DISSERTATIONES
MEDICINAE
UNIVERSITATIS
TARTUENSIS
377

IHOR FILIPPOV

Single-cell data analysis in immunology: from technology to applications

DISSERTATIONES MEDICINAE UNIVERSITATIS TARTUENSIS
377

DISSERTATIONES MEDICINAE UNIVERSITATIS TARTUENSIS
377

IHOR FILIPPOV

Single-cell data analysis in immunology: from technology to applications

UNIVERSITY OF TARTU
Press
1632

Institute of Biomedicine and Translational Medicine, University of Tartu, Tartu, Estonia

This dissertation has been accepted for the commencement of the degree of Doctor of Philosophy in Medicine on March 19th, 2025, by the Council of the Faculty of Medicine, University of Tartu, Estonia

Supervisor: Professor Pärt Peterson, PhD
Institute of Biomedicine and Translational Medicine,
Faculty of Medicine
University of Tartu, Tartu, Estonia

Reviewers: Associate Professor Alar Aints, PhD
Institute of Biomedicine and Translational Medicine,
Faculty of Medicine
University of Tartu, Tartu, Estonia

Associate Professor Kaur Alasoo, PhD
Institute of Computer Science,
Faculty of Science and Technology
University of Tartu, Tartu, Estonia

Opponent: Associate Professor Nuno Luís Barbosa Morais, PhD
Faculty of Medicine
University of Lisbon, Lisbon, Portugal

Commencement: May 28th, 2025

This research was supported by:

1. European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No.: 955321
2. Estonian Research Council grants PRG377, PRG1177 and PRG2011

ISSN 1024-395X (print)
ISBN 978-9916-27-856-7 (print)
ISSN 2806-240X (pdf)
ISBN 978-9916-27-857-4 (pdf)

University of Tartu Press
www.tyk.ee

TABLE OF CONTENTS

LIST OF ORIGINAL PUBLICATIONS

I **Filippov, I.**, Philip, C. S., Schauser, L., & Peterson, P. (2024). Comparative transcriptomic analyses of thymocytes using 10x Genomics and Parse scRNA-seq technologies. *BMC Genomics*, *25*(1), 1069. https://doi.org/10.1186/s12864-024-10976-x

II **Filippov, I.**, Schauser, L., & Peterson, P. (2024). An integrated single-cell atlas of blood immune cells in aging. *Npj Aging*, 10(1), 59. https://doi.org/10.1038/s41514-024-00185-x

III Türk, L., **Filippov, I.**, Arnold, C., Zaugg, J., Tserel, L., Kisand, K., & Peterson, P. (2024). Cytotoxic CD8+ Temra cells show loss of chromatin accessibility at genes associated with T cell activation. *Frontiers in Immunology*, *15*, 1285798. https://doi.org/10.3389/fimmu.2024.1285798

IV Sjøgren, T., Islam, S., **Filippov, I.**, Jebrzycka, A., Sulen, A., Breivik, L. E., Hellesen, A., Jørgensen, A. P., Lima, K., Tserel, L., Kisand, K., Peterson, P., Ranki, A., Husebye, E. S., Oftedal, B. E., & Wolff, A. S. B. (2024). Single cell characterization of blood and expanded regulatory T cells in autoimmune polyendocrine syndrome type 1. *iScience*, *27*(4), 109610. https://doi.org/10.1016/j.isci.2024.109610

The author's personal contribution:

Paper I: Study conceptualization and design, data analysis, writing the manuscript.

Paper II: Study conceptualization and design, data collection and analysis, writing the manuscript.

Paper III: Analyzing the single-cell data and writing the manuscript.

Paper IV: Analyzing the single-cell data of fresh cells and writing the manuscript.

ABBREVIATIONS

aAbs	Autoantibodies
ABC	Aging-associated (atypical) B cells
APC	Antigen-presenting cells
APECED	Autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy
APS-1	Autoimmune polyendocrine syndrome type 1
CCA	Canonical correlation analysis
cDNA	Complimentary DNA
CMV	Cytomegalovirus
cTEC	Cortical thymic epithelial cells
CTL	Cytotoxic lymphocytes
DCs	Dendritic cells
DN	Double-negative thymocytes
DP	Double-positive thymocytes
DP-A	Double-positive thymocytes undergoing apoptosis
DP-P	Double-positive thymocytes undergoing undergoing proliferation
DP-R	Double-positive thymocytes undergoing undergoing recombination
DP-S	Double-positive thymocytes undergoing selection
FACS	Fluorescence-activated cell sorting
GEM	Gel bead-in-emulsions
GEO	Gene Expression Omnibus
GPU	Graphics processing unit
HSCs	Hematopoietic stem cells
HTO	Hashtag oligo
HVGs	Highly variable genes
IFN-α	Interferon alpha
IFN-β	Interferon beta
IFN-γ	Interferon gamma
ILs	Interleukins
MAIT	Mucosal-associated invariant T cells
MHC-II	Major histocompatibility complex class II
MNN	Mutual nearest neighbours
mTEC	Medullary thymic epithelial cell
NKs	Natural killer cells
NKT	Natural killer T cells
pDCs	Plasmacytoid dendritic cells
qPCR	Quantitative polymerase chain reaction
RNA	Ribonucleic acid
ROS	Reactive oxygen species
scRNA-seq	Single-cell RNA sequencing

SP	Single-positive thymocytes
t-SNE	t-distributed stochastic neighbor embedding
Tact	Activated T cells
Tcm	Central memory T cells
TCR	T cell receptor
Tem	Effector memory T cells
Temra	Effector memory T cell re-expressing CD45RA
Tgd	Gamma-delta T cells
Th1	T helper 1 cells
Th17	T helper 17 cells
Th2	T helper 2 cells
Tn	Naïve T cell
Treg	Regulatory T cell
UMAP	Uniform manifold approximation and projection
UMI	Unique molecular identifier

1. INTRODUCTION

Single-cell technology has revolutionized biomedical research by allowing scientists to study molecular profiles of individual cells, providing unprecedented insight into cellular diversity and function. This technology enables the identification of rare cell types, tracks dynamic cellular states, and uncovers subtle changes in gene expression that are masked in bulk analyses (Davidson et al., 2020; Miao et al., 2020; K. Sun et al., 2022). However, the complexity and volume of data generated by single-cell sequencing require sophisticated bioinformatic analysis to interpret results accurately. Advanced data analysis tools are essential to distill meaningful patterns from high-dimensional data, identify disease-specific biomarkers, and gain a deeper understanding of developmental and disease processes at the cellular level. Together, single-cell technology and robust data analysis are critical for driving forward precision medicine and personalized therapeutic strategies.

Aging is a complex and multifaceted phenomenon that is associated with an increased susceptibility to diseases over time and a gradual deterioration of physiological functions. Despite extensive research and a wealth of data from different experimental systems, the nature of aging is still debated (De Magalhães, 2023; Hayflick, 2007; Timmons & Brenner, 2024; J.-H. Yang et al., 2023). The interplay between genetic factors and environmental influences complicates our understanding, as does the variability in aging patterns among individuals. The role of external stimuli, such as chronic infections, on aging must be clarified. Aging profoundly affects the immune system, leading to immunosenescence, which reduces the body's ability to fight infections and contributes to the increased prevalence of autoimmune diseases and chronic inflammation in older adults (Montecino-Rodriguez et al., 2013).

Autoimmune Polyendocrine Syndrome Type 1 (APS-1) is a rare, monogenic disorder caused by mutations in the *AIRE* gene, leading to immune system dysregulation and multiple endocrine deficiencies (Peterson & Peltonen, 2005). Unlike aging, which is a multifactorial process, the basis of APS-1 is well-defined. While APS-1 can be traced to a single genetic anomaly, aging involves a complex interplay of genetic, environmental, and lifestyle factors. However, the seemingly "simple" cause of APS-1 should not obscure the fact that a single mutation can lead to a wide range of downstream effects (Kisand & Peterson, 2011)

The goal of this thesis was to explore applications of single-cell RNA sequencing (scRNA-seq) in immunological research, focusing on transcriptomic signatures of the aging immune system and the dysregulated transcriptome associated with APS-1. A central objective was to demonstrate the utility of scRNA-seq for uncovering cell states and transcriptional changes across different contexts. In Paper I, a comparative transcriptomic analysis of thymocytes was conducted using 10x Genomics and Parse Bio scRNA-seq technologies. Paper II presented an integrated single-cell atlas of blood immune cells in aging, identifying transcriptomic biomarkers across age groups. Paper III analyzed $CD8^+$

Temra cells in young individuals, accounting for their cytomegalovirus (CMV) seropositivity status. Finally, Paper IV examined the heterogeneity of immune cells in APS-1, highlighting differences from healthy controls.

2 REVIEW OF THE LITERATURE

2.1 Single-cell transcriptomics

2.1.1 The rise of single-cell genomics

The scRNA-seq represents a significant advancement in the field of transcriptomics, allowing for the high-resolution analysis of gene expression at the individual cell level. Unlike traditional bulk RNA sequencing, which provides an averaged gene expression profile from a mixed population of cells, scRNA-seq allows for the dissection of cellular heterogeneity by capturing the transcriptomic signatures of single cells. This high-resolution approach is crucial for understanding the complexity of biological systems, as it reveals the diverse cell types, states, and functional roles often masked in bulk analyses. The scRNA-seq provides whole-transcriptome profiling, offering a comprehensive view of gene expression in individual cells. In contrast, flow cytometry relies on a predefined antibody panel, limiting its analysis to known cell surface markers. This flexibility enables scRNA-seq to discover novel cell types and states that might be missed by flow cytometry, offering more profound insights into cellular diversity and function.

Single-cell gene expression profiling emerged in the early 1990s, with pioneering efforts focused on manually isolating individual cells, followed by qPCR (Eberwine et al., 1992; Lambolez et al., 1992). The first protocol for single-cell isolation followed by sequencing was demonstrated by Tang and colleagues in 2009 (Tang et al., 2009). These initial methods were labor-intensive, requiring meticulous manual handling, and had limited throughput, making it challenging to analyze a large number of cells efficiently. Since then, the landscape of scRNA-seq has been dramatically transformed with the advent of high-throughput techniques such as droplet-based microfluidics, microwell-based platforms, and combinatorial indexing (Svensson et al., 2018). These technological advancements have significantly increased the scalability, accuracy, and cost-effectiveness of scRNA-seq, enabling the analysis of thousands to millions of cells in a single experiment.

The scRNA-seq has profoundly impacted multiple areas of biological research by providing unprecedented insights into cellular diversity and function. In immunology, scRNA-seq has been instrumental in elucidating the heterogeneity of immune cell populations, revealing novel cell types and states, and mapping the dynamic responses of immune cells to infection and vaccines (Stubbington et al., 2017). In aging research, scRNA-seq has uncovered molecular signatures of aging at the cellular level, identified biomarkers, and provided insights into age-related decline in immune function (Mogilenko et al., 2022). Furthermore, in the context of autoimmunity, scRNA-seq has shed light on the mechanisms of immune dysregulation, identified pathogenic cell populations, and highlighted potential therapeutic targets by analyzing the molecular pathways involved in autoimmune diseases (L. Zeng et al., 2022). By capturing the transcriptomes of individual cells, scRNA-seq enables researchers to uncover new cell types, trace

lineage relationships, and explore dynamic processes such as differentiation, immune responses, and disease progression at an unparalleled resolution.

2.1.2 Major single-cell sequencing steps

The scRNA-seq involves a series of elaborate steps, each profoundly impacting the accuracy and depth of the resulting data. The process initiates with isolating single cells, which can be achieved using multiple different techniques such as fluorescence-activated cell sorting (FACS), microfluidic devices, and plate-based isolation (Hwang et al., 2018). FACS is frequently utilized for its ability to sort cells based on specific surface markers, enabling targeted cell population studies. Microfluidic devices, such as those using droplet encapsulation, offer high-throughput capabilities by isolating individual cells into nanoliter-scale droplets, each containing the necessary reagents for subsequent RNA processing (A. M. Klein et al., 2015; Macosko et al., 2015). These droplets serve as reaction chambers where cell lysis, reverse transcription of RNA to cDNA, and barcoding occur sequentially within each droplet (Figure 1). 10x Genomics single-cell platform is the most popular droplet-based commercial solution (G. X. Y. Zheng et al., 2017).

