Quantifiable Blood TCR Repertoire Components Associate with Immune Aging

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

1: Department of Pathology and Laboratory Medicine, Perelman School of Medicine, University of Pennsylvania, PA

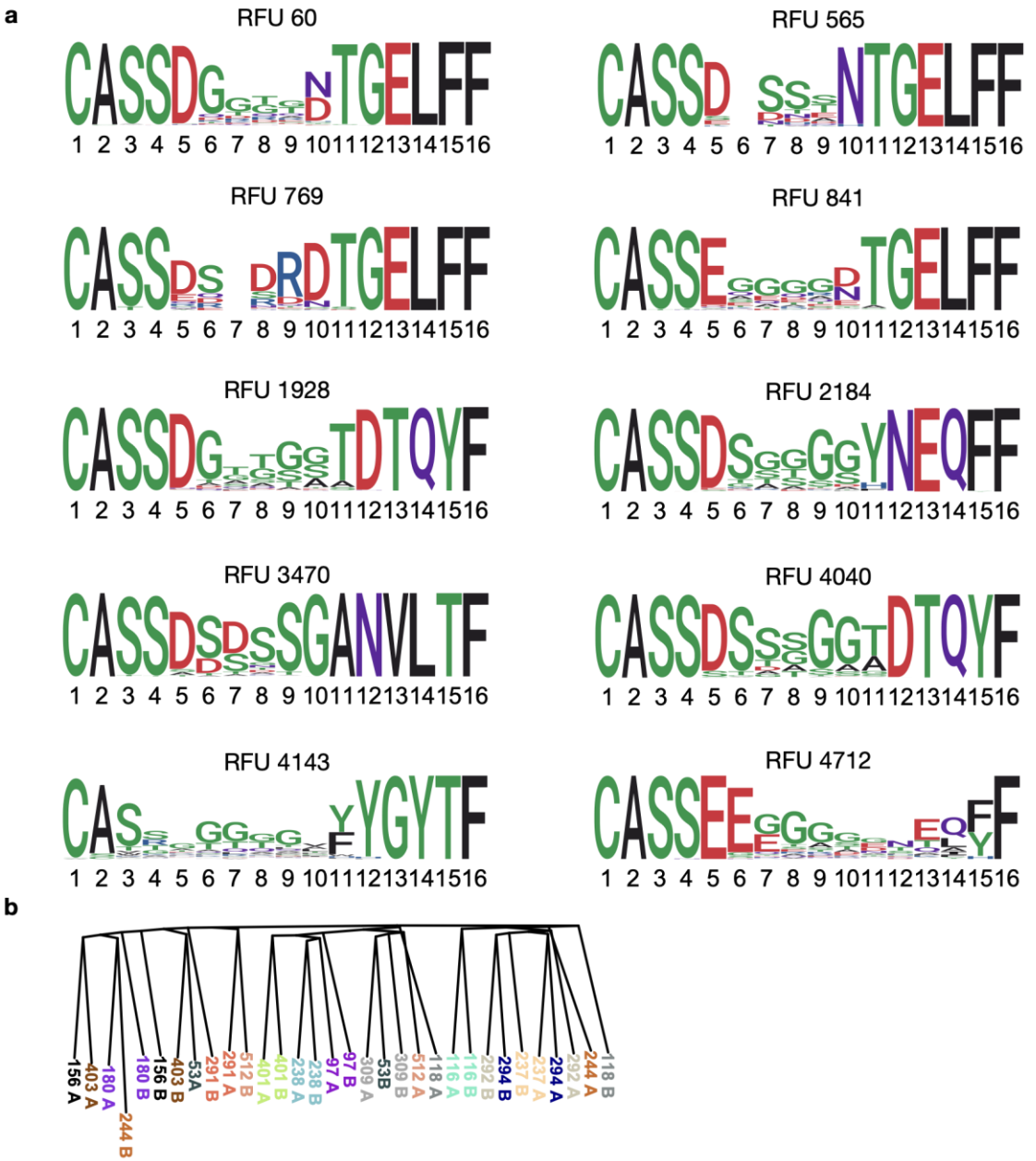
2: Department of Pathology and Laboratory Medicine, Children's Hospital of Philadelphia, PA

