance of the difference in **d** and **e** was evaluated using two-sided Wilcoxon rank sum test. **f** Monocle pseudotime trajectory plot of all the MAIT cells with direction of differentiation determined by CytoTRACE, where lower value indicates higher differentiation status. **g** Percentage plot showing eRFU cells enrichment in state 3. Statistical significance was evaluated using two-sided Fisher's exact test. Source data are provided as a Source Data file.

Fig. 4b), though statistical significance was not reached due to small sample size. These results suggested that total count of eRFUs decreases by approximately 2% each year in the general population aged from 30 s till 80 s.

All above cohorts were derived from blood samples. We further investigated the aging trends of diverse organs using a postmortem patient cohort with TCR repertoires sequenced for skin, liver, spleen, and gastrointestinal (GI) track[44]. We observed that blood, skin and esophagus shared similar eRFU levels, whereas colon, liver and spleen were closely correlated (Fig. 3f), though no p value passed FDR level at 0.05 due to small sample size.

Accelerated loss of eRFUs under immunosuppressive conditions

We next investigated eRFUs loss among individuals with HCMV, HIV or COVID-19, in contrast to the healthy donors. For HCMV infection, HCMV+ was defined as patients with a positive serology. Interestingly, the age-association curves of patients with HCMV or SARS-CoV2 infections were similar to the healthy controls, while HIV infection was associated with a 37% drop in the overall eRFU levels (Fig. 4a). Of the 13 eRFUs, 2428 and 4712 showed highest age-adjusted reduction in the HIV+ group (Fig. 4b). These HIV patients (Towlerton 2020) received anti-retroviral treatment (ART)[45], allowing us to investigate if therapeutic intervention could reverse this trend. Indeed, we observed a significant increase of eRFUs 2 years post ART (Fig. 4c), which became more significant after correcting for participant age and individual baseline levels using a mixed effect model (Fig. 4d). Based on these results, we speculated that immunosuppressive conditions lead to accelerated immune aging, manifested by faster reduction of the eRFU levels that could be rescued by therapy.

To test this hypothesis, we leveraged the Nolan 2020 cohort, which comprehensively documented the medical records of over 1400 COVID-19 patients. A subset of the patients reported usage of immunosuppressant drugs, such as corticosteroids to treat asthma[46]. We observed significantly reduced levels of eRFUs in those patients after adjustment for age (Fig. 4e). In addition, 24 patients in this cohort were diagnosed with cancer before the time of blood draw. These patients may have been immunocompromised due to cancer-induced immune exhaustion[47] or exposure to immunosuppressive therapies that are common for cancers[48]. Indeed, we observed a significant decrease of overall eRFU counts in these patients (Fig. 4f). Interestingly, we noticed that use of immunosuppressant had a stronger influence among younger (35–50 yrs) individuals, while cancer occurrence in the older (65–80 yrs) patients was associated with higher reduction of eRFUs (Fig. 4e, f). Together, our analysis indicated that immunosuppressive conditions might further lower the eRFU levels.

Clinical impact of eRFUs in COVID-19 patients

To further understand the functional impact of eRFUs in the context of diseases, we performed a systematic analysis of TCR repertoire samples from several recent COVID-19 patient cohorts[22,42,49]. First, among the adult COVID-19 patients, higher overall eRFU level was associated with reduced risk of ICU admission when controlled for patient age (Fig. 5a). Within hospitalized patients, we observed a suggestively higher eRFU count associated with fewer days in the hospital (Fig. 5b). Quantitatively, at the same age, an increment of one eRFU clone reduced the odds of ICU admission by 6% and hospitalization by 0.29 days. Next, we investigated a cohort of pediatric patients[49] who developed multisystem inflammatory syndrome in children (MIS-C), a rare but serious condition associated with COVID-19. Consistent with the analysis in adults, higher eRFU level in children was associated with lower risk of MIS occurrence (Fig. 5c). Further, within the patients who developed MIS-C, higher eRFU count appeared to be associated with lower risk of cardiac involvement (Fig. 5d), although this observation did not hold for neurologic involvement (Supplementary Fig. 5a). Notably, eRFUs showed different correlations with age in adults (inverse) and children (positive), yet in both groups, eRFUs were associated with better clinical outcomes. Finally, we studied the outcome of COVID-19 mRNA vaccination in a patient cohort with hematologic malignancies[42]. Total eRFU count was positively correlated with anti-Spike protein antibody titer (Fig. 5e; Spearman's $\rho = 0.21$, $p = 2.3 \times 10^{-6}$), for both types of mRNA vaccines studied (BNT and Moderna) (Supplementary Fig. 5b). We controlled for related factors (age, gap between 1st and 2nd vaccines and COVID-19 status) using linear regression and confirmed that eRFU levels independently predicted higher antibody production after vaccination (Fig. 5f). Together, our results suggested that higher eRFU counts were related to better outcomes in patients with COVID-19 and improved vaccine response.

Secondary expansion of eRFUs following bone marrow transplantation

Given the potential predictive value of eRFUs in the clinical outcomes during acute infection, we next sought to investigate which factors influenced the expansion of eRFUs. For this purpose, we obtained TCR-seq samples from leukemia patients who received bone marrow transplantation, which allowed us to separately investigate the age associations of both donor and recipient. We first analyzed the Kanakry 2016 cohort[50], including blood samples from 16 healthy donors and 46 recipients. We confirmed that recipient blood eRFU levels were negatively associated with age in the donors, and yet uncorrelated with recipient patient age (Fig. 6a). In addition, using another cohort[51] (Pagliuca 2021), we observed that in the patients who received peripheral blood stem cell (PSC) transfer, the association with donor age was no longer significant, in contrast to those who received bone marrow transplant (BM) (Fig. 6b). These findings, together with our previous observations (Fig. 3, Supplementary Fig. 3), suggested that senescence of the hematopoietic stem cells (HSC)[52] may be a cause for the decreased eRFU levels in older populations with hematologic malignancies.