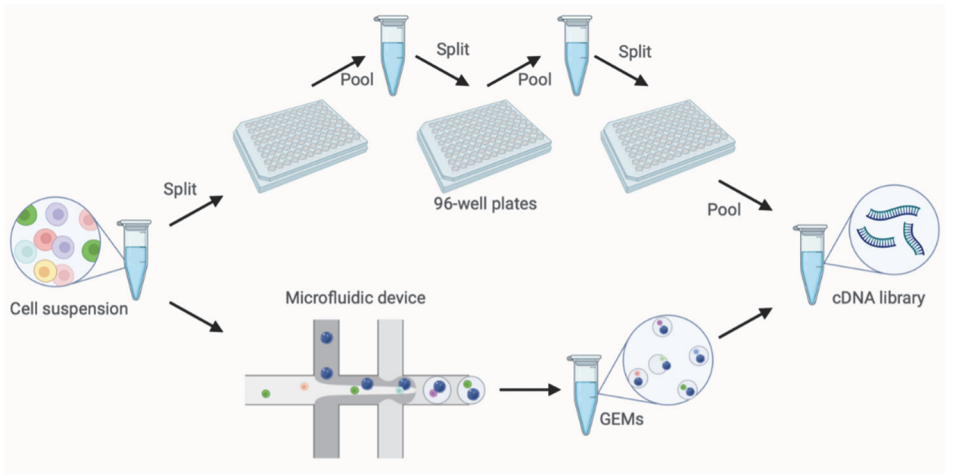


Figure 1. Comparison of split-pool and microfluidic scRNA-seq workflows. In the split-pool approach (upper), a cell suspension is divided across 96-well plates multiple times with pooling between splits to encode cell-specific barcodes. In the microfluidic approach (lower), individual cells are encapsulated into Gel Bead-in-Emulsions (GEMs) using a microfluidic device to associate each cell with unique barcodes. Both methods generate cDNA libraries for downstream sequencing and transcriptomic analysis.

Split-pool barcoding is an alternative high-throughput scRNA-seq technique that does not require a microfluidic instrument (Cao et al., 2017). This method involves the iterative splitting and pooling of cells in 96-well plates (Figure 1). The

samples are distributed into wells where the first-round barcode is introduced into each cell. Then, the cells from all wells are pooled together. In subsequent rounds, the barcoding procedure is repeated, which results in the ligation of additional barcodes. Due to the probabilistic nature of cell distribution in wells, each cell acquires its specific barcode. The split-pool barcoding approach has been commercialized by Parse Biosciences in the Evercode kits.

After isolating the cells, RNA extraction and reverse transcription are crucial processes. The RNA extracted from each cell is reverse-transcribed into complementary DNA (cDNA), often incorporating unique molecular identifiers (UMIs) to reduce amplification bias and enhance quantification accuracy (Islam et al., 2014). This reverse transcription can be carried out within the same droplets or wells where the cells were initially isolated, preserving the single-cell resolution. Including UMIs during cDNA synthesis allows for precise RNA molecule counting, ensuring accurate downstream quantification and analysis.

The next phase involves library preparation, where the cDNA is amplified, and sequencing adapters are added. This process generates libraries compatible with high-throughput sequencing platforms such as Illumina. The prepared libraries are then subjected to next-generation sequencing, which produces large volumes of data representing the transcriptomic landscape of individual cells. Sequencing outputs, known as reads, are short DNA sequences corresponding to original RNA molecule fragments. These reads are then processed and mapped back to a reference genome, enabling the gene expression quantification within each cell. High-throughput sequencing technologies allow the parallel sequencing of thousands to millions of cells, providing a comprehensive view of cellular heterogeneity and gene expression dynamics.

2.1.3 Challenges of single-cell

When planning a single-cell sequencing experiment, cost-effective strategies are paramount to ensure the feasibility and success of the study. One approach to reducing costs is through sample multiplexing, where multiple samples are barcoded and pooled together, allowing for simultaneous processing and sequencing. These include labelling cells with hashtag oligos (HTOs) or lipid-conjugated oligonucleotides prior to droplet formation (McGinnis et al., 2019; Stoeckius et al., 2018). Sample multiplexing saves on reagent and sequencing costs and helps balance the sequencing depth across samples. However, researchers must carefully consider the trade-offs between sequencing depth per cell, the number of cells per sample, and the number of samples in each group (Davis et al., 2019; Wu et al., 2017). Achieving an optimal balance is crucial: deeper sequencing provides more comprehensive data for each cell but reduces the number of cells that can be analyzed within a given budget.

Batch effects present a significant challenge in the scRNA-seq experiments, arising when samples are processed on different days or under varying conditions. These effects can introduce unwanted variability that confounds actual biological differences, complicating downstream data analysis. RNA-seq, in general, is

particularly sensitive to such batch effects, necessitating rigorous experimental design and data normalization techniques (McIntyre et al., 2011). Strategies to mitigate batch effects include using randomized sample processing, incorporating control samples across batches, and employing computational methods for batch effect correction (Tung et al., 2017). Ensuring consistency in sample handling, reagent preparation, and sequencing conditions is crucial to minimizing these artifacts.

The complexity of scRNA-seq protocols adds another layer of consideration, making the process labor-intensive and technically demanding. Microfluidics-based methods, while offering high-throughput and precise cell handling, come with their own set of challenges, such as device clogging, variability in droplet formation, and the need for specialized equipment and expertise (Arceneaux et al., 2023; Simmons & Lau, 2022). Additionally, the choice of tissue processing and RNA extraction techniques can significantly impact the quality and integrity of the resulting data (Wohnhaas et al., 2019). Different tissues may require specific dissociation protocols to preserve cell viability and RNA integrity, and suboptimal extraction methods can lead to RNA degradation or loss of specific cell populations (Denisenko et al., 2020; Madissoon et al., 2020). Therefore, careful optimization and validation of each step are essential to ensure robust and reproducible results in single-cell sequencing experiments.

2.2 Single-cell data analysis

2.2.1 Overview of the key bioinformatics techniques

Single-cell technologies have significantly enhanced our ability to analyze heterogeneous cell populations with exceptional resolution, revealing complex molecular interactions and previously undiscovered cellular states (Potter, 2018). As a result, computational methods designed explicitly for single-cell data have become indispensable, allowing researchers to extract meaningful insights from vast and multidimensional datasets. The development of computational methodologies for single-cell genomics has kept pace with the rapid advancements in sequencing technologies (Hie et al., 2020). Considering the abundance of available tools, researchers often seek the best pipelines for conducting single-cell analysis. The latest update of the scRNA-tools database includes over 1,700 published algorithms and utilities for single-cell gene expression analysis alone (Zappia et al., 2018). Several valuable guidelines have been published to assist in the tool selection process (Amezquita et al., 2020; Andrews et al., 2021; Heumos et al., 2023). However, the bioinformatics of single-cell genomics still needs to be standardized, with most popular workflows requiring extensive fine-tuning through trial and error.

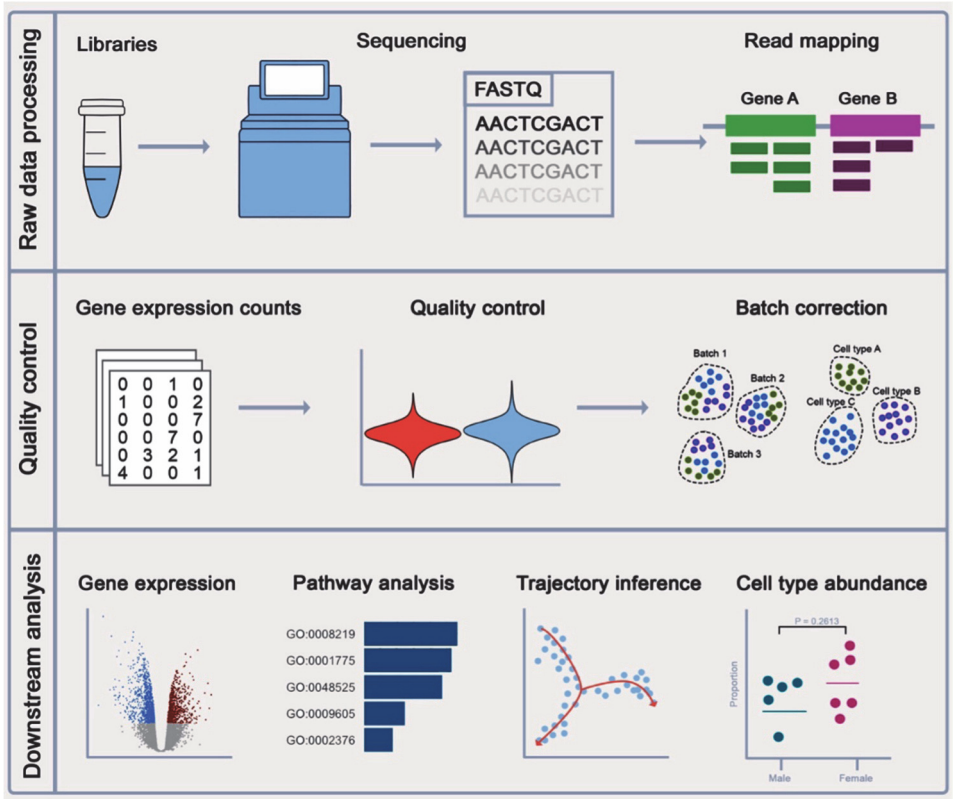


Figure 2. Overview of scRNA-seq data analysis pipeline. Raw data processing involves sequencing libraries, generating FASTQ files, and mapping reads to genes. Quality control ensures accurate gene expression counts, removes low-quality cells, and corrects batch effects. Downstream analysis includes identifying gene expression changes, pathway enrichment, trajectory inference for cellular dynamics, and cell type abundance comparisons across conditions.

The overview of scRNA-seq processing workflow is provided in Figure 2. The scRNA-seq analysis starts with processing raw sequencing data in FASTQ format, containing reads that must be prepared and organized for meaningful biological analysis. First, the data undergoes quality control (QC) to filter out low-quality reads, removing those with poor scores or potential sequencing errors (Haque et al., 2017). Next, the reads are aligned to a reference genome, which links the sequences to specific genes, enabling the quantification of gene expression levels for each cell. For commercial single-cell RNA sequencing assays, such as those offered by 10x Genomics, proprietary software solutions like *CellRanger* are commonly provided to facilitate the alignment and quantification processes through streamlined and integrated workflows. Alternatively, open-source tools, including *STARsolo*, *kallisto*, and *alevin-fry*, are available, enabling researchers to customize analytical pipelines (He et al., 2022; Kaminow et al., 2021; Sullivan et al., 2024). These open-source approaches can exhibit substantial

variability in computational efficiency and runtime, as well as differences in gene quantification outcomes, which arise from variations in their underlying methods and algorithms (Brüning et al., 2022; You et al., 2021). However, the application of open-source pipelines to commercial assays often necessitates additional effort, including the adaptation of preprocessing steps and resolution of compatibility issues, to ensure accurate and reproducible results (Kuijpers et al., 2024).

After initial processing, quality control extends to the cell level, where low-quality cells (often identified by high mitochondrial gene expression or low total gene counts) are filtered out to ensure only high-quality cells are included in downstream analyses (Andrews et al., 2021). Normalization techniques adjust for differences in sequencing depth and other factors to handle technical variability and enable cross-sample comparisons. At the same time, batch effect correction methods are applied to mitigate technical biases across samples.

Once the data is prepared, dimensionality reduction techniques like UMAP or t-SNE reduce complex high-dimensional data to two or three dimensions, making visualization and interpretation more accessible (Cakir et al., 2020). Clustering algorithms are then used to group cells with similar expression patterns, revealing distinct cell populations (Kiselev et al., 2019). This clustering step is essential for identifying known and novel cell types within the dataset. Additional analyses, such as cell type annotation (assigning identities to clusters), differential expression analysis (identifying genes that vary between conditions or clusters), and trajectory analysis (modeling dynamic processes like differentiation), provide further insights (Saelens et al., 2019). Together, these analyses enable an understanding of cellular heterogeneity, lineage relationships, and underlying biological processes.

2.2.2 Cell quality control

Quality control in scRNA-seq is essential due to the extensive processing cells undergo during isolation and library construction, which can introduce technical variability and artifacts (Amezquita et al., 2020). Rigorous QC measures are necessary to filter out damaged cells, contaminants, and inaccuracies that arise during these preparatory stages, ensuring the integrity of gene expression profiles for accurate downstream analysis. Challenges common to scRNA-seq protocols include distinguishing genuine cell-associated signals from background noise, addressing contamination from ambient RNA, and identifying artifacts such as multiplets or “doublets” where multiple cells are erroneously grouped (Hong et al., 2022). These steps are integral to maintaining data quality across diverse scRNA-seq methodologies.