3: Department of Cancer Epidemiology, Moffitt Cancer Center, Tampa, FL

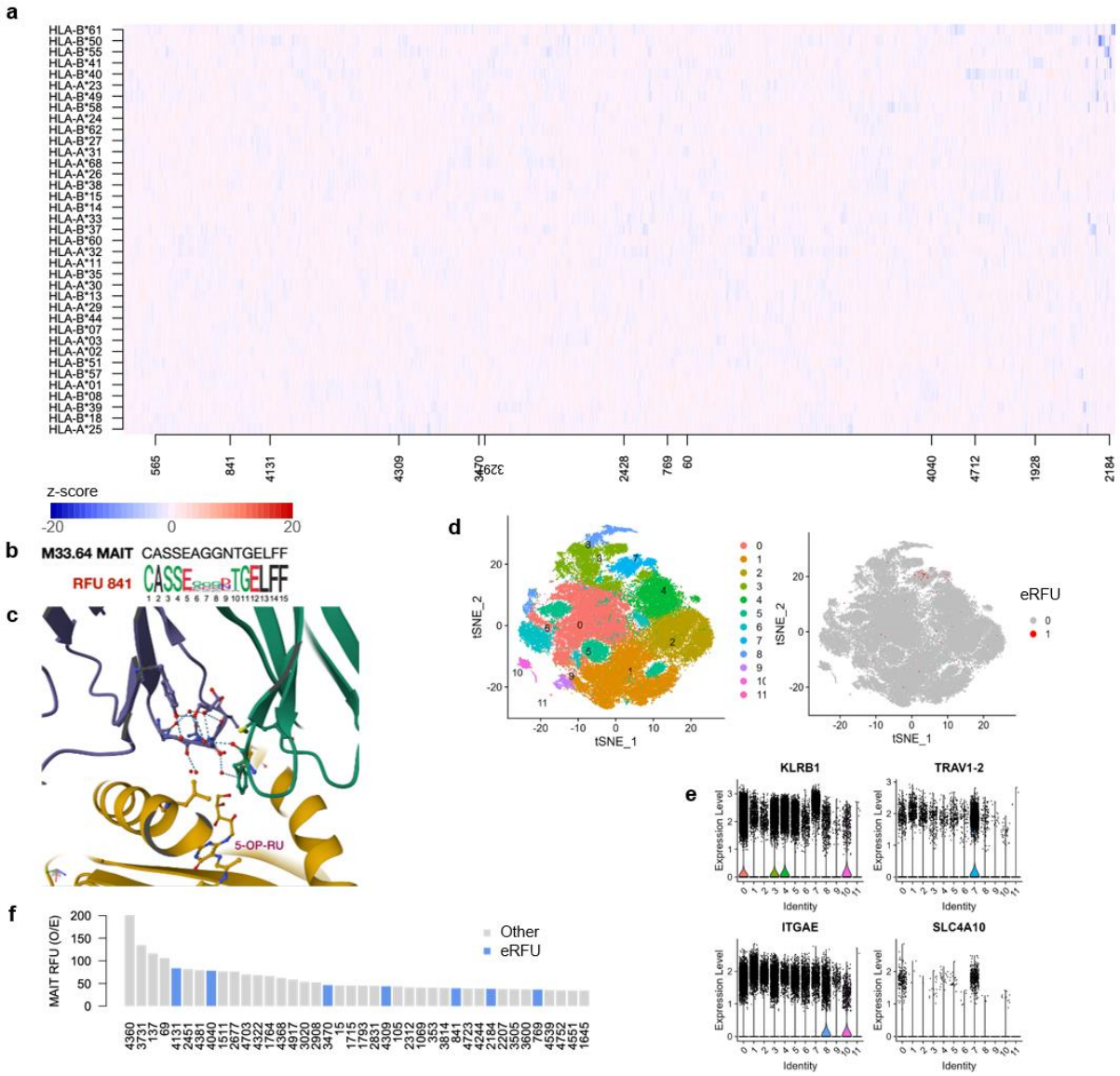
4: Knight Cancer Institute and Division of Oncological Sciences, Oregon Health and Science University, Portland, OR

*: corresponding author: Bo Li (lib3@chop.edu)