We next studied the changes of eRFUs after bone marrow transplantation. Interestingly, in both cohorts, we observed a significant increase of eRFUs by 2-fold within the first 100 days (Fig. 6c, d). RFUs 60, 2428 and 4712 showed highest positive correlations with post-transplantation time (Supplementary Fig. 6). Notably, of these, RFUs 2428 and 4712 also displayed largest decrease in the Towlerton cohort of HIV patients (Fig. 4b). We used the term 'primary expansion' to refer the increase of eRFUs after birth, and 'secondary expansion' the

https://doi.org/10.1038/s41467-024-52522-z

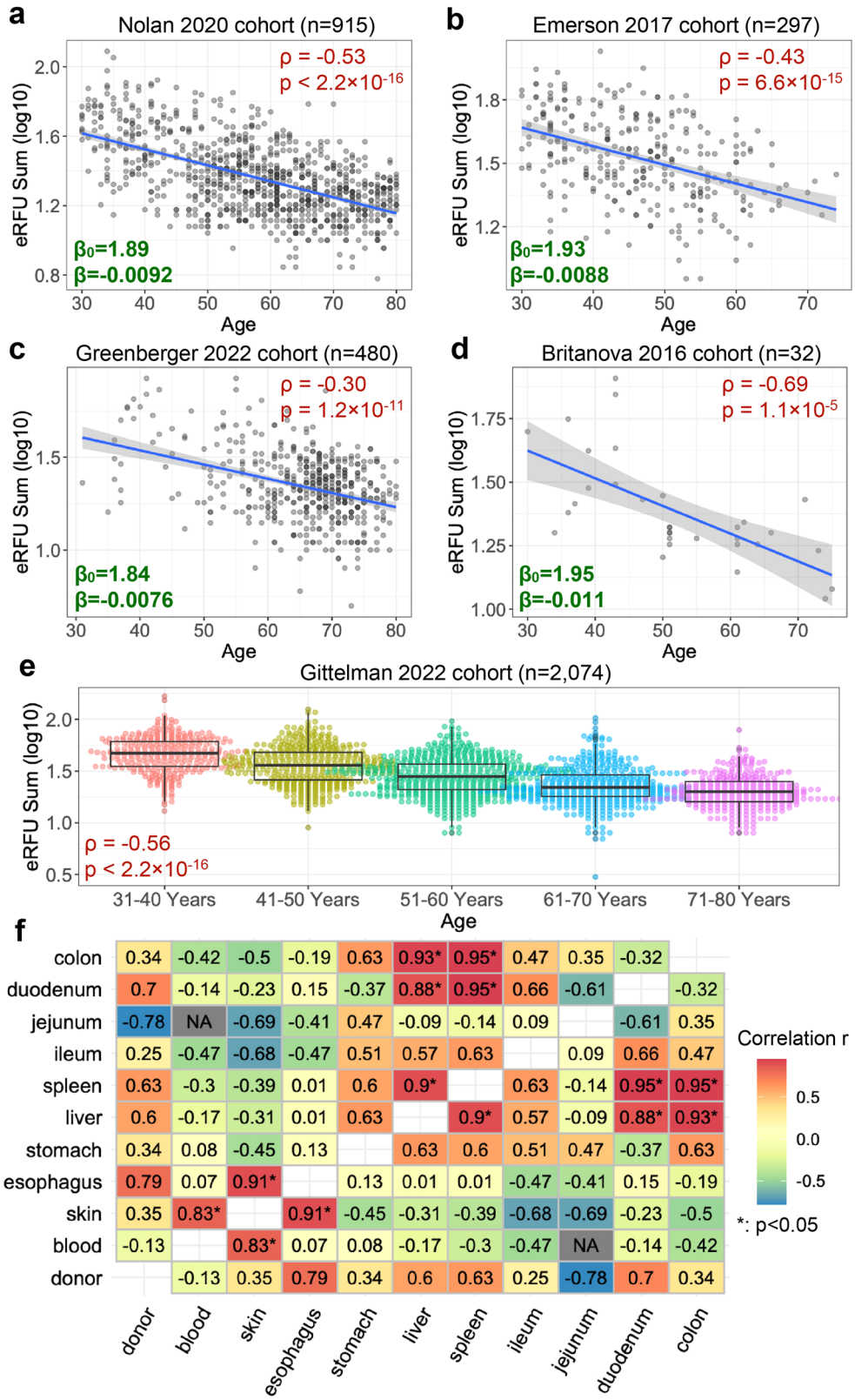


Fig. 3 | Quantification of eRFU loss with age in adult cohorts. a–d Regression analysis of eRFU sum on age in four different adult cohorts. All individuals are within the range of 30 to 80 years. The solid lines denote linear regression fits between eRFU sum and age. The shaded areas show 95% confidence intervals, which were calculated using 1.96 standard deviation estimated from the model. Top right text boxes show the Spearman's correlation and p values between eRFU sum and age. **e** Association of eRFU sum with age group information. All box plots display the median, interquartile range (IQR), whiskers extending up to 1.5 times the IQR, and individual data points representing the minimum and maximum values. Two-sided Spearman's correlation test was applied to evaluate statistical significance. The *p*-values in **a** and **e** are less than 2.2×10^{-16} and are denoted as $< 2.2 \times 10^{-16}$. **f** Heatmap showing correlation of eRFU sums between a pair of tissue samples collected from 10 patients (postmortem). Statistical significance was evaluated using two-sided Pearson's correlation test. Source data are provided as a Source Data file.