In droplet-based scRNA-seq data (like 10x Genomics), the initial quality control phase includes distinguishing cell-associated barcodes from those of empty droplets. EmptyDrops, a popular algorithm embedded in the *CellRanger* workflow, is used for this purpose (Lun et al., 2019). A significant challenge in droplet data is the capture of cell-free “ambient” RNA, which occurs due to cell lysis during sample preparation. High levels of ambient RNA can make it difficult

to accurately differentiate between droplets containing cells and those with ambient RNA (Y. Zhang et al., 2023). A high proportion of reads not linked to cell barcodes suggests background RNA contamination, which requires correction. Software packages like *CellBender*, *SoupX*, and *DecontX* can estimate and correct the background RNA profile in the count matrices (Fleming et al., 2023; S. Yang et al., 2020; M. D. Young & Behjati, 2020). However, fine-tuning these algorithms may require expert knowledge of tissue-specific expression profiles. High-throughput single-cell assays can also produce multiplets where multiple cells share the same molecular barcode. In microfluidics-based protocols, this occurs when more than one cell is captured in a droplet. The performance of doublet detection tools has been thoroughly benchmarked, revealing significant variability in their effectiveness across different experimental settings (Germain et al., 2022; Xi & Li, 2021).

Damaged cells can have altered mRNA content, often showing increased mitochondrial transcript counts, a decreased number of detected genes and UMIs, and an increased nuclear RNA fraction (Muskovic & Powell, 2021). These changes can lead to spurious clusters that complicate downstream analyses. Tissue handling and cell isolation can also introduce specific gene signatures, like stress responses, further contributing to misleading results (Marsh et al., 2022). Key quality control metrics include the number of detected genes, UMIs, and the percentage of mitochondrial gene expression (Andrews et al., 2021). Thresholds for these metrics can be set based on expert knowledge or adaptively determined from the data, with consideration for the sample's cellular composition, as different cell types may naturally vary in mRNA content (Amezquita et al., 2020).

2.2.3 Accounting for technical variation

Addressing technical variation in single-cell sequencing analyses is crucial for accurate data interpretation. Factors such as variability in mRNA capture, cDNA conversion, amplification efficiency, sequencing depth, and stochastic molecular sampling can obscure actual biological differences between cells (L. Lun et al., 2016; Vallejos et al., 2017). The relationship between gene abundance and gene variance in scRNA-seq datasets can result in highly expressed genes disproportionately influencing the definition of cellular states in downstream analyses (Eling et al., 2018). Normalization techniques, such as library size normalization, are employed to standardize expression values to correct these disparities. Library size normalization allows the scaling of gene expression values by dividing the counts for each cell by a "library size factor," which is proportional to the total counts across all genes in that cell. This scaling is typically followed by log transformation that emphasizes more pronounced gene expression changes during downstream applications (Amezquita et al., 2020). Alternatively, *sctransform* package employs a model-based approach using a negative binomial distribution, estimating dataset-specific overdispersion parameters to stabilize the variance (Choudhary & Satija, 2022).

Unlike normalization, which adjusts for differences in sequencing depth and other within-sample factors, batch effect correction specifically targets discrepancies that arise when samples are processed at different times, under varying conditions, or with different assay types. Early methods like *limma* and *ComBat*, which rely on fitting linear models, were adapted from bulk RNA-seq but are not explicitly designed for single-cell data (Johnson et al., 2007; Ritchie et al., 2015). One of the fundamental challenges in scRNA-seq is that cells cannot be sequenced in multiple batches, as the process of sequencing destroys the cells. This limitation prevents direct comparisons of the same cells across batches, requiring computational methods to identify and correct batch effects indirectly by analyzing population-level gene expression patterns and relationships. To better handle scRNA-seq's unique challenges, newer approaches like Mutual Nearest Neighbors (MNN) and Canonical Correlation Analysis (CCA) have been developed. MNN corrects batch effects by identifying similar cells across batches and using expression differences for adjustment, while CCA uncovers shared gene correlation structures for data integration (Butler et al., 2018; Haghverdi et al., 2018). Recent advancements include deep learning-based algorithms like *scVI*, *scANVI*, and *scGen*, which model gene expression with greater accuracy, considering factors like latent cellular states and batch covariates, improving correction in complex datasets (Lopez et al., 2018; Lotfollahi et al., 2019; Xu et al., 2021). These methods enhance the ability to combine data from multiple sources and technologies and use hardware acceleration (such as GPUs) to process large-scale datasets quickly. The performance of these methods can vary depending on the tissue type and the specific characteristics of the dataset. While they enhance the ability to combine data from multiple sources and technologies, their application requires careful interpretation of results, as there is a risk of overcorrection, which can inadvertently obscure genuine biological variation (Luecken et al., 2022).

2.2.4 Cell type annotation

Cell type annotation, which involves classifying each cell into a specific type or state based on its molecular profile, is central to scRNA-seq analysis workflows (Andrews & Hemberg, 2018). This process enables the contextualization of data within existing biological knowledge by assigning the observed cell communities to known or novel cell types or states. Unlike techniques such as flow cytometry, which depends on predefined gating strategies and known markers to identify cell populations, scRNA-seq offers greater flexibility in exploring the cellular composition of a sample (See et al., 2018). Flow cytometry's reliance on existing antibodies may limit its ability to detect novel or intermediate cell states that do not have specific markers. In contrast, scRNA-seq can reveal these previously uncharacterized states.

However, the ability of scRNA-seq to define cell populations at varying levels of granularity can lead to differing interpretations of cellular composition in similar samples. For example, researchers may select different clustering resolu-

tions, identifying broad cell types or more precise subtypes (Z. Wang et al., 2020). The concept of "cell type" remains a topic of debate, intensified by the extensive use of single-cell technologies. Arendt suggested an evolutionary perspective, defining cell types as cells sharing a set of transcription factors regulating gene expression profiles (Arendt, 2008; Arendt et al., 2016). Morris proposed that cell identity should be understood regarding function, lineage, and potential responses to stimuli (Morris, 2019). More recently, the importance of perturbation experiments and multi-omics data integration in defining cell types has been highlighted (Fleck et al., 2023; H. Zeng, 2022). While these discussions have advanced our understanding of cell type definition in the single-cell era, practical and actionable definitions are still needed.

Published guidelines for cell annotation recommend integrating manual annotation with automatic classification tools (Z. A. Clarke et al., 2021). This "expert in the loop" approach involves assigning cell type labels using reference datasets and refining these labels based on domain expertise. Validation of cell type classifications is crucial, as novel cell types may not be accurately identified using existing references. The choice of a reference database depends on the specific organism and tissue being studied, and numerous databases and cell annotation tools have been developed for scRNA-seq (Aran et al., 2019; Domínguez Conde et al., 2022; Hao et al., 2021; Kiselev et al., 2018). Differential expression analyses can also be performed for each cell cluster to identify marker genes that define distinct cell communities.

2.2.5 The power of single-cell atlases

The concept of single-cell atlases emerged with the technological developments and increased availability of scRNA-seq assays. These represent large-scale datasets that comprise comprehensive maps of cellular diversity within an organism, tissue, or disease state at a single-cell resolution. The atlases provide remarkable scale and detail, allowing researchers to study the intricacies of cellular processes and interactions. By offering insights into cell type-specific gene expression, signaling pathways, and molecular characteristics, single-cell atlases facilitate a deeper understanding of biological systems and their dynamic changes in health and disease. The benefits of these atlases are manifold, including the potential to identify novel cell types, uncover mechanisms of disease progression, and characterize population heterogeneity at the single-cell level (Wilbrey-Clark et al., 2020).

The availability of public datasets has accelerated the construction of single-cell atlases. Researchers can build comprehensive resources encompassing various biological contexts by integrating and analyzing the data made available by others. Leveraging public data or using it to supplement an in-house dataset enhances the scope and scale of single-cell atlases and promotes reproducibility (Wilbrey-Clark et al., 2020). The Human Cell Atlas is one of the pioneering and most extensive collections of published data encompassing a reference catalog of human cells across multiple tissues (Regev et al., 2017). More focused atlas pro-

jects, among others, include the atlas of ageing tissues in the mouse (Tabula Muris Senis), the murine Liver Cell Atlas, Human Lung Cell Atlas, and Mouse Kidney Atlas (Guilliams et al., 2022; Novella-Rausell et al., 2023; Sikkema et al., 2023; The Tabula Muris Consortium et al., 2020).

Large-scale scRNA-seq projects provide comprehensive cellular resolution, enabling precise cell type annotation by capturing individual cells' heterogeneity and dynamic states within complex tissues. This high-resolution data facilitates the discovery of novel cell types and rare subpopulations, enhancing our understanding of cellular functions, interactions, and their roles in health and disease (Rood et al., 2024). Therefore, cell type annotation-related tasks are often a primary focus in scRNA-seq atlas projects. Annotations from projects like Tabula Muris have been utilized as references for automated cell type annotation (Kimmel et al., 2019; The Tabula Muris Consortium et al., 2018, 2020).

2.2.6 Data re-analysis and reproducibility

Bioinformatics data re-analysis is essential for advancing scientific discovery and ensuring reproducibility (Sielemann et al., 2020). Leveraging published data repositories such as the Gene Expression Omnibus (GEO), ArrayExpress, and others provide researchers access to extensive secondary analysis datasets. These repositories facilitate the re-use of data for novel research questions and allow for the cross-validation of findings and the refinement of computational methodologies. The ability to re-analyze existing datasets enables the validation of prior results and the identification of new biological insights, thereby enhancing the reliability and depth of scientific conclusions.

The importance of metadata sharing cannot be overstated in this context (Puntambekar et al., 2021; Sarfraz et al., 2024). Comprehensive metadata, encompassing details about experimental design, sample characteristics, and data processing techniques, is essential for the accurate interpretation and reproducibility of datasets (Füllgrabe et al., 2020). This practice enables other researchers to understand the experimental nuances, replicate studies, and integrate data across different studies. Code sharing in bioinformatics further improves reproducibility and transparency. The open dissemination of code used for data analysis enables the scientific community to replicate analytical workflows, validate computational methods, and extend existing research (Abdill et al., 2024; Cadwallader et al., 2022). In single-cell omics, where intricate computational pipelines are often employed, sharing code is critical for verifying the procedures underlying scientific findings. Platforms like Zenodo and GitHub facilitate code sharing and collaborative development. By promoting the open exchange of data, metadata, and code, the bioinformatics field fosters a culture of transparency and collaboration, thereby enhancing the robustness and reproducibility of scientific research (Cokelaer et al., 2023).

2.3 Biological aging

2.3.1 Theories of aging

Aging is a complex biological process marked by a gradual decline in physiological function, increased susceptibility to disease, and mortality. Theories of aging generally fall into two categories: error accumulation theories and programmed aging theories (Lipsky & King, 2015). Error theories suggest that aging arises from random cellular damage or errors, such as somatic mutations, DNA sequence loss, protein synthesis errors, or oxidative damage (Harman, 1992; Medvedev, 1972; Orgel, 1963; Szilard, 1959). In contrast, programmed theories propose that aging is genetically regulated, potentially as an evolutionary adaptation, through delayed effects of harmful genetic programs, or pleiotropic genes with age-specific trade-offs (Blagosklonny, 2013; Williams, 1957). A distinction must be made between the genetic determinants of the aging process and those that influence longevity (Hayflick, 2007). Evidence for the latter in model organisms, such as mice, shows that specific genes influence longevity and healthspan (Bou Sleiman et al., 2022; Z. Zhang et al., 2023).

2.3.2 Hallmarks of aging

López-Otín and colleagues proposed the hallmarks of aging, a comprehensive framework that organizes the biological processes contributing to the gradual decline in cellular and organismal function over time (López-Otín et al., 2013, 2023). The hallmarks of aging are broadly categorized into primary, antagonistic, and integrative (Figure 3). Primary hallmarks are the underlying causes of cellular damage, including genomic instability, telomere attrition, epigenetic alterations, and loss of proteostasis. Antagonistic hallmarks, such as deregulated nutrient sensing, mitochondrial dysfunction, and cellular senescence, reflect the body's compensatory responses to damage but can become harmful when dysregulated. Integrative hallmarks are the combined effects of primary and antagonistic hallmarks, leading to stem cell exhaustion and chronic inflammation. The hallmarks of aging are closely linked to the hallmarks of health, which encompass processes that maintain physiological integrity and resilience (López-Otín & Kroemer, 2021). For example, the integrity of barriers, such as intracellular membranes, is crucial for protecting against environmental insults and pathogens; however, genomic and epigenomic instability can compromise this integrity (Kelley et al., 2011).