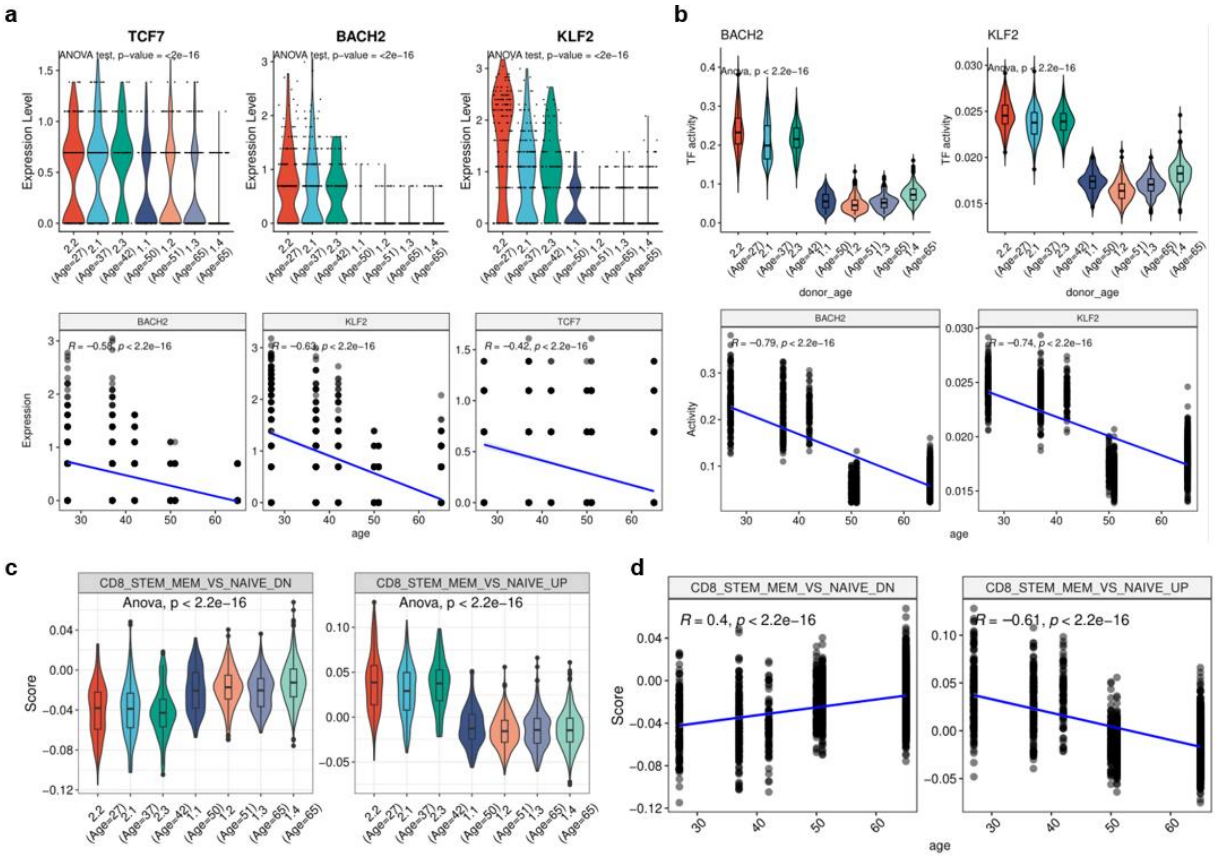
Supplementary Information Content:



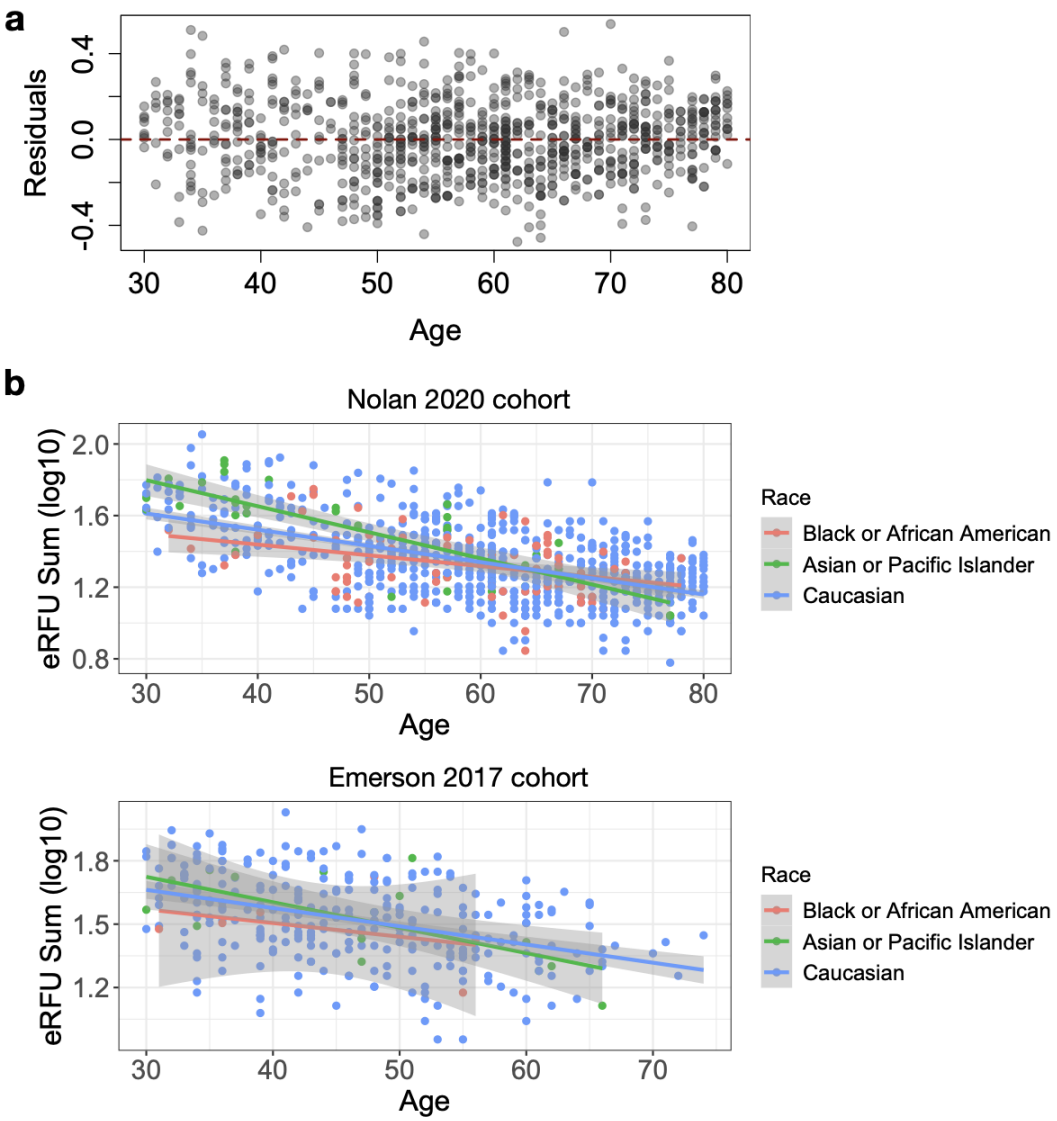
Supplementary Fig. 1: Additional features of eRFUs. **a**) Sequence logo plot for 10 eRFUs not shown in Figure 1. **b**) Neighbor joining tree showing the relationship between all the twins as in Figure 1h using all RFUs to calculate the distance matrix. Source data are provided as a Source Data file.



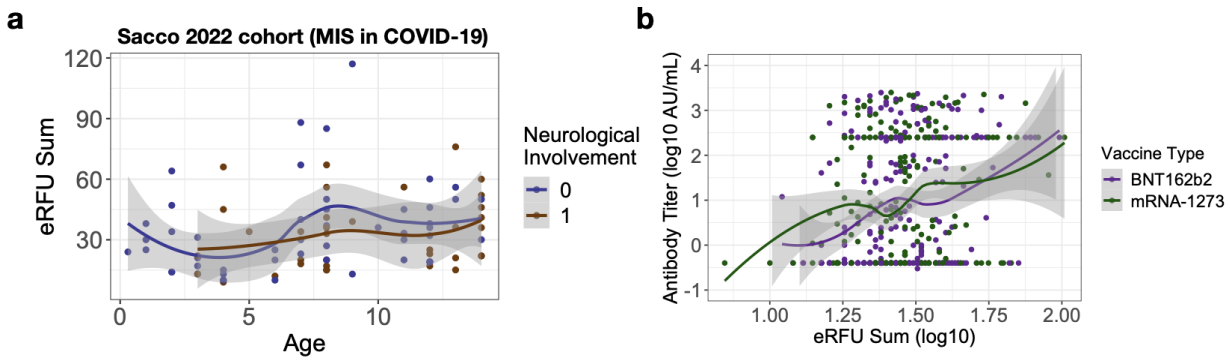
Supplementary Fig. 2: eRFUs displayed mucosal-associated invariant T cell signatures. **a**) Heatmap showing the z-scores of 5,000 RFUs and 37 HLA alleles. eRFU numbers are marked on x-axis. Details see Methods. **b**) Sequence motif comparison of the M33.64 TCR with RFU 841. **c**) Zoom-in view of the TCRβ chain (purple) of M33.64 interacting with MR1 (dijon yellow) and the antigen molecule (PDB accession: 5D7J). **d**) tSNE plot and clustering of T cells using single cell RNA-seq data (left) and same plot highlighting eRFU T cells with red color (right). **e**) Gene signature plot showing overexpression of selected MAIT cell markers in cluster 7. **f**) Enrichment analysis of MAIT RFUs among the T cells in the same single cell dataset. Source data are provided as a Source Data file.