https://doi.org/10.1038/s41467-024-52522-z

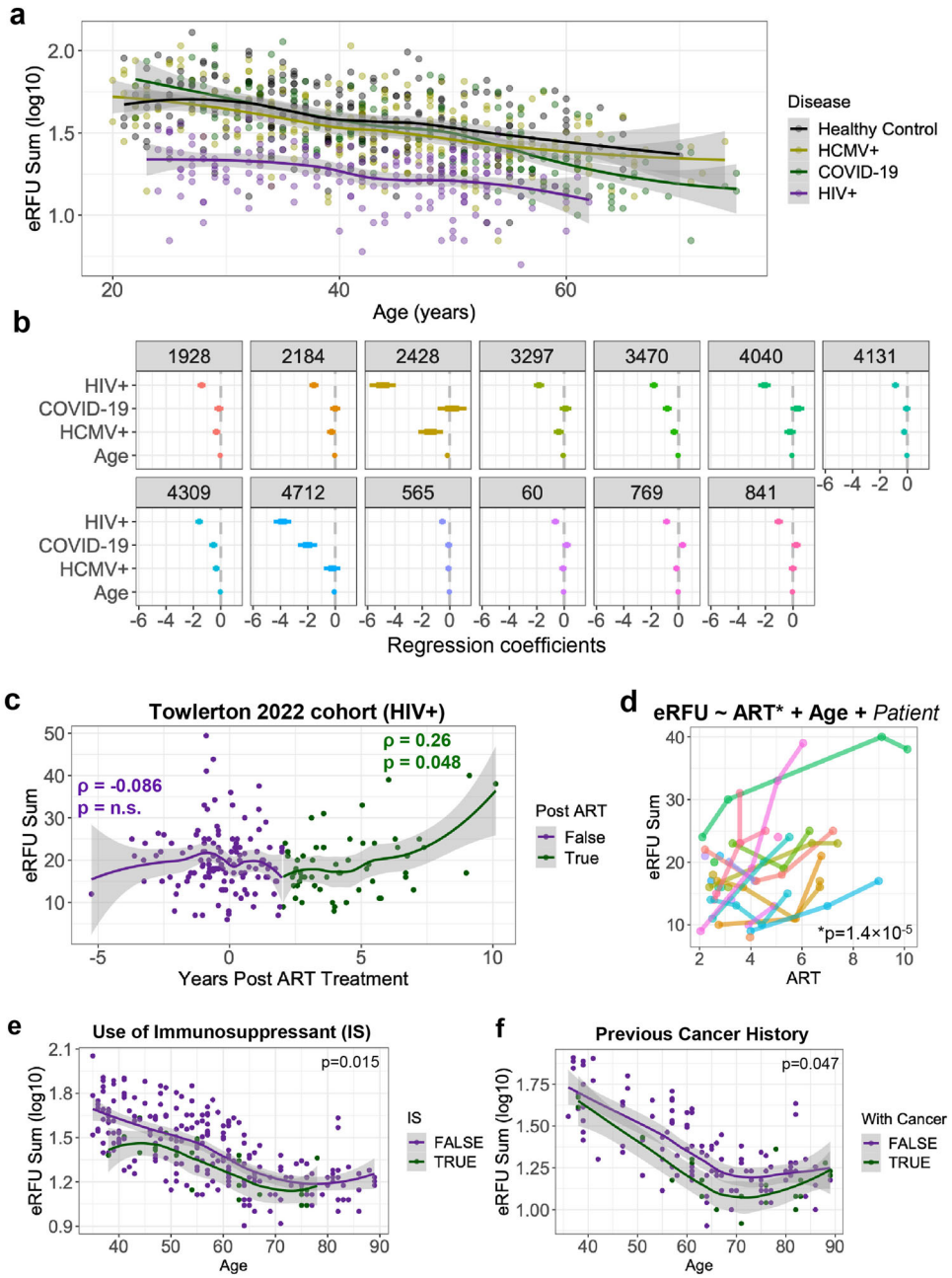


Fig. 4 | eRFUs are decreased in patients under immunosuppressive conditions. **a** Scatter plot showing eRFU loss with age in different patient groups. **b** Coefficient plot for all 13 eRFUs from linear regression with age and viral infection categories as covariates. **c** eRFU sum dynamics pre- or post- ART treatment. Color change marks 2-year post ART. Two-sided Spearman's correlation test was implemented to estimate statistical significance. **d** Dynamics of eRFU sum post ART treatment for each patient. Statistical significance was estimated using linear mixed effect model. eRFU dynamics with age stratified by the use of immunosuppressant (**e**) or existence of past cancer diagnosis (**f**). Statistical significance was estimated by logistic regression controlled for age. The shaded areas in **a**, **c**, **e** and **f** show 95% confidence intervals. Source data are provided as a Source Data file.

increase after bone marrow transplantation. We observed that even though selected eRFUs may reach a high magnitude in a secondary expansion, the overall count of eRFUs is lower than the primary expansion (Fig. 6e). On average, the sum of eRFUs 5 years after transplantation was 59.6% of that in 5 yr old children.

Discussion

In this study, we identified quantifiable TCR repertoire units, or eRFUs, that robustly correlate with chronologic age, with a bell-shaped association peaking around 27 years old. Systematic investigation over 6500 TCR-seq samples led to several intriguing discoveries: 1) immunosuppressive conditions, such as HIV infection or cancer, were associated with more loss of eRFUs than aging alone, and this trend was partially reversible with certain therapeutic interventions; 2) higher intrinsic eRFU levels were predictive of better outcomes during an acute viral infection in both pediatric and adult patients; 3) HSC transplant from younger donors might lead to a partial restoration of eRFUs in older recipients.

eRFU clones phenotypically belong to MAIT cells that recognize MR1-bound small molecules derived from bacteria[30], which also are activated during viral infections[53,54]. Although the overall percentage of MAIT cells among CD3 + T cells in the blood is reported to dynamically