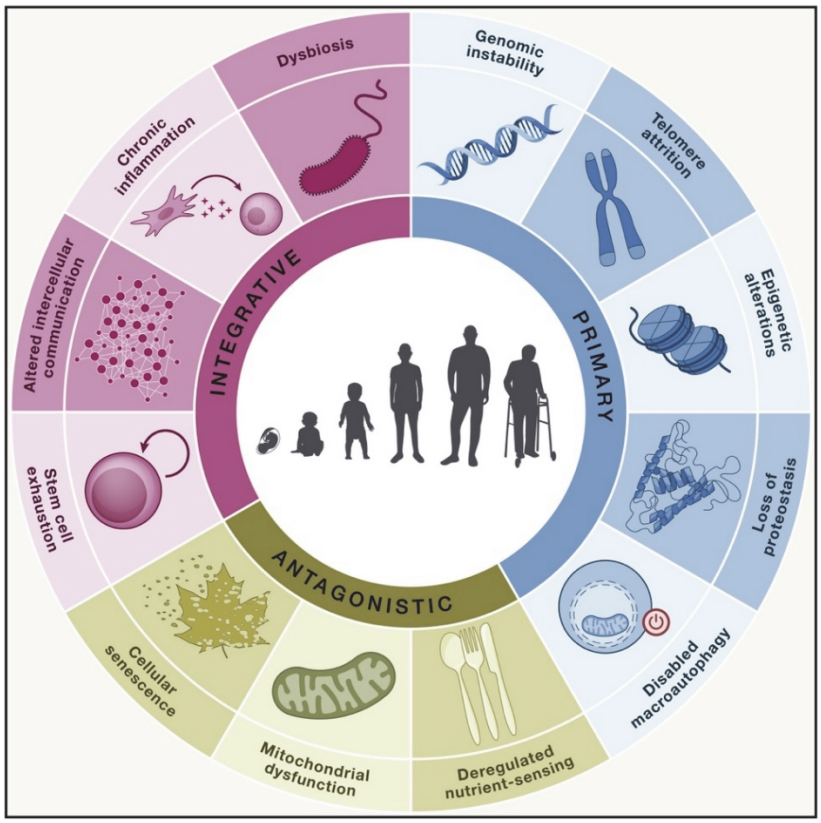


Figure 3. Hallmarks of aging categorized into primary, antagonistic, and integrative processes. Primary hallmarks, including genomic instability, telomere attrition, epigenetic alterations, and loss of proteostasis, are the root causes of cellular damage. Antagonistic hallmarks, such as mitochondrial dysfunction, cellular senescence, deregulated nutrient sensing, and disabled macroautophagy, act as responses to damage. Integrative hallmarks, including stem cell exhaustion, altered intercellular communication, chronic inflammation, and dysbiosis, drive functional decline and organismal aging. Figure from López-Otín et al., 2023 (© 2022 Elsevier Inc.)

Furthermore, there is a strong connection between the hallmarks of aging and the development of age-related diseases. Many of the same molecular pathways that drive aging also contribute to the onset and progression of diseases such as cancer, neurodegenerative disorders, cardiovascular diseases, and metabolic conditions (Blank & Bellizzi, 2008; Donato et al., 2018; Hou et al., 2019). For example, shorter telomeres can reduce the regenerative capacity of certain cells, diminishing the ability to repair tissues (X. Zhang et al., 2022). This decline in regenerative ability can lead to tissue remodeling and increase the risk of conditions such as fibrosis or emphysema (Alder & Armanios, 2022). In addition, methylation patterns have been demonstrated to be predictive of age across various species and can also serve as indicators of mortality risk (A. T. Lu et al., 2019, 2023).

2.4 Aging immunity

2.4.1 Components of immune system

Broadly, the mammalian immune system can be divided into innate and adaptive immunity. Innate immunity protects the organism by mounting a rapid response to pathogens. It includes various cells, such as macrophages, neutrophils, dendritic cells (DCs), and natural killer (NK) cells, as well as the complement systems and barrier immunity components. The innate immune response is non-specific and provides a broad defense against various pathogens through mechanisms like phagocytosis and the release of antimicrobial substances (Turvey & Broide, 2010). This system recognizes common pathogen-associated molecular patterns and quickly responds to eliminate invading pathogens. The innate immune cells, such as macrophages, have a vital role in tissue repair and cellular debris clearance (Murray & Wynn, 2011).

On the other hand, adaptive immunity develops a targeted response to specific pathogens. B cells produce antibodies that bind to and neutralize pathogens and activate the complement system (Nothelfer et al., 2015). Upon activation, T cells can coordinate immune responses by releasing cytokines that activate other immune cells, directly kill infected cells, and help maintain immune tolerance and prevent autoimmune responses (Dong, 2021; Golstein & Griffiths, 2018; Mohr et al., 2018). In contrast to innate immunity, adaptive immunity also has a memory component, allowing the immune system to respond more efficiently upon subsequent exposures to the same pathogen. This memory aspect is the basis for the effectiveness of vaccines, which prepare the adaptive immune system to fight specific infections by introducing antigens in a controlled manner, leading to long-term immunity (Cobey, 2024).

Interactions between innate and adaptive immunity are critical in mounting a robust immune response, with antigen presentation as a conduit between the two subsystems (Cherrier et al., 2020). Through cells such as DCs and macrophages, the innate immune system initially recognizes and engulfs pathogens. These cells process the foreign antigens and present them on their surface using major histocompatibility complex (MHC) molecules. DCs, in particular, migrate to lymphoid organs, where they interface with the T cells. The MHC-antigen complex is recognized by T cell receptors (TCRs), thus activating them (Lever et al., 2014). This activation is essential for the clonal expansion and differentiation of T cells into effector cells, which then coordinate a targeted immune response. Furthermore, co-stimulatory signals and cytokine secretion enhance the interaction between innate antigen-presenting cells (APCs) and T cells, ensuring a precise and effective adaptive immune response (Dong, 2021). The interactions highlight the necessity of close collaboration between adaptive and innate immunity, promoting a comprehensive defense mechanism against pathogens.

2.4.2 Immunosenescence

The central role of immunity in aging was recognized in the 1960s by Roy Walford (Walford, 1964). He introduced the term "immunosenescence", which refers to the gradual decline of the immune system associated with aging. This process impacts both the innate and adaptive branches of the immune system, leading to increased susceptibility to infections and autoimmune diseases, as well as reduced efficacy of vaccinations in older adults (Liu et al., 2023). Aging is associated with functional defects and inefficiencies across multiple immune cell subsets, including T cells, B cells, macrophages, and other cells, contributing to a significant decline in the overall immune response (Simon et al., 2015). This decline compromises the body's ability to mount effective defenses against pathogens and to maintain immune homeostasis (Kennedy et al., 2016; Zinatizadeh et al., 2023).

Aging is linked to the alterations in hematopoietic stem cells (HSCs) and concurrent changes in the bone marrow microenvironment (De Haan & Lazare, 2018; Ho & Méndez-Ferrer, 2020). As the hematopoietic system ages, the differentiation potential of HSCs declines, resulting in changes in the production of multiple immune cell types. Studies in mice have indicated that even in middle-aged animals, there is already a pronounced bias toward the myeloid lineage in hematopoietic differentiation (Konturek-Ciesla et al., 2023; K. Young et al., 2021). The lineage tracing and scRNA-seq experiments carried out by Konturek-Ciesla and colleagues showed that this bias arises at the level of the primitive lymphoid-primed multipotent progenitor cells, which are depleted in aged animals (Konturek-Ciesla et al., 2023).

Aging is also associated with chronic low-grade inflammation, termed inflammaging, marked by the continuous release of pro-inflammatory cytokines and other immune mediators (Franceschi et al., 2000; Fülöp et al., 2019). Unlike acute inflammation, which is a response to infection or injury, inflammaging represents a persistent, systemic state of immune activation independent of invading pathogens. The pro-inflammatory factors characteristic of inflammaging include various interleukins (ILs), which are cytokines that play a role in regulating immune responses, such as IL-1, IL-6, IL-8, IL-13, and IL-18, as well as C-reactive protein, interferons (IFN-α and IFN-β), transforming growth factor-beta, tumor necrosis factor, and others (Ferrucci & Fabbri, 2018). Inflammation is a high-cost trait, and its benefits are most significant early in life, while the costs (resulting from the chronic inflammatory process) are often delayed until later life when they have less impact on reproductive success, exemplifying antagonistic pleiotropy (Okin & Medzhitov, 2012).

2.4.3 Innate immunity ageing

Aging affects the function and composition of both innate and adaptive immune cell compartments by altering the number, diversity, and efficiency of these cells. This can lead to a diminished immune response and increased susceptibility to infections due to changes in such innate populations as monocytes, macrophages, NK cells, and others. Hematopoiesis in aging is associated with a shift in output, favoring the production of myeloid lineage cells over lymphoid lineage cells, a phenomenon referred to as myeloid bias (Pang et al., 2011). In this process, myeloid lineage is increasingly favored, contributing to a decline in adaptive immune function and an increase in myeloid cell-related pathologies, such as chronic inflammation and myeloid malignancies, as observed in elderly populations (Chung & Park, 2017).

Monocytes circulate in the bloodstream and historically were defined as precursor cells that can differentiate into macrophages or myeloid-derived DCs. However, they are no longer viewed merely as immature precursors but as distinct cell types that can rapidly exert effector function (Guilliams et al., 2018). Research in humans has revealed changes in both the composition and function of peripheral monocytes. Hearps and colleagues found that the proportion of intermediate and non-classical monocytes increases with age, with impaired function evidenced by reduced phagocytic activity (Hearps et al., 2012). Recently, Reitsema and colleagues discovered a higher frequency of classical and non-classical monocytes in the peripheral blood of the elderly, accompanied by increased expression of activation markers in classical and intermediate monocytes (Reitsema et al., 2024). The proportion of classical monocytes in the airways of adult donors was previously found to be higher than in children but declines with advanced age (Byrne et al., 2020). Studies of tissue-resident macrophages in mice have also demonstrated a functional decline in this population (Kelly et al., 2007; McQuattie-Pimentel et al., 2021).

DCs are antigen-presenting immune cells crucial for initiating and regulating the body's adaptive immune response (Cabeza-Cabrerizo et al., 2021). DC subsets also experience alterations in abundance and function as they age. A lower proportion of plasmacytoid dendritic cells (pDCs) was observed in the peripheral blood of older donors compared to young individuals (Reitsema et al., 2024). The investigation into impaired humoral responses to vaccination in older adults revealed defects in conventional dendritic cells type 2 (cDC2) caused by decreased type 1 interferon signaling (Stebegg et al., 2020).

NK cells are innate lymphoid cells essential for the body's defense against tumors and virally infected cells (Lam & Lanier, 2017; Shimasaki et al., 2020). Like the alterations observed in other innate immune cells mentioned earlier, the NK compartment also experiences remodeling with aging. It has been observed that aging is associated with a reduction in the $CD56^{high}$ subset and a simultaneous increase in the $CD56^{low}$ NK cells (Campos et al., 2014; Chidrawar et al., 2006; Solana et al., 2014).

2.4.4 B cell ageing

B cell-mediated immunity undergoes pronounced changes with aging, characterized by the diminished response to pathogens and vaccines, which compromises the body's ability to mount effective immune responses in older adults. Studies in mice showed a notable reduction in the number of B cell precursors, highlighting the impact of aging on the early stages of B cell development (Alter-Wolf et al., 2009). A decrease in the number of B cell precursors has also been observed in studies of human bone marrow (McKenna et al., 2001). Alterations in peripheral cell type composition also mark B cell aging. Recent study of human PBMCs have demonstrated an increase in the proportion of memory B cells, accompanied by a decrease in peripheral naive B cells (Jalali et al., 2022). In addition, as organisms age, B cells acquire intrinsic defects that impair their ability to mount effective defenses against invading pathogens. Studies in mice and humans revealed that the expression of activation-induced cytidine deaminase is reduced in B cells with aging, which negatively affects the ability to generate high-affinity antibodies through somatic hypermutation (Frasca et al., 2004, 2008).

Atypical B cells (also referred to as aging-associated B cells, or ABCs), are a subset of B lymphocytes that increase with age in mice and are linked to autoimmunity in humans (Cancro, 2020). Characterized by the expression of T-bet and CD11c, ABCs rely on TLR7 or TLR9 signaling for their development. In autoimmune-prone mouse strains, ABCs emerge earlier and exhibit more pronounced expansion, with TLR7 driving their proliferation (Nickerson et al., 2023; Rubtsov et al., 2011). Comparable subsets of $CD11c^+$ B cells in human blood share similarities with murine ABCs (Dorfman et al., 2004; Jöhrens et al., 2006). Elevated ABC levels have been observed in patients with autoimmune diseases such as systemic lupus erythematosus and rheumatoid arthritis, reinforcing their association with autoimmunity (Holla et al., 2021; McGrath et al., 2024). Studies show that TLR9 or TLR7 activation primes B cells for ABC differentiation, with subsequent cytokine signals like IFN-γ or IL-21 promoting this fate (Naradikian et al., 2016). Evidence suggests that ABCs arise through antigen-driven responses requiring T cell help and germinal center involvement, with somatic mutations and diverse immunoglobulin gene usage consistent with their germinal center origin (Russell Knode et al., 2017).