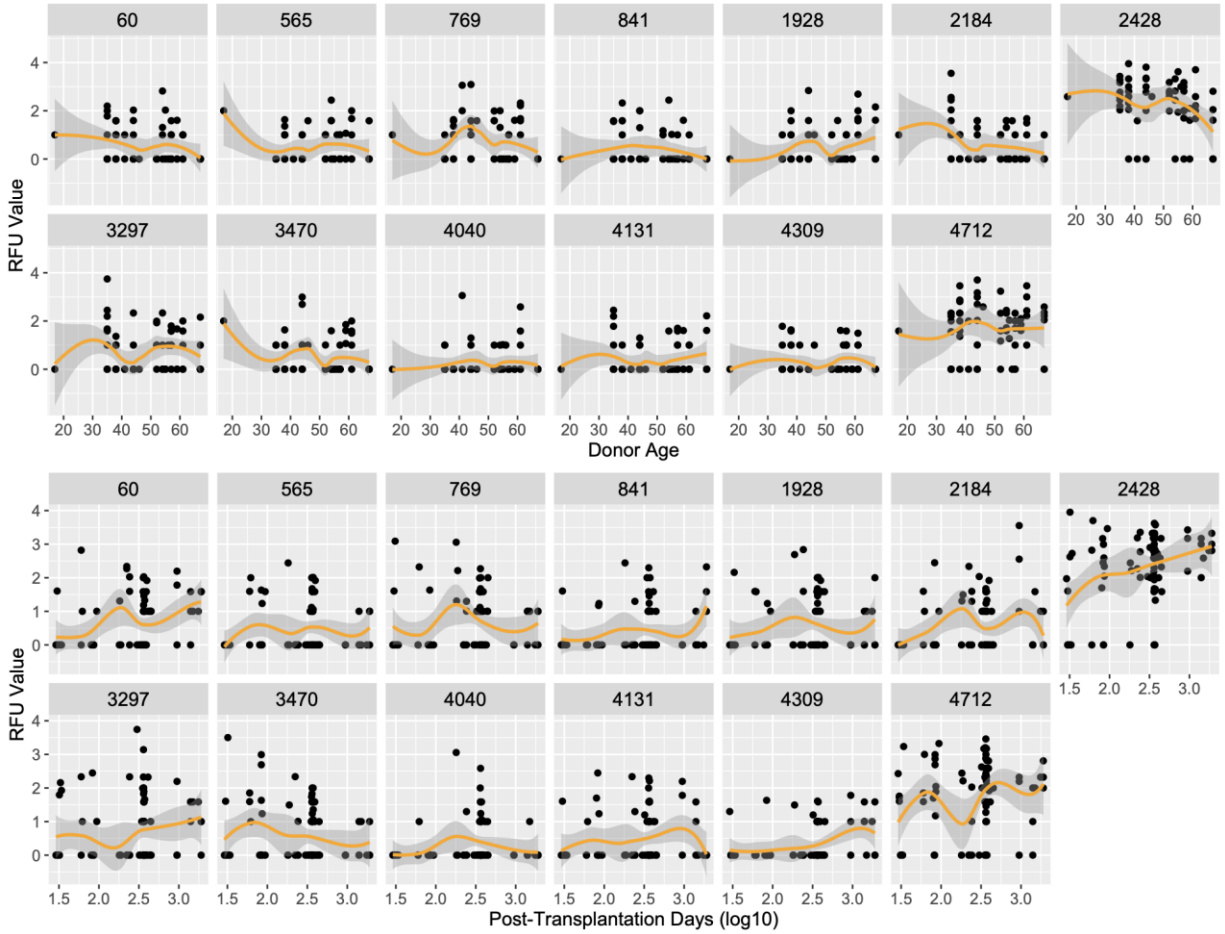
Supplementary Fig. 3: Age association of stemness markers in the eRFU cells. **a**) Violin (upper) or scatter (lower) plots showing the distributions of selected T cell stemness genes trending with age. Statistical significance in the upper panels was evaluated using one-way ANOVA, while the lower panels using Pearson's correlation test. **b**) Similar analysis as in A) performed for transcription factor activity for stemness markers BACH2 and KLF2. **c-d**) Similar analysis performed for the scores measuring the signature genes down- or up- regulated in CD8 stem memory T cells compared to naïve T cells. Source data are provided as a Source Data file.