https://doi.org/10.1038/s41467-024-52522-z

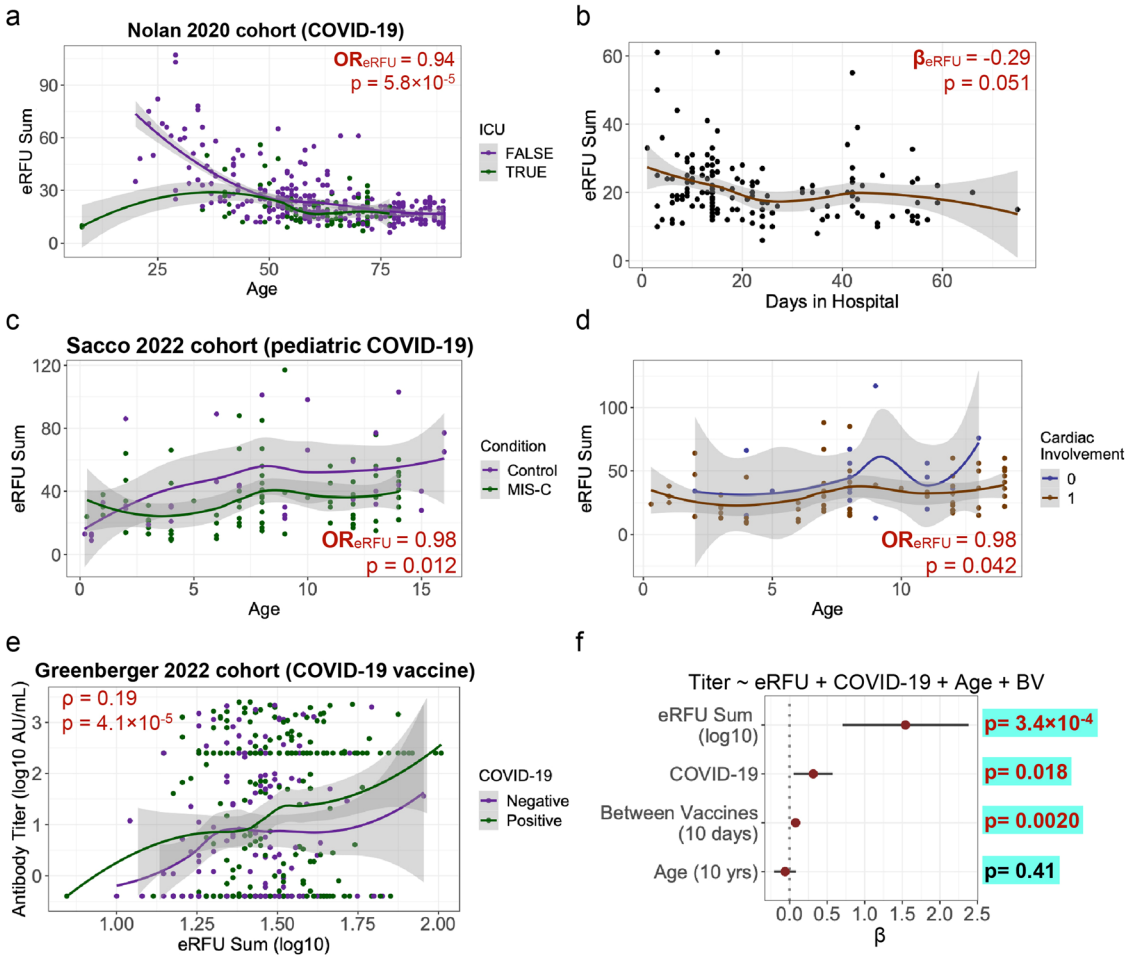


Fig. 5 | Clinical impacts of eRFUs in COVID-19 patients. a Trending of eRFU sum with age stratified by ICU admission status. Odds ratio and *p* value were estimated using logistic regression controlled for patient age. **b** eRFU sum trending with days in hospital. Linear regression controlled for age was used to estimate the impact of eRFU sum and *p* value. eRFU sum over age among pediatric COVID-19 patients stratified by disease groups: healthy control vs MIS-C (**c**) or cardiac involvement status (**d**). Odds ratio and *p* value were estimated using logistic regression controlled for patient age. **e** Scatter plot showing the relationship between eRFU sum and antibody titer after COVID-19 vaccination in an adult cohort. Statistical significance was evaluated using two-sided Spearman's correlation test for both groups combined. The shaded areas in **a**–**e** show 95% confidence intervals. **f** Coefficient plot of linear regression analysis with related clinical covariates in the COVID-19 vaccination cohort. Estimate of coefficients are shown as points, and 95% confidence intervals as bars. Source data are provided as a Source Data file.

change with age[55], there has never been a systematic investigation of which MAIT clones have strong age associations. Our results suggested that only a subset of CD8 + MAIT clones expand after birth, and contract with aging. The age-associations of the 13 eRFUs defined using the Nolan 2020 and Emerson 2017 cohorts are repeatedly seen across all the TCR sample cohorts in our analysis. This is a surprising result, given the high diversity, high plasticity and potentially unseen batch effects of the immune repertoire samples. This reproducibility suggests that there might exist a driving immunologic factor, which could be the exposure to a common pathogen, a stimulating cytokine[56] or an immune-related hormone[57].

A likely explanation to the non-linear age-dynamics of eRFUs is that these T cell clones are expanded with antigen exposure during childhood, while contraction follows thymic involution[58] and HSCs senescence[59] in adulthood. A critical remaining question is, what antigens do these TCRs recognize? One hypothesis is that these RFUs are against common viral or bacterial infections given their constant exposures after birth. Through analyzing sequence logo and protein structure data, we identified 5-OP-RU as one potential bacterial target for eRFU 841. This finding is supportive for the above hypothesis. However, it should be noted that the pool of known antigen targets remains extremely limited, and it is unclear whether eRFUs, which mostly present in the form of CD8 + T cells, will directly kill the target cells to exert their functions.

Within the same age group, the total eRFU counts exhibited nearly 4-fold difference in the healthy donor cohort (Fig. 3e). Genetics might be a contributing factor to this large variation, such as the polymorphism of the TRB variable gene region or the MHC loci[60]. Environmental factors, such as pathogen exposures, use of medications or supplements, lifestyle, disease history, psychosocial stress or other factors, could also be influential to eRFUs, though a systematic investigation is beyond the scope of this study. Notably, many of these factors, particularly diet, distress (e.g., depression), and smoking, among others, have been associated with other age-related phenotypes[61–64]. Leveraging epidemiologic studies will be critical in identifying these relationships.

There are several limitations of this study. First, the eRFUs were identified using cross-sectional data, which limits our ability to assess how changes within individuals over long periods might replicate the patterns observed. To better understand the age-related dynamics of eRFUs, future research should include large longitudinal cohorts with blood specimens collected from multiple timepoints within the same individuals. Second, the eRFU analysis was focused on the βCDR3 region, which cannot fully determine the antigen specificity without