2.4.5 T cell ageing

The thymus is a crucial part of adaptive immunity as it plays an integral role in the development and maturation of T lymphocytes. Aging is associated with thymic involution, a process in which the thymus progressively decreases in size and function, leading to a decline in the production of naïve T lymphocytes and impacting immune system efficacy (Palmer et al., 2018). In humans, the export rate of naïve T cells from the thymus reaches its highest point at one year of age, then rapidly decreases until eight years old, after which it continues to decline

gradually until the age of twenty (Bains et al., 2009). Studies in mice showed noticeable morphological changes in the thymus observed at four weeks of age, such as cortical thinning and the merging of medullary islands, which paralleled changes in the cellularity of thymocytes and thymic epithelial cells (Baran-Gale et al., 2020).

The impact of reduced thymic activity on maintaining a naive T cell compartment is highly species-specific, with young mice relying heavily on thymic output for T cell production. In contrast, in young adult humans, most T cells are generated through peripheral proliferation rather than thymic activity (Goronzy & Weyand, 2019). In humans, homeostatic proliferation maintains a sufficient number of naive $CD4^+$ T cells, whereas the decline in circulating naive $CD8^+$ T cells with age is more pronounced and is a marker of immune aging in healthy older adults (Duggal et al., 2018; Thome et al., 2016).

As individuals age, T cells also accumulate intrinsic defects that result in a diminished ability to respond to antigens and contribute to a decline in overall immune function. Studies of aged T cells point to mitochondrial dysfunction marked by dysregulation of oxidative phosphorylation and energy metabolism (Bektas et al., 2019). In mice, dysfunctional mitochondrial transcription factor A (TFAM) leads to premature aging and various aging-associated defects, along with the onset of chronic inflammation (Desdín-Micó et al., 2020). Mitochondrial reactive oxygen species (ROS) in human $CD8^+$ T cells are correlated with telomere length which further emphasizes the role of mitochondria in T cell aging (Sanderson & Simon, 2017). The ROS-induced damage is further exacerbated by reduced autophagy, a decline in the cellular process responsible for removing damaged mitochondria and other cellular debris (A. J. Clarke & Simon, 2019).

Cellular senescence is characterized by a permanent state of cell cycle arrest, altered phenotype, and a pro-inflammatory secretory profile, contributing to tissue dysfunction and aging. T cell senescence becomes a significant factor in the decline of immune function. In T cells, senescence manifests as reduced proliferative capacity, CD27 and CD28 expression loss, shorter telomere length, increased apoptosis, and increased expression of inhibitory receptors like KLRG1 (Akbar et al., 2016). One notable T subset that exhibits senescent cell features is the terminally differentiated effector memory T cells re-expressing CD45RA (Temra cells). Temra cells accumulate with age and are characterized by loss of TCR signaling and increased expression of NK cell receptors and cytotoxicity molecules (Pereira et al., 2020; Quinn et al., 2018). Interestingly, $CD4^+$ Temra cells maintain higher mitochondrial content than $CD8^+$ counterparts, enabling greater metabolic flexibility and enhanced proliferation and migration (Callender et al., 2020).

2.4.6 CMV and immunosenescence

Chronic infections in aging significantly impair immune function, reducing the individual's ability to respond to new infections and establish immunological memory (Ripa et al., 2017). This prolonged immune response induces exhaustion

and senescence in immune cells, compromising their functionality (Utzschneider et al., 2020). Chronic infections promote clonal expansion of specific T cell populations, thereby reducing their repertoire diversity and impairing the immune system's capacity to respond to novel antigens effectively (Schober et al., 2018). These immunological alterations contribute to the overall decline in immune competence observed in the elderly.

CMV is a ubiquitous herpesvirus that establishes lifelong latency following primary infection, often remaining asymptomatic in healthy individuals. CMV seroprevalence varies globally, with an estimated 66% in the European region and up to 90% in the Eastern Mediterranean region (Zuhair et al., 2019). CMV has the ability to manipulate host immune responses, for example, by undermining NK activating receptors or hampering antigen presentation (Berry et al., 2020). Although CMV infection is generally asymptomatic due to the immune system's ability to contain the virus, it can still trigger a chronic inflammatory response, potentially leading to long-term adverse effects (García-Torre et al., 2021).

Memory inflation of T cells is a phenomenon where specific T cell populations, especially in response to persistent infections like CMV, expand and maintain high frequencies over time (Kim et al., 2015; Snyder et al., 2011). Unlike typical memory T cells, which contract after infection clearance, these inflated T cell populations persist in large numbers due to ongoing antigen stimulation, potentially compromising the immune response to new infections.

2.5 Autoimmune polyendocrine syndrome type 1

2.5.1 Disease origins and clinical manifestations

Autoimmune Polyendocrine Syndrome Type 1 (APS-1), also known as Autoimmune Polyendocrine Syndrome Type 1 (APECED), is a rare recessively inherited autoimmune disorder caused by mutations in the autoimmune regulator (AIRE) gene. The *AIRE* gene encodes a transcriptional regulator crucial for establishing central immune tolerance. This protein is primarily expressed in medullary thymic epithelial cells (mTECs), where it promotes the ectopic expression of various tissue-specific antigens, a process vital for the negative selection of autoreactive T cells. Symptoms of APS-1 typically appear during childhood or adolescence. The prevalence of APS-1 is higher in certain isolated populations, such as Finns, Iranian Jews, and Sardinians, compared to the general population (Husebye et al., 2009).

The diagnosis is typically based on the presence of at least two components out of the classical triad (Whitaker's triad): chronic mucocutaneous candidiasis, adrenal insufficiency, and hypoparathyroidism. Other manifestations are also recognized, including primary ovarian failure, type 1 diabetes, enamel hypoplasia, alopecia, vitiligo, and autoimmune hepatitis, with patients having five components, on average. The mortality rates are significantly increased (up to 10-fold)

in APS-1 patients compared to general population with endocrine diseases having the highest rates. APS-1 is also characterized by high levels of autoantibodies (aAbs), and their presence can be used for diagnostic and prognostic purposes. Most commonly, patients possess autoantibodies (aAbs) targeting type I interferons as well as cytokines specific to T helper 17 (Th17) cells, which play a critical role in orchestrating inflammatory responses (Fishman et al., 2017; Orlova et al., 2017). The aAbs against interferon alpha (IFN-α) are most commonly shared between the APS-1 patients, however, many patients possess a unique aAbs repertoire (Fishman et al., 2017).

2.5.2 Central tolerance and AIRE

Central tolerance is an essential process for maintaining immune homeostasis by eliminating autoreactive immune cells (L. Klein & Petrozziello, 2024). Dysregulation of this process can lead to autoimmune diseases, such as APS-1, as autoreactive T cells escape central selection, infiltrate peripheral tissues, and cause immune-mediated pathology. The selection of T cells in the thymus ensures the development of a diverse yet self-tolerant T cell repertoire, vital for effective immune responses. Early T cell progenitors originate from HSCs and, upon entering the thymus, undergo TCR gene rearrangement through V(D)J recombination, creating a diverse range of TCRs capable of recognizing numerous antigens.

Positive selection occurs in the thymic cortex, where cortical thymic epithelial cells (cTECs) present antigens to developing thymocytes, ensuring that only T cells with functional TCRs capable of recognizing peptide-MHC complexes survive (Figure 4). Thymocytes that fail to interact successfully with cTECs undergo apoptosis. Surviving single-positive thymocytes then migrate to the thymic medulla for negative selection, where mTECs present tissue-restricted antigens (TRAs). Additionally, DCs and B cells also contribute to antigen presentation during this stage, further aiding in the negative selection process (Hadeiba et al., 2012; Yamano et al., 2015). Thymocytes with high reactivity to self-pMHC are either induced to apoptosis or directed toward the Treg lineage (L. Klein et al., 2019).

The role of AIRE in tolerance and autoimmunity has been recognized since its discovery, as it was initially identified through studies analyzing data from APS-1 patients (Aaltonen et al., 1997; Nagamine et al., 1997). There is evidence that AIRE functions as a DNA-binding protein, as it forms dimers and tetramers capable of binding to specific oligonucleotides (Kumar et al., 2001; Purohit et al., 2005). Notably, AIRE exerts its regulatory function by forming a multiprotein complex through interactions with partners such as CBP, P-TEFb, and DNA-PK, among others (Liiv et al., 2008; Oven et al., 2007; Pitkänen et al., 2000, 2005). In the thymus, *AIRE* is expressed mainly in mTECs, where it controls the expression of TRAs (Chen et al., 2008; Sin et al., 2023; Sousa Cardoso et al., 2006).

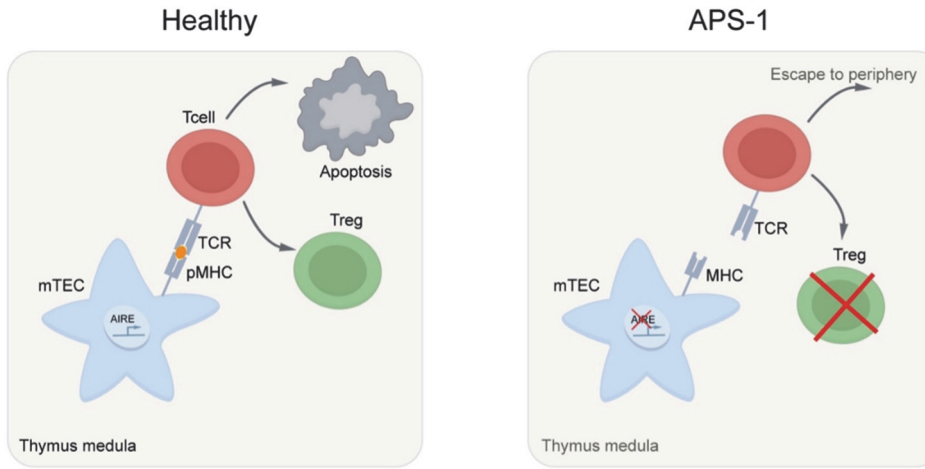


Figure 4. The role of AIRE in thymic selection. AIRE regulates the expression of tissue-restricted antigens in the thymus, promoting the negative selection of T cells. Through this process, self-reactive T cells are either directed toward the Treg lineage or undergo apoptosis. In APS-1 patients, defects in AIRE lead to a failure of effective negative selection.

2.5.3 Peripheral T cells in AIRE deficiency

Given the significant impact of AIRE deficiency on the development of conventional and regulatory T cells, there is considerable interest in understanding the downstream effects of impaired thymic function in APS-1. Studies in transgenic AIRE-deficient mice and rats have shown that a lack of AIRE leads to lymphocyte infiltration in the periphery (Hubert et al., 2009; Ossart et al., 2018). In both models, an increase in $CD25^+$ activated T cells was observed, while no significant differences were noted in other T cell subsets across various organs. Data on peripheral T cell subsets in APECED patients is more limited compared to the information available from animal models of the disease. In the peripheral blood of APS-1 patients, an increase in cells with an effector phenotype has been reported, accompanied by dysregulation of the IL7-IL7R pathway (Laakso et al., 2011).

Notable differences were observed in the $CD4^+$ T cell compartment of APS-1 patients compared to healthy individuals. Reports on the frequency of various conventional CD4 T cell subsets are limited. One group reported no changes in the Th1, Th2, and Th17 subsets in APS-1 patients; however, they observed an increase in cells with an effector phenotype and a decrease in naive cells (Heikkilä et al., 2016). More recently, a decrease in Th1 precursors has been reported in APS-1 patients (Sng et al., 2019). The frequency and phenotype of Treg cells in peripheral blood have been reported to be altered in APS-1 patients. Treg cells from APS-1 patients exhibit downregulation of FOXP3, the master regulator of Treg cell identity (Kekäläinen et al., 2007; Laakso et al., 2010). These intrinsic cellular changes are also accompanied by an overall decrease in the number of Tregs in circulation (Sng et al., 2019; Wolff et al., 2010).

3 SUMMARY OF THE LITERATURE REVIEW

The role of single-cell genomics in advancing our understanding of immunity in health and disease contexts is increasingly recognized. This technology enables the identification of biomarkers for various immune-related conditions. By providing high-resolution insights into cellular heterogeneity, single-cell genomics allows for a more detailed exploration of immune system functions and dysfunctions. Bioinformatics plays a crucial role in processing and analyzing the high-dimensional data generated by single-cell technologies, enabling researchers to interpret complex cellular behaviors and interactions. This has led to significant breakthroughs in understanding immune responses in various diseases and has improved our ability to track changes in immune cells over time. Despite these advancements, substantial gaps in knowledge and technical challenges remain.