Supplementary Fig. 4: eRFU residual plot and distribution across different racial groups. a) Regression residual plot against age. Linear regression model log10(eRFU sum)~Age was estimated using Nolan 2020 cohort. **b**) Trends of eRFU sum with age stratified by racial groups. Source data are provided as a Source Data file.



Supplementary Fig. 5: Additional results on eRFU functional impact. **a**) Scatter plot showing eRFU sum with age in pediatric COVID-19 cohort stratified with whether the patient developed neurological complications or not. **b**) Antibody titer increased with eRFU sum stratified by vaccine types. Source data are provided as a Source Data file.



Supplementary Fig. 6: eRFU dynamics with donor age (upper panel) or with post-transplantation days (lower panel). Source data are provided as a Source Data file.

Supplementary Table 1: Data sets used in this study.

Study	Sample Size*	Disease	PubMed ID	Link
Emerson et al., 2013	50	Healthy Donor, Multiple Sclerosis	23428915	https://clients.adaptivebiotech.com/pub/emerson-2013-jim
Nolan et al., 2020	1,414	Healthy Donor (COVID-19)	32793896	https://clients.adaptivebiotech.com/pub/covid-2020
Pagliuca et al., 2021	135	Leukemia (BM transplant)	34748628	https://clients.adaptivebiotech.com/pub/pagliuca-2021-b
Kanakry et al., 2016	104	Healthy Donor and GVHD patients	27213183	https://clients.adaptivebiotech.com/pub/Kanakry-2016-JCIInsight
Emerson et al., 2017	666	Healthy Donors	28369038	https://clients.adaptivebiotech.com/pub/emerson-2017-natgen
Britanova et al., 2016	79	Healthy Donors	27183615	http://mitcr.milaboratory.com/datasets/
Sacco et al., 2022	131	Pediatric MIS-C (COVID-19)	35177862	https://clients.adaptivebiotech.com/pub/sacco-2021-misc
Towlerton et al., 2022	192	HIV patients	35585986	https://clients.adaptivebiotech.com/pub/towlerton-2022-hiv
Gittelman et al., 2022	2,887	Healthy Donor (COVID-19)	35439174	https://clients.adaptivebiotech.com/pub/gittelman-2022-jci
Greenberger et al., 2022	505	Leukemia (COVID-19)	36074641	https://clients.adaptivebiotech.com/pub/greenberger-2022-bcd
Cader et al., 2020	102	Hochkin lymphoma	32778827	https://clients.adaptivebiotech.com/pub/cader-2020-nm
DeWolf et al., 2022	108	Post-mortem	37494469	https://clients.adaptivebiotech.com/pub/dewolf-2022-gvhd
Fischer et al., 2021	65	Twins with CF	33708199	https://clients.adaptivebiotech.com/pub/fischer-2021-fi
Mitchell et al., 2022	359	Healthy Donor (Pediatric)	35998036	https://clients.adaptivebiotech.com/pub/mitchell-2022-jcii
Dong et al., 2021	18 donors, 75,820 cells	Type 1 Diabetes	34324441	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE178991
Garner et al., 2023	12 donors, 89,456 cells	Mixed conditions	37580605	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE194189

*: Actual sample size of the original cohort might be larger. The number in this table reflect the sample size used in this study. Preprocessing and sample filtering criteria are described in Methods.