https://doi.org/10.1038/s41467-024-52522-z

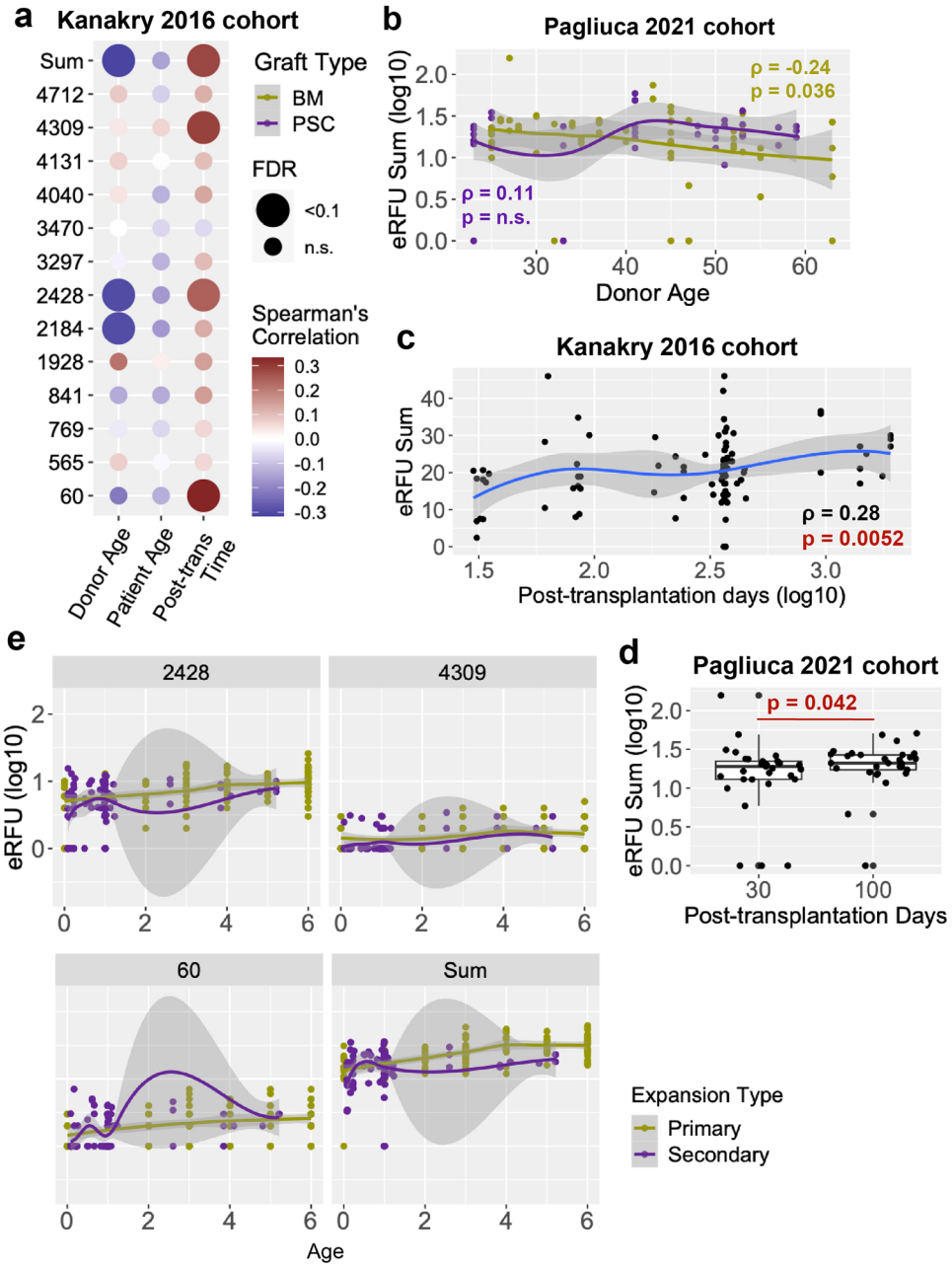


Fig. 6 | Secondary expansion of eRFUs after bone marrow transplantation. **a** Heatmap showing the correlation of eRFU values and age (donors and recipient patients) or post-transplantation time (patients only). Statistical significance was evaluated using two-sided Spearman's correlation test with FDR corrected using the Benjamini-Hochberg approach. Graft types included bone marrow (BM) or peripheral stem cell (PSC). **b** Association of eRFU sum with donor age stratified by graft types. **c** Increase of eRFU sum with post-transplantation days. For both (**c**) and (**d**), statistical significance was evaluated using two-sided Spearman's correlation test. **d** Boxplot showing eRFU increase after transplantation in Pagliuca 2021 cohort. The boxplot display the median, interquartile range (IQR), whiskers extending up to 1.5 times the IQR, and individual data points representing the minimum and maximum values. One-sided Wilcoxon rank sum test was implemented to estimate statistical significance. **e** Comparisons of primary and secondary expansions of selected eRFU clones or the summed value. Age (x-axis) was measured in years. The shaded areas in **b**, **c** and **e** show 95% confidence intervals. Source data are provided as a Source Data file.

the pairing of the α chain. As paired TCR sequences from known antigens become more prevalent, it has been observed that certain antigen specific TCRs predominantly utilize TCRα[65]. Future high-throughput single T cell sequencing efforts to pair with the α chains are needed to understand the functional significance of the eRFUs. Third, while multiple different cohorts were included in our analysis, their racial/ethnic diversity was limited. Also, many of the individuals in the analysis were experiencing diseases, which could influence the results. Finally, our study remains descriptive. The findings in this work are suggestive of selected MAIT clones in healthy immune aging, but in vivo experiments would be required to prove their functional impact during viral or bacterial infections in aged individuals.

In the future, more research is needed to investigate eRFUs and their potential clinical implications. The MR1-restricted eRFU clones are unlikely to be directly specific to SARS-Cov2 antigens, thus, their role in acute infections may transcend the type of pathogens involved, for example, boosting B cell responses through cytokine production[66]. Hence, it will be critical to understanding how the eRFUs are related to other metrics of immune system health and the potential role that they play in driving immunity in the context of infections and cancer.

Additional validation in other populations as well as evaluating relationships with other age-related markers will be important to understand which aspects of the aging phenotype are captured by the eRFU measurement. Finally, our findings might provide novel opportunities to develop approaches for immune rejuvenation, including vaccination approach to achieve higher eRFU levels[67].

Methods

Datasets and preprocessing procedures

All TCR repertoire sequencing samples were accessed from the immuneAccess database (https://clients.adaptivebiotech.com/immuneaccess) managed by Adaptive Biotechnology by Nov 27th, 2023, or downloaded from literature-provided repositories. Samples collected from immuneAccess were profiled using the immunoSEQ platform developed by the company. Zip files were directly downloaded through the 'Export' function and selecting 'v2'. Accession numbers for each cohort are available in Supplementary Table 1. For each repertoire sample, sequences with missing variable genes or nonproductive CDR3 regions were removed. The top 10,000 TCRs with most abundant clonality were selected for RFU calculation. We exclusively used the top 10,000 clones to minimize the batch effects caused by differences in sequencing depths. These preprocessing criteria were applied to all the TCR-seq samples throughout this study. Associated clinical information for each cohort were individually accessed through literature. Single cell RNA sequencing paired with TCR sequencing data were downloaded from GEO by accession numbers GSE178991 and GSE194189.

Repertoire functional units

We applied GIANA[68] to perform clustering of over 20 million TCRs using both the CDR3 sequences and variable gene alleles obtained from public domain. These samples covered a wide spectrum of disease context, including healthy individuals and patients with cancer, autoimmune disorders as well as viral infections. Previous work, including ours, have demonstrated that TCRs clustered using such strategy are highly specific (>=95%) to the same antigen epitopes[69], with smaller ($n <= 5$) clusters being more likely to share antigen-specificity[68].

From GIANA output, we identified a total of 821 K such clusters. An example of a typical cluster of two sequences, CSARQGA***R***TYEQYF and CSARQGA***Y***TYEQYF, bear a mismatch R/Y in position 8. We considered the amino acids flanking this mismatch and extracted the trimer sequences from both TCRs. As the two TCRs likely share antigen-specificity, the two trimers, ART and AYT, are thus considered 'replaceable' in the context of antigen recognition. We then traversed all 821 K clusters and built the 8000-by-8000 trimer-substitution matrix (TSM) by calculating the number of replacements of each trimer pairs. We calculated the Spearman's correlation matrix using TSM and converted it into a Euclidean distance matrix (EDM). Next, similar as in GIANA, we obtained the isometric embedding vector for each of the trimers using multi-dimensional scaling based on the EDM. This approach allowed us to use a numeric vector to represent each trimer, with similar trimers located closely in the Euclidean space. The embedding of each CDR3 sequence is then calculated as the average of all the vectors from consecutive trimers. This embedding is a continuous representation of TCR similarity.