The use of single-cell genomics in specific contexts, such as APS-1, is still limited, and further investigation is needed. Analyzing novel scRNA-seq datasets could significantly advance our understanding of the mechanisms underlying the breakdown of immune tolerance in this disease, offering detailed insights into the cellular and molecular interactions involved. Similarly, our knowledge of immunity at the intersection of CMV infection and aging remains incomplete, with single-cell approaches offering the potential to untangle the complex relationship between chronic viral infections and immune senescence. Moreover, the quality and reliability of aging biomarkers can be greatly enhanced through the large-scale re-analysis of single-cell datasets, enabling the identification of more robust and reproducible markers of immune aging. This could lead to a better understanding of immune decline and more precise interventions to address age-related immune dysfunction.

4 AIMS OF THE RESEARCH

This thesis focuses on two major aspects: benchmarking commercial scRNA-seq technologies and the applications of scRNA-seq in immunology. The first part presents a comprehensive evaluation, assessing the performance and limitations of two high-throughput scRNA-seq assays to inform their use in research. The second part explores the applications of scRNA-seq in immunology, uncovering novel biomarkers associated with various conditions. Together, these efforts advance our understanding of both the technological capabilities and the biological applications of scRNA-seq.

The specific aims were the following:

1. To conduct a comparative analysis of two commercial scRNA-seq assays for studying the thymus, a complex organ central to immune function (Paper I).
2. To identify immune aging biomarkers that are reproducible across different cohorts using published scRNA-seq aging datasets (Paper II).
3. To investigate the heterogeneity of $CD8^+$ Temra cells in the context of aging and CMV infection using a novel scRNA-seq dataset (Paper III).
4. To characterize the transcriptome of Tregs in APS-1 patients and compare it to healthy controls using scRNA-seq data from peripheral blood (Paper IV).

5 MATERIALS AND METHODS

5.1 Ethical approval

The study in Paper I was conducted in accordance with the permission from the Ministry of Regional Affairs and Agriculture (Estonia) and approved by the Animal Experiments Ethics Committee at the Ministry (Protocol No. 224). The ethical permissions for studies included in Paper II were obtained by the original authors. The study from Paper III was approved by the Ethics Review Committee of Human Research of the University of Tartu (permissions no 272/T-12, 275M-17, and 368/M-4).

The study from Paper IV was conducted with research permission HUS/1127/2016, including ethical review board approval from the Helsinki University Hospital Medical Ethical Review Board. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

5.2 Murine thymus data generation and processing (Paper I)

The thymi from two female C57BL/6N mice aged two months were isolated, and each lobe was separated and processed separately as a technical replicate. Each sample was divided equally and further processed by Chromium Next GEM Single Cell 3' Reagent Kit v3.1 (10x Genomics; 10x_A1, 10x_A2, 10x_B1, 10x_B2) and Parse Evercode Whole Transcriptome (WT) kit v2 (Parse Bio; parse_A1, parse_A2, parse_B1, parse_B2), respectively. The four 10x Genomics samples were multiplexed using TotalSeq™-B hashtag antibodies (BioLegend). The Parse Bio samples were first fixed using the Fixation Kit (Parse Bio). The samples were multiplexed and processed using the Parse WT kit as per manufacturer's protocol. All libraries were sequenced as paired-end 150-bp reads using the Illumina NovaSeq 6000 platform. The experimental design is summarized in Figure 5.

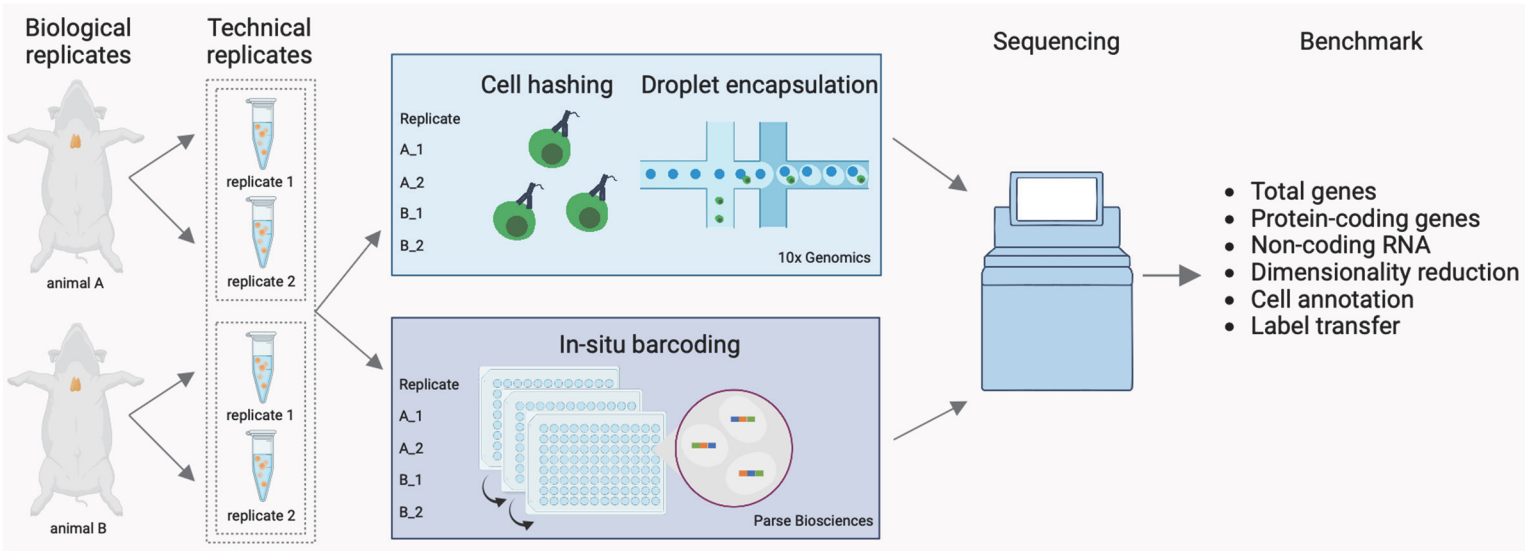


Figure 5. The benchmark experiment design. Thymi from two mice were divided into two technical replicates each. The resulting four samples were further divided between 10x Genomics and Parse Bio workflows.

The 10x Genomics data FASTQ files were mapped to the mm10 mouse reference (downloaded from the 10x Genomics website) using *CellRanger* software. For Parse Bio data, we first generated an indexed genome using the mm10 mouse reference downloaded from Ensembl. Next, we used the *split-pipe* tool from Parse Bio to process files from each sub-library. The results from each sublibrary were then combined into one gene expression matrix.

5.3 Public aging data collection and processing (Paper II)

We first performed a comprehensive literature review to identify the scRNA-seq studies of PBMCs from old and young human donors (Figure 6). We selected seven studies that matched our criteria and downloaded the gene expression matrices from the Luo et al., Zhu et al., and Thomson et al. from the Gene Expression Omnibus (accessions GSE157007, GSE213516, and GSE214546). The raw sequencing data from Zheng et al., Hung et al., and Li et al. were obtained from the National Genomics Data Center (accessions HRA000203, HRA000624, and HRA003766). The Mogilenko et al. raw data files were downloaded from Synapse (accession syn22255433).

Table 1. Cracteristics of published aging datasets from Paper II.

	Ages young	Ages Old	Young	Old	Female	Male	Reference
GSE157007	23–27	72–100	2	6	3	5	(Luo et al., 2022)
GSE213516	28–40	60–77	5	7	7	5	(Zhu et al., 2023)
GSE214546	27–35	60–63	16	9	13	12	(Thomson et al., 2023)
HRA000203	20–40	60–70	7	5	5	7	(Y. Zheng et al., 2020)
HRA000624	20–30	60–80	10	10	10	10	(Huang et al., 2021)
HRA003766	23–25	61–63	3	3	6	-	(J. Li et al., 2023)
syn22255433	26–29	62–70	10	10	-	20	(Mogilenko et al., 2021)

The data for HRA000203, HRA000204, HRA003766, and syn22255433 studies were obtained as FASTQ read files. Therefore, we first ran the *CellRanger* pipeline from 10x Genomics to perform the gene expression quantification. To this end, we used the pre-built GRCh38 human reference obtained from 10x Genomics. After acquiring the gene expression matrices for all the studies under investigation, we created a combined *AnnData* object using the *Scanpy* package. We have also added the original metadata information to the *AnnData* object (Wolf et al., 2018). The original age group definitions varied between studies, so we applied additional sample exclusion criteria. For the young group, we excluded samples from donors older than 40 years; for the old group, we included samples from donors over 60 years. The final combined dataset contained gene expression data from from 53 young and 50 old donors (Table 1).

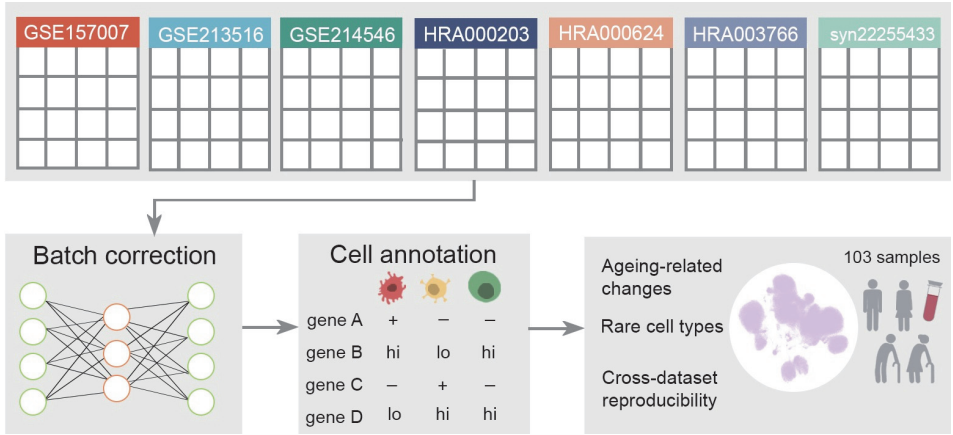


Figure 6. The PBMC ageing atlas study design. The gene expression data and sample metadata were retrieved for seven aging human immune system studies.

5.4 Patient data generation and processing (Papers III, IV)

The $CD8^+$ Temra cells for Paper III were sorted as $CCR7^{lo}$ $CD45RA^{hi}$ cells. The samples were barcoded using TotalSeq-B Hashtags (BioLegend) and combined into two different reactions. For Paper III, Tregs were sorted as $CD4^+CD25^+CD127^{low}$ cells. The cells for both Paper II and III were loaded onto a Chromium controller and cDNA was generated using Chromium Next GEM Single Cell 3' GEM, Library & Gel Bead Kit v3.1 and Chromium Next GEM Chip G Single Cell Kit (both from 10x Genomics) according to the manufacturer's protocol. The resulting gene libraries were sequenced on an Illumina NovaSeq 6000 at Novogene.

The raw sequencing data were processed using the *CellRanger* pipeline with the GRCh38 reference obtained from 10x Genomics. The scRNA-seq alignment and quantification were performed on the Tartu University High-Performance Computing Center Rocket cluster. In Papers I and III, the samples were demultiplexed using the *HTODemux* function from the Seurat package (Stuart et al., 2019).

5.5 Quality control

A rigorous quality control procedure was implemented to ensure that only high-quality cells were retained. Cells were evaluated based on the number of detected UMIs and the percentage of mitochondrial gene expression. This approach ensured the exclusion of potential multiplets and low-quality cells from the analysis. In all studies, plots displaying the quality control metrics were analyzed to determine the appropriate thresholds for excluding low-quality cells. These metrics were assessed for each 10x Genomics reaction to ensure suitable thresholds were selected for every batch.

In Papers I and II, additional quality control steps included computational doublet detection using the *scDblFinder* R package (Germain et al., 2022; Xi & Li, 2021). Ten independent runs were performed to account for the dependence on random initialization in *scDblFinder* results. A cell was marked as a potential doublet if it was flagged in more than half of the runs. High-confidence doublet predictions were then manually reviewed using UMAP and gene expression plots to check for isolated cell clusters, which indicated co-expression of markers from different lineages (e.g., both T and B cell markers). This approach helped minimize false-positive doublet assignments. Such clusters were excluded from further analysis.