We pooled 1.2 million TCRs from 120 healthy donors from a previous study[23], and projected them onto the Euclidean space with trimer-based embedding. We divided the TCR sequences in this space into 5000 groups with the k-means method. We referred the centroid of each group as a 'Repertoire Functional Unit', or RFU. To calculate the RFU vector of a new TCR repertoire sample, we first select the top 10,000 most abundant TCRs based on clonal frequencies. We consistently used the top 10,000 clones for all the datasets to minimize the influence of sequencing depth. For each TCR, we calculate the embedding vector and assign it to the closest centroid from 5000 RFUs. The value of each RFU is determined by the number of TCRs assigned to its centroid. We chose 5000 as the group number so that the expected count for each RFU is 2.

Regression analysis

The relationship between eRFU sum (log10 scale) and age was quantified with ordinary linear regression for all the cohorts with numeric age information (Fig. 3a–d). Constant variance of residuals (homoscedasticity) was visually confirmed on the residual plot using the Nolan 2020 cohort (Supplementary Fig. 4a). Linear regression with viral infection status and age was performed to evaluate the impact of each viral infection over eRFU sum (Fig. 4b). Linear mixed effect model was implemented with patient age and time post ART therapy as fixed effect, and patient ID as random effect (Fig. 4d). Regression was performed using R package *lme4* (v1.1-33). Statistical significance of ART over eRFU sum was evaluated using the *lmerTest* package (v3.1-3)[70]. Logistic regression controlled for patient age and/or other covariates was implemented to assess the impact of eRFU sum over binary clinical outcomes in COVID-19 patients (Fig. 5). Odds ratio and statistical significance were both estimated from the R function *glm*.

Public TCR analysis

4425 public TCRs were defined as occurring in 90% of an age-representative population[12]. For each public TCR, we assigned its RFU number based on the method described above. To determine if eRFUs were enriched for the public TCRs, we randomly selected 13 RFUs (without replacement) from the 5000 pool and calculated the number of public TCRs that had been assigned with an RFU within the random set. This number could be larger than 13 since it was the TCR count. The overlapping TCR count for eRFU was then compared to 1000 random RFU sets to evaluate statistical significance. RFUs that were assigned with any public TCRs were compared to all or eRFUs for their age associations using the correlation values estimated in the Nolan cohort.

HLA allele association analysis

We applied the Emerson 2017 cohort to investigate RFU association with HLA alleles (A or B loci) at 2-digits resolution. Specifically, we selected alleles occurring in 10 or more individuals in the full cohort. For each allele and for each RFU, we performed a one-sided *t*-test, to investigate if individuals with the allele have higher RFU count than those without, and recorded the z-score. We obtained and visualized the allele-by-RFU matrix of z-scores from this analysis (Supplementary Fig. 2a).

Single cell RNA-seq data analysis for MAIT cells

An analysis of the gene expression pattern of MAIT cells was conducted using the Dong cohort[33] (Access number: GSE178991) containing paired scRNA-seq and scTCR-seq data. We primarily used R/4.1.1 and R Package Seurat/4.0.5 for scRNA-seq data analysis. A total of 75,820 T cells passed quality control, with a mitochondrial gene percentage of less than 5% and a ribosomal gene ratio of less than 2%, allowing them to proceed to downstream analysis. When finding clusters, T cells were divided into 12 subgroups using a 0.3 resolution. T cells with age-associated RFU were highlighted in red in the tSNE plot and were found to be significantly enriched in cluster 7. The expression levels of four MAIT markers[34] (*TRAV1-2*, *CD161*, *ITGAE*, and *SLC4A10*) were evaluated across all clusters. Notably, T cells in cluster 7 displayed typical MAIT gene signatures and were thus designated as MAIT cells.

For the Garner cohort[34] (GSE194189), we directly downloaded the processed data from GEO which had been quality-controlled and normalized. We used Seurat v5.0.3 for single-cell data analysis for this cohort. This dataset contains paired scRNA-seq and scTCR-seq

samples from flow-sorted MAIT cells and conventional memory T cells (Tmem). There were 12 donors in this dataset, aged from 27 to 65, comprising 89,456 cells in total. We further kept cells with only one pair of productive TCRα and TCRβ chains for both cell types. We further discarded some potential contaminating MAIT cells from Tmem cell population, which expressed *TRAV1-2* paired with *TRAJ33*, *TRAJ12* or *TRAJ20* and had a 12 amino acid CDR3α region. Finally, we retained 30,058 MAIT cells and 25,927 Tmem cells. We assigned RFU number to each T cell based on the TCRβ chain and obtained 1971 T cells expressing eRFUs. Enrichment of eRFUs in specific cell subsets/states was performed using Fisher's exact test. Differential analysis was done using the *FindMarker* function in Seurat package. Pseudotime trajectory was constructed using Monocle2[39] v2.30.1, and the direction of pseudotime was determined by CytoTRACE[40] v0.3.3. Visualization of volcano plot was performed using EnhancedVolcano package v1.22.0, and violin plots and bar plots were made with ggpubr v0.6.0, ggstatsplot v0.11.0 and patchwork v1.2.0.

Statistical analysis

Computational and statistical analyses in this work were performed using the R programming language v4.3.0. FDR control was using the Benjamini-Hochberg method. Sequence logos were generated using package *ggseqlogo* (v0.1), by performing multiple sequence alignment (*msa*, v1.32.0) using CDR3s with length 16. Neighbor joining trees were calculated based on pairwise correlation matrix by RFU values and visualized using R package *ape* (v5.7-1). Subpanels of main figures were produced using *ggplot2* (v3.4.2). Loess smooth lines were estimated using the *loess* function in R with default parameters, which was automatically implemented in the *ggplot* function. Beeswarm plots were generated using R package *beeswarm* (v0.4.0). Violin plots were produced using R package *vioplot* (v0.4.0). Due to limited sample size ($n = 10$), Pearson's correlation test was implemented to calculate the eRFU associations between each pair of somatic organs and statistical significance (Fig. 3f). Visualization of the heatmap was performed using *ggplot*.

Reporting summary

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

Data availability

All the TCR repertoire sequencing data can be downloaded at https://clients.adaptivebiotech.com/immuneaccess, and single-cell RNA sequencing and TCR sequencing data can be downloaded at GEO database, with accession links listed in Supplementary Table 1. Source data are provided with this paper.

Code availability

RFU codebase is available at: https://github.com/s175573/RFU.

References

1. Ventura, M. T., Casciaro, M., Gangemi, S. & Buquicchio, R. Immunosenescence in aging: between immune cells depletion and cytokines up-regulation. *Clin. Mol. Allergy* **15**, 21 (2017).
2. Shaw, A. C., Joshi, S., Greenwood, H., Panda, A. & Lord, J. M. Aging of the innate immune system. *Curr. Opin. Immunol.* **22**, 507–513 (2010).
3. Weiskopf, D., Weinberger, B. & Grubeck-Loebenstein, B. The aging of the immune system. *Transpl. Int* **22**, 1041–1050 (2009).
4. Chen, K. & Kolls, J. K. T cell-mediated host immune defenses in the lung. *Annu Rev. Immunol.* **31**, 605–633 (2013).
5. Farhood, B., Najafi, M. & Mortezaee, K. CD8(+) cytotoxic T lymphocytes in cancer immunotherapy: A review. *J. Cell Physiol.* **234**, 8509–8521 (2019).
6. 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).
7. Yano, T. et al. Surgical Outcomes of Postural Instability in Patients With Cervical Myelopathy. *Clin. Spine Surg.* **33**, E466–E471 (2020).
8. Liang, Z., Dong, X., Zhang, Z., Zhang, Q. & Zhao, Y. Age-related thymic involution: Mechanisms and functional impact. *Aging Cell* **21**, e13671 (2022).
9. Britanova, O. V. et al. Age-related decrease in TCR repertoire diversity measured with deep and normalized sequence profiling. *J. Immunol.* **192**, 2689–2698 (2014).
10. Colonna-Romano, G. et al. Impairment of gamma/delta T lymphocytes in elderly: implications for immunosenescence. *Exp. Gerontol.* **39**, 1439–1446 (2004).
11. van der Geest, K. S. M. et al. Impact of Aging on the Frequency, Phenotype, and Function of CD161-Expressing T Cells. *Front Immunol.* **9**, 752 (2018).
12. Britanova, O. V. et al. Dynamics of Individual T Cell Repertoires: From Cord Blood to Centenarians. *J. Immunol.* **196**, 5005–5013 (2016).
13. Qi, Q. et al. Diversity and clonal selection in the human T-cell repertoire. *Proc. Natl. Acad. Sci. USA* **111**, 13139–13144 (2014).
14. Zhuo, Y. et al. Evaluation and comparison of adaptive immunity through analyzing the diversities and clonalities of T-cell receptor repertoires in the peripheral blood. *Front Immunol.* **13**, 916430 (2022).
15. Mittelbrunn, M. & Kroemer, G. Hallmarks of T cell aging. *Nat. Immunol.* **22**, 687–698 (2021).
16. de Greef, P. C. et al. The naive T-cell receptor repertoire has an extremely broad distribution of clone sizes. *Elife* **9,** e49900 (2020).
17. Weng, N. P. Numbers and odds: TCR repertoire size and its age changes impacting on T cell functions. *Semin Immunol.* **69**, 101810 (2023).
18. Yoshida, K. et al. Aging-related changes in human T-cell repertoire over 20years delineated by deep sequencing of peripheral T-cell receptors. *Exp. Gerontol.* **96**, 29–37 (2017).
19. Sun, X. et al. Longitudinal analysis reveals age-related changes in the T cell receptor repertoire of human T cell subsets. *J. Clin. Invest* **132,** e158122 (2022).
20. Laydon, D. J., Bangham, C. R. & Asquith, B. Estimating T-cell repertoire diversity: limitations of classical estimators and a new approach. *Philos Trans. R Soc. Lond B Biol. Sci.* **370,** (2015).
21. Yu, X. et al. Quantifiable TCR repertoire changes in prediagnostic blood specimens among patients with high-grade ovarian cancer. *Cell Rep. Med.* **5**, 101612 (2024).
22. Nolan, S. et al. A large-scale database of T-cell receptor beta (TCRbeta) sequences and binding associations from natural and synthetic exposure to SARS-CoV-2. *Res Sq*, Preprint (2020).
23. Emerson, R. O. et al. Immunosequencing identifies signatures of cytomegalovirus exposure history and HLA-mediated effects on the T cell repertoire. *Nat. Genet* **49**, 659–665 (2017).
24. Mitchell, A. M. et al. Temporal development of T cell receptor repertoires during childhood in health and disease. *JCI Insight* **7** (2022).
25. Madi, A. et al. T cell receptor repertoires of mice and humans are clustered in similarity networks around conserved public CDR3 sequences. *Elife* **6,** (2017).
26. Fischer, S., Stanke, F. & Tummler, B. VJ Segment Usage of TCR-Beta Repertoire in Monozygotic Cystic Fibrosis Twins. *Front Immunol.* **12**, 599133 (2021).
27. Cader, F. Z. et al. A peripheral immune signature of responsiveness to PD-1 blockade in patients with classical Hodgkin lymphoma. *Nat. Med.* **26**, 1468–1479 (2020).
28. Emerson, R. et al. Estimating the ratio of CD4+ to CD8+ T cells using high-throughput sequence data. *J. Immunol. Methods* **391**, 14–21 (2013).