5.6 Data integration

Computational methods were applied to correct batch effects and address technical differences between batches. In Papers I, III and IV, the *Seurat* R package was used for data integration (Stuart et al., 2019). Seurat's integration algorithm leverages CCA to align datasets. First, it identifies shared sources of variation between datasets by finding correlated gene expression patterns. Based on these correlations, CCA projects these datasets into a shared low-dimensional space, allowing for identifying cell pairs, or "anchors," that link similar cells across datasets. These anchors are then used to integrate the datasets, correcting for batch effects while preserving biological variation.

In Paper II, data integration was performed using the *scvi-tools* Python package (Lopez et al., 2018). This tool uses a deep neural network to learn a low-dimensional latent representation of scRNA-seq data, capturing the underlying biological variability while correcting for batch effects. Gene expression matrices from seven datasets were used as input, with batch labels and the percentage of mitochondrial gene expression included as covariates to account for technical differences. The resulting low-dimensional representation of the combined dataset was then used for downstream tasks. Additionally, the *Harmony* R package was employed to validate the *scVI* data integration in Paper II, ensuring the robustness of the batch correction process (Korsunsky et al., 2019). *Harmony* performs batch correction by taking cell coordinates from a PCA-reduced space and iteratively adjusting for dataset-specific effects through fuzzy clustering, calculating correction factors based on centroids, and applying a cell-specific correction.

5.7 Dimensionality reduction and clustering

In Paper II, the *Scanpy* Python workflow was utilized for data clustering and visualization. The neighborhood graph was computed using the *scVI* latent representation as the input. The UMAP embeddings were generated using the *umap-learn* implementation, based on the computed neighborhood information (McInnes et al., 2018). The clustering was performed on the *scVI* latent space using Leiden algorithm (Traag et al., 2019). Similar procedures were followed in Papers I, III and IV, utilizing the *Seurat* R package.

5.8 Differential expression analysis

Differential expression analysis was performed using the Wilcoxon rank-sum test. Significance thresholds were set at 0.25 for log-fold change and 0.05 for FDR. In Paper II, the pseudobulk method was used to validate DE results. To this end, the *glmGamPoi* R package was used (Ahlmann-Eltze & Huber, 2021). Pseudobulk profiles for each cell type and sample were created, and profiles with fewer than ten cells were excluded. Pseudobulk tests were conducted for cell types with at least three pseudobulk profiles per age group. g:Profiler was used for functional enrichment in Paper II, and Ingenuity Pathway Analysis was applied in Paper IV (Kolberg et al., 2023).

6 RESULTS

6.1 Comparative transcriptomic analyses of thymocytes using 10x Genomics and Parse Bio scRNA-seq technologies (Paper I)

6.1.1 Parse Bio Exhibits Greater Variability in QC Metrics Between Replicates Compared to 10x Genomics

We used mouse thymus as a model tissue to evaluate the performance of the 10x Genomics and Parse Bio platforms and compare their quality control metrics for capturing single-cell transcriptomes. We utilized thymi from two female C57BL/6N mice (aged two months), with each mouse's (biological replicates A and B) thymic lobes treated technical replicates (1 and 2). After enzymatic digestion, the four samples were split and processed separately using the 10x Genomics and Parse Bio kits. Hereafter, the 10x Genomics samples are referred to as 10x_A_1, 10x_A_2, 10x_B_1, and 10x_B_2. Similarly, Parse Bio samples were designated as parse_A_1, parse_A_2, parse_B_1, and parse_B_2. The 10x Genomics samples were further labelled using cell hashing and subsequently processed using the 10x Genomics microfluidics device. The Parse Bio samples were first processed with the cell fixation kit and the cells barcodes were added using the split-pool barcoding.

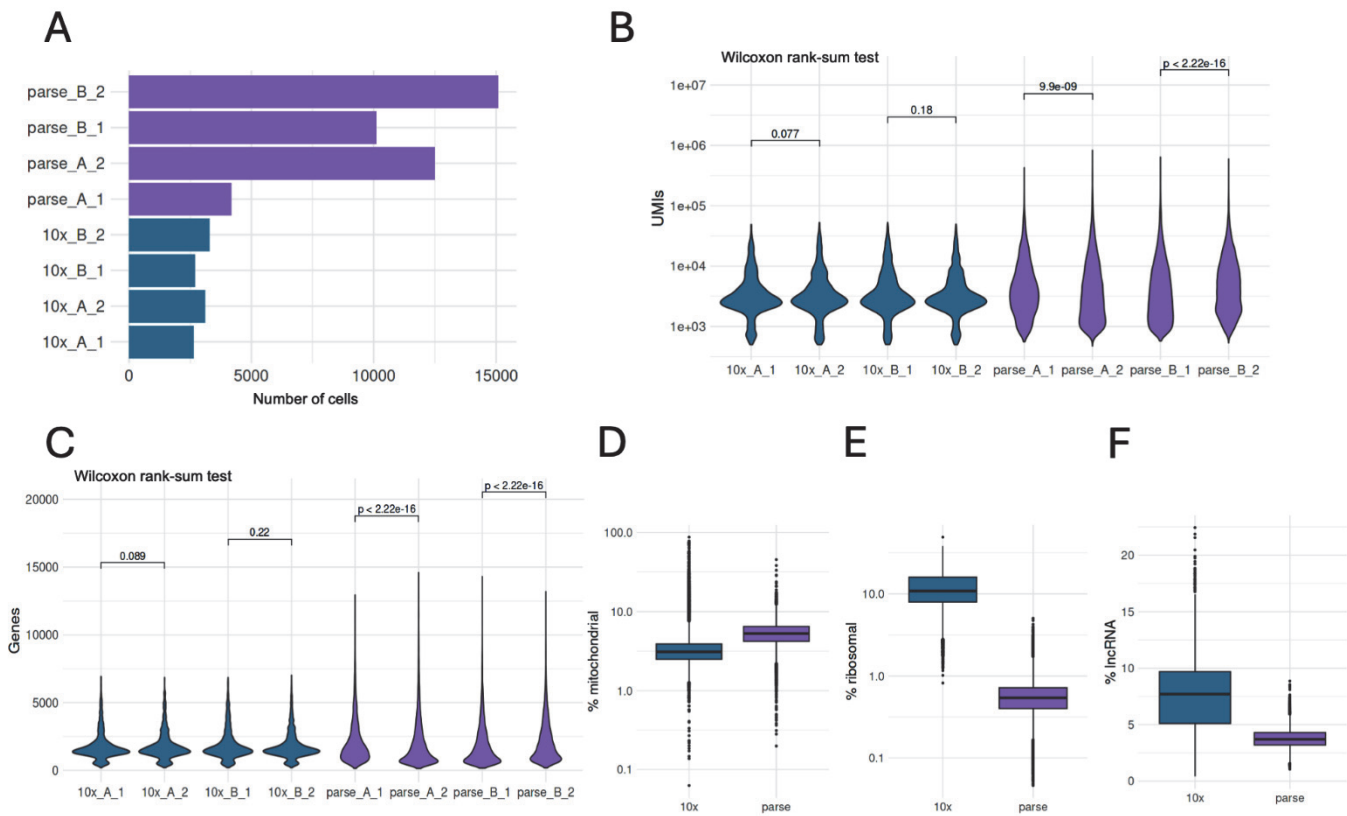


Figure 7. Cell recovery and gene detection. (A) Number of cells recovered from each replicate, (B) UMI counts/cell for each sample, (C) Genes detected/cell in each sample, (D) Percentage of expression mapped to mitochondrial genes, (E) Percentage of expression mapped to ribosomal genes, (F) Percentage of expression mapped to long non-coding RNA genes.

Approximately 3,000 cells per sample were targeted for collection using 10x Genomics, and around 5,000 cells per sample for Parse Bio. To achieve this, ~5,100 cells per sample were loaded for 10x Genomics and ~19,200 cells for Parse Bio. Following sequencing, an average of 3,477 cells per sample was recovered for 10x Genomics, with minimal inter-sample variability and a recovery rate of 56.5% (Figure 7A; Tables S1,S2 in Paper I). In contrast, an average of 10,460 cells per Parse sample was recovered, showing significant inter-sample variability and a recovery rate of 54.4% (Figure 7A; Tables S1,S2 in Paper I). Thus, 10x Genomics exhibited higher cell capture efficiency and more consistent results across samples.

We analyzed UMI and detected gene count distributions per cell, common quality control metrics in scRNA-seq analysis. Parse Bio data showed longer tails in the violin plots compared to 10x Genomics, indicating more outlier cells with high gene expression (Figures 7B, C). When comparing technical replicates, 10x Genomics displayed no significant differences in UMI and gene counts, while Parse Bio showed notable variation. Interestingly, the average percentage of mitochondrial reads was slightly lower in 10x Genomics than in Parse Bio (Figure 7D). However, 10x Genomics had a much higher fraction of ribosomal gene counts and a greater proportion of lncRNA counts (Figures 7E,F). Taken together, these results show differences in technical metrics between 10x Genomics and Parse Bio platforms, as well as highlight variability between samples and replicates in Parse Bio data.

6.1.2 Parse Bio and 10x Genomics are enriched for distinct sets of genes

We then examined the differences in captured genes between the datasets. Despite similar sequencing depth, the Parse Bio dataset detected a larger number of unique genes. Of the 16,390 genes shared between both platforms, 14,731 were unique to Parse Bio, while only 578 were unique to 10x Genomics (Figure 8A). We further investigated this difference in highly expressed genes, finding that only 364 of the top 1,000 most expressed genes overlapped between the two datasets (Figure 8B). In the 10x Genomics data, the most highly expressed gene was the lncRNA *Malat1*, while in Parse Bio data, the top gene was the ribosomal RNA *Rn18s-rs5*, with *Malat1* ranking sixth (Figures 8C,D). Both datasets showed high expression of mitochondrial genes, but these varied: 10x Genomics data featured *mt-Co3, mt-Atp6, mt-Co1,* and *mt-Co2*, while Parse Bio data included *mt-Rnr1* and *mt-Rnr2*. Interestingly, both platforms detected high expression of *Themis*, a gene involved in T cell selection.

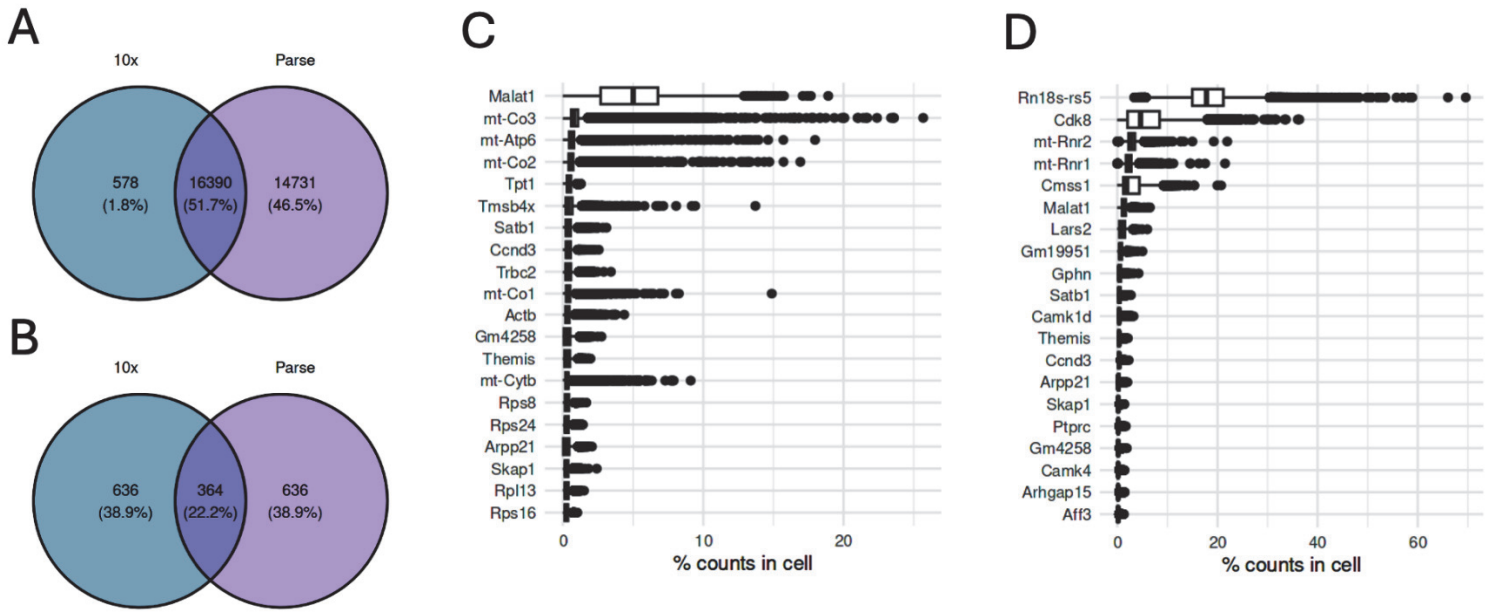


Figure 8. Gene detection in 10x and Parse datasets. (A) Overlap of all genes detected in both libraries, (B) Overlap of top 1000 expressed genes in both libraries, (C) Top expressed genes in 10x Genomics library, (D) Top expressed genes in Parse Bio library.