29. DeWitt, W. S. et al. Human T cell receptor occurrence patterns encode immune history, genetic background, and receptor specificity. *Elife* **7** (2018).
30. Le Bourhis, L. et al. Mucosal-associated invariant T cells: unconventional development and function. *Trends Immunol.* **32**, 212–218 (2011).
31. Matsuoka, T. et al. The effects of 5-OP-RU stereochemistry on its stability and MAIT-MR1 axis. *Chembiochem* **22**, 672–678 (2021).
32. Xiao, X. & Cai, J. Mucosal-Associated Invariant T Cells: New Insights into Antigen Recognition and Activation. *Front Immunol.* **8**, 1540 (2017).
33. Dong, S. et al. The effect of low-dose IL-2 and Treg adoptive cell therapy in patients with type 1 diabetes. *JCI Insight* **6,** e147474 (2021).
34. Garner, L. C. et al. Single-cell analysis of human MAIT cell transcriptional, functional and clonal diversity. *Nat. Immunol.* **24**, 1565–1578 (2023).
35. Pais Ferreira, D. et al. Central memory CD8(+) T cells derive from stem-like Tcf7(hi) effector cells in the absence of cytotoxic differentiation. *Immunity* **53**, 985–1000 e11 (2020).
36. Yao, C. et al. BACH2 enforces the transcriptional and epigenetic programs of stem-like CD8(+) T cells. *Nat. Immunol.* **22**, 370–380 (2021).
37. Vodnala, S. K. et al. T cell stemness and dysfunction in tumors are triggered by a common mechanism. *Science* **363** (2019).
38. Gattinoni, L. et al. A human memory T cell subset with stem cell-like properties. *Nat. Med.* **17**, 1290–1297 (2011).
39. Qiu, X. et al. Single-cell mRNA quantification and differential analysis with Census. *Nat. Methods* **14**, 309–315 (2017).
40. Gulati, G. S. et al. Single-cell transcriptional diversity is a hallmark of developmental potential. *Science* **367**, 405–411 (2020).
41. Van de Sande, B. et al. A scalable SCENIC workflow for single-cell gene regulatory network analysis. *Nat. Protoc.* **15**, 2247–2276 (2020).
42. Greenberger, L. M. et al. Anti-spike T-cell and Antibody Responses to SARS-CoV-2 mRNA Vaccines in Patients with Hematologic Malignancies. *Blood Cancer Discov.* **3**, 481–489 (2022).
43. Gittelman, R. M. et al. Longitudinal analysis of T cell receptor repertoires reveals shared patterns of antigen-specific response to SARS-CoV-2 infection. *JCI Insight* **7**, e151849 (2022).
44. DeWolf, S. et al. Tissue-specific features of the T cell repertoire after allogeneic hematopoietic cell transplantation in human and mouse. *Sci. Transl. Med.* **15**, eabq0476 (2023).
45. Towlerton, A. M. H., Ravishankar, S., Coffey, D. G., Puronen, C. E. & Warren, E. H. Serial Analysis of the T-Cell Receptor beta-Chain Repertoire in People Living With HIV Reveals Incomplete Recovery After Long-Term Antiretroviral Therapy. *Front Immunol.* **13**, 879190 (2022).
46. Kearns, N., Maijers, I., Harper, J., Beasley, R. & Weatherall, M. Inhaled Corticosteroids in Acute Asthma: A Systemic Review and Meta-Analysis. *J. Allergy Clin. Immunol. Pr.* **8**, 605–617 e6 (2020).
47. Jiang, Y., Li, Y. & Zhu, B. T-cell exhaustion in the tumor microenvironment. *Cell Death Dis.* **6**, e1792 (2015).
48. Chabner, B. A. & Roberts, T. G. Jr Timeline: Chemotherapy and the war on cancer. *Nat. Rev. Cancer* **5**, 65–72 (2005).
49. Sacco, K. et al. Immunopathological signatures in multisystem inflammatory syndrome in children and pediatric COVID-19. *Nat. Med.* **28**, 1050–1062 (2022).
50. Kanakry, C. G. et al. Origin and evolution of the T cell repertoire after posttransplantation cyclophosphamide. *JCI Insight* **1**, e86252 (2016).
51. Pagliuca, S. et al. The similarity of class II HLA genotypes defines patterns of autoreactivity in idiopathic bone marrow failure disorders. *Blood* **138**, 2781–2798 (2021).
52. Hellmich, C., Moore, J. A., Bowles, K. M. & Rushworth, S. A. Bone Marrow Senescence and the Microenvironment of Hematological Malignancies. *Front Oncol.* **10**, 230 (2020).
53. Sandberg, J. K., Leeansyah, E., Eller, M. A., Shacklett, B. L. & Paquin-Proulx, D. The Emerging Role of MAIT Cell Responses in Viral Infections. *J. Immunol.* **211**, 511–517 (2023).
54. van Wilgenburg, B. et al. MAIT cells are activated during human viral infections. *Nat. Commun.* **7**, 11653 (2016).
55. Walker, L. J., Tharmalingam, H. & Klenerman, P. The rise and fall of MAIT cells with age. *Scand. J. Immunol.* **80**, 462–463 (2014).
56. Junnila, R. K., List, E. O., Berryman, D. E., Murrey, J. W. & Kopchick, J. J. The GH/IGF-1 axis in ageing and longevity. *Nat. Rev. Endocrinol.* **9**, 366–376 (2013).
57. Arlt, W. Dehydroepiandrosterone and ageing. *Best. Pr. Res Clin. Endocrinol. Metab.* **18**, 363–380 (2004).
58. Thomas, R., Wang, W. & Su, D. M. Contributions of Age-Related Thymic Involution to Immunosenescence and Inflammaging. *Immun. Ageing* **17**, 2 (2020).
59. Ahmed, A. S., Sheng, M. H., Wasnik, S., Baylink, D. J. & Lau, K. W. Effect of aging on stem cells. *World J. Exp. Med.* **7**, 1–10 (2017).
60. Sharon, E. et al. Genetic variation in MHC proteins is associated with T cell receptor expression biases. *Nat. Genet* **48**, 995–1002 (2016).
61. Freitas-Simoes, T. M., Ros, E. & Sala-Vila, A. Nutrients, foods, dietary patterns and telomere length: Update of epidemiological studies and randomized trials. *Metabolism* **65**, 406–415 (2016).
62. Oblak, L., van der Zaag, J., Higgins-Chen, A. T., Levine, M. E. & Boks, M. P. A systematic review of biological, social and environmental factors associated with epigenetic clock acceleration. *Ageing Res Rev.* **69**, 101348 (2021).
63. Ridout, K. K. et al. Early life adversity and telomere length: a meta-analysis. *Mol. Psychiatry* **23**, 858–871 (2018).
64. Ryan, J., Wrigglesworth, J., Loong, J., Fransquet, P. D. & Woods, R. L. A Systematic Review and Meta-analysis of Environmental, Lifestyle, and Health Factors Associated With DNA Methylation Age. *J. Gerontol. A Biol. Sci. Med Sci.* **75**, 481–494 (2020).
65. Choy, C. et al. SARS-CoV-2 infection establishes a stable and age-independent CD8(+) T cell response against a dominant nucleocapsid epitope using restricted T cell receptors. *Nat. Commun.* **14**, 6725 (2023).
66. Boulouis, C. et al. MAIT cell compartment characteristics are associated with the immune response magnitude to the BNT162b2 mRNA anti-SARS-CoV-2 vaccine. *Mol. Med.* **28**, 54 (2022).
67. Wu, R., Sun, F., Zhang, W., Ren, J. & Liu, G. H. Targeting aging and age-related diseases with vaccines. *Nat. Aging* **4**, 464–482 (2024).
68. Zhang, H., Zhan, X. & Li, B. GIANA allows computationally-efficient TCR clustering and multi-disease repertoire classification by isometric transformation. *Nat. Commun.* **12**, 4699 (2021).
69. Dash, P. et al. Quantifiable predictive features define epitope-specific T cell receptor repertoires. *Nature* **547**, 89–93 (2017).
70. Kuznetsova, A., Brochkoff, P. B. & Christensen, R. H. B. lmerTest Package: Tests in Linear Mixed Effects Models. *J. Statistical Software* **82**, 1–26 (2017).

Acknowledgements

This work is supported by NCI R01 grants CA258524 (B.L.) and CA245318 (B.L.).

Author contributions

B.L. conceived the project. B.L. and J.H. performed the analysis and wrote the manuscript. S.T., M.P. and B.R. helped with data analysis.

Competing interests

The authors declare no competing interest.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-024-52522-z.

Correspondence and requests for materials should be addressed to Bo Li.

Peer review information *Nature Communications* thanks Jorg Goronzy, Nan-ping Weng and the other anonymous reviewer(s) for their contribution to the peer review of this work. A peer review file is available.