6.1.3 Computational analysis identifies more doublets in Parse Bio data

Doublets or multiplets arise when two or more cells are captured within a single droplet, resulting in mixed or hybrid transcriptomic profiles. Including these aggregates in downstream analysis can lead to incorrect identification of rare cell types, intermediate states, or specific transcriptomic signatures. In our study, we used cell hashing for the 10x Genomics samples, allowing us to distinguish doublets from single cells. Cells positive for more than one sample hashtag were identified as doublets and excluded from further analysis. Doublets accounted for 14% of the 10x Genomics dataset (Table 2). In contrast, the Parse Bio toolkit lacks a comparable method for distinguishing doublets, as it does not use hashtags for sample multiplexing. To address this, we applied the computational doublet detection method in *scDblFinder*, which identifies doublets by comparing cells against computationally simulated doublet profiles (Germain et al., 2022; Xi & Li, 2021). Notably, doublets comprised 31% of the Parse Bio library (Table 2).

Table 2. Doublet detection results in 10x Genomics and Parse Bio.

	10x Genomics	**Parse Bio**
Total cells	13820	40415 (post QC)
HTO-negative	259	-
Singlets	11688	27990
Doublets	1873	12425
% Doublets	14%	31%

6.1.4 Parse Bio data has a considerable batch effect

We next investigated whether batch effects contributed to the significant variability in detected gene and UMI counts across Parse Bio samples. Such technical effects could interfere with downstream analyses, like cell type annotation or differential expression analysis. We applied a standard scRNA-seq analysis workflow, including normalization, dimensionality reduction, and clustering, separately on the 10x Genomics and Parse Bio datasets. To assess batch effects, we examined the alignment of samples on the UMAP plot. In the 10x Genomics data, cells from both biological and technical replicates were well-aligned (Figure 9A). However, in the Parse Bio data, the replicates formed distinct cell communities (Figure 9B), which could lead to artificial clusters being mistaken for genuine biological states. To address this, we performed batch effect correction to remove variability between the Parse Bio replicates, and after data integration, the replicates were well-aligned (Figure 9C).

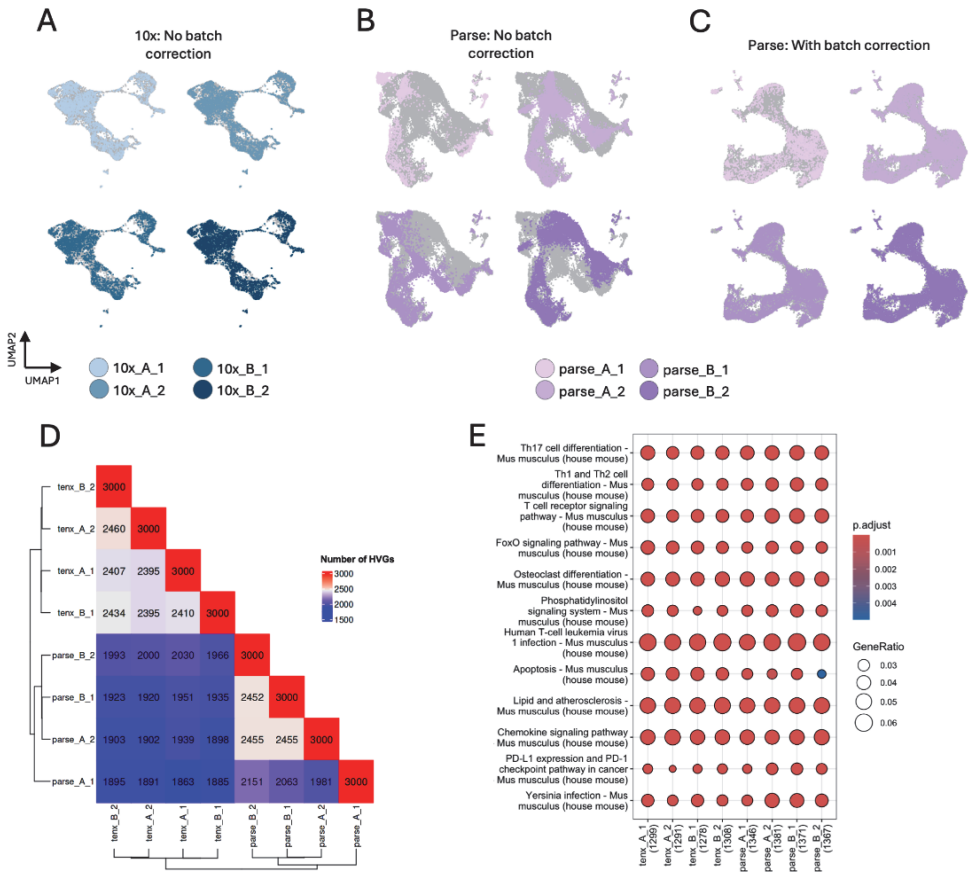


Figure 9. Batch effect exploration. (A) UMAP representation of 10x Genomics data without batch effect correction, (B) UMAP representation of Parse Bio data without batch effect correction, (C) UMAP representation of Parse Bio data with batch effect correction, (D) Overlap in HVGs between samples, (E) KEGG pathways enriched in HVGs for each sample.

We further explored the differences between replicates within each dataset. Most scRNA-seq workflows reduce data dimensionality by selecting highly variable genes (HVGs) in each sample. After performing this step, we compared the selected HVGs across replicates and platforms. The overlap of HVGs between the two technologies was lower than the overlap within each platform (Figure 9D). In the 10x Genomics samples, pairwise HVG overlaps were consistent, while one Parse Bio sample (parse_A_1) showed fewer shared HVGs with other Parse Bio samples. We then conducted KEGG enrichment analysis on the HVGs from each sample. Gene ontology analysis revealed similar pathway representation across both 10x Genomics and Parse Bio, with enrichment for the same biological processes (Figure 9E). The only notable difference was in the "Apoptosis" pathway, which had a slightly lower score in Parse Bio samples compared to 10x Genomics.

6.1.5 10x Genomics is more sensitive to lineage-defining genes than Parse Bio

Single-cell datasets often contain a variety of cell types and subsets, making detailed cell type and state annotation a crucial step in scRNA-seq analysis. To begin, we analyzed the expression levels of *Cd3d*, *Cd4*, and *Cd8a* in the unsupervised clustering of both 10x Genomics and Parse Bio data, as these markers are key for T cell identification (Figures 10A,B). A striking difference in marker gene expression was observed between the datasets (Figure 10C). In the 10x Genomics data, *Cd3d* was consistently and highly expressed across nearly all clusters, while in the Parse Bio data, *Cd3d* expression was undetectable in any cluster. A similar pattern was noted for *Cd3g*, which appeared in fewer clusters and at lower levels in the Parse Bio data compared to 10x Genomics.

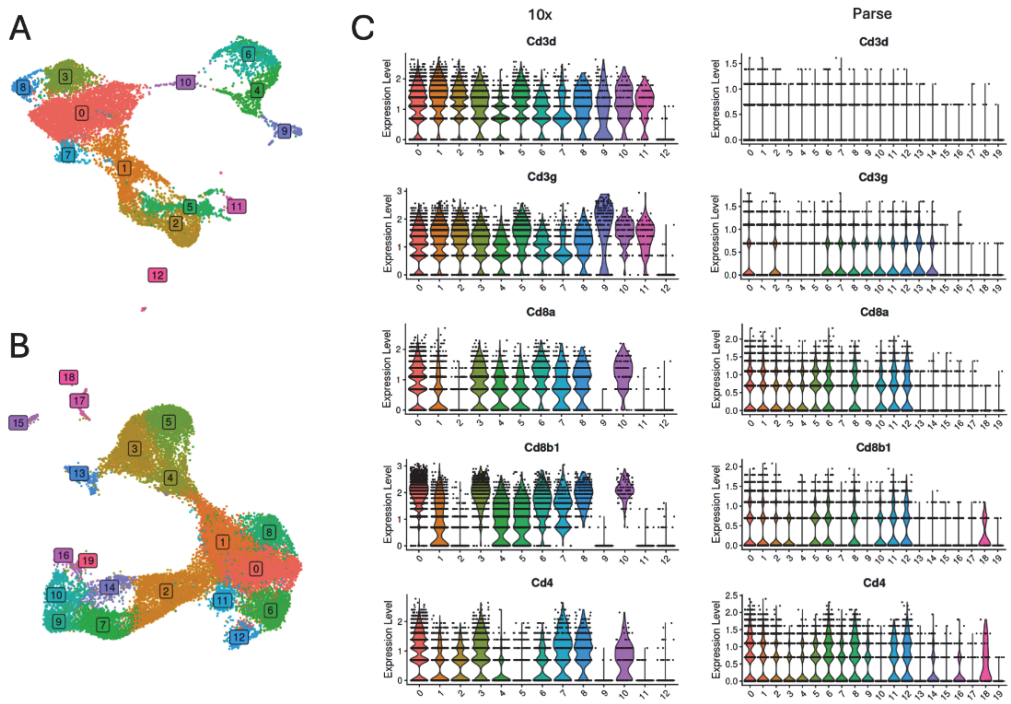


Figure 10. T marker gene expression in 10x Genomics data and Parse Bio. UMAP representation and clustering of (A) 10x Genomics and (B) Parse Bio data. (C) Comparison of T marker gene expression between 10x Genomics and Parse Bio.

6.1.6 10x Genomics resolves thymocyte populations better than Parse Bio

We next grouped the clusters according to the stages of thymic differentiation based on the expression of *Cd4* and *Cd8a* genes (Figures 11A-D). These are represented by double-negative (DN; $Cd4^{-}Cd8^{-}$), double-positive (DP; $Cd4^{+}Cd8^{+}$), and single-positive (SP; $Cd4^{+}Cd8^{-}$ and $Cd4^{-}Cd8^{+}$) cell populations. In the 10x Genomics data, we identified a $Cd3d^{+}Cd4^{-}Cd8a^{-}$ cluster that could be annotated as DN thymocytes. This cluster had a high *Il2ra* (CD25) expression, a marker of DN cells. Next, we analyzed the Parse Bio data and found that major thymocyte populations were congruent with the 10x Genomics data.

We then annotated all $Cd3d^{+}Cd4^{+}Cd8a^{+}$ clusters in the 10x Genomics data. The DP group was subdivided into proliferating (DP-P) and recombining (DP-R) populations based on the expression of genes associated with the cell cycle (*Mki67*, *Top2a*) and V(D)J recombination (*Rag1*), respectively. The third stage in the differentiation was a DP undergoing selection cluster (DP-S) representing thymocytes in the post-recombination positive selection stage (*Tox*). In addition, we observed a DP apoptotic cluster (DP-A) that was enriched in mitochondrial gene expression. We used the same markers to annotate DP-P, DP-R, and DP-S thymocyte subsets in the Parse Bio data. However, we did not observe a DP-A cluster in the Parse Bio dataset, as there were no cell communities with a pronounced increase in mitochondrial gene expression. The SP clusters represent the end stage of thymocyte development and have mutually exclusive *Cd4* and *Cd8a* expression patterns. In 10x Genomics data, we identified the $Cd3d^{+}Cd4^{+}Cd8^{-}$ SP-CD4 and $Cd3d^{+}Cd4^{-}Cd8a^{+}$ SP-CD8 populations, and these were also distinguishable in the Parse Bio data.

The remaining clusters were identified as APC and other $CD3^{+}$ cells in 10x Genomics data based on distinct *Cd3d* expression level and MHC-II related genes. However, we could not identify these clusters in Parse Bio solely by *Cd3*-related gene expression. We annotated the APC cluster in Parse Bio by expression of *Cd4*, *Cd8a*, and MHC-II related genes but other $CD3^{+}$ cells could not be annotated due to a lack of distinct *Cd3d* or *Cd3g* gene expression in the clusters. We wondered whether we could leverage the 10x Genomics data annotations to identify the similar clusters and proportions in Parse Bio data. For this, we trained a classifier on 10x Genomics gene expression values and annotations and obtained predictions for the Parse Bio data. Based on this model, we could accurately annotate the minor thymocyte population, other $CD3^{+}$, in Parse Bio data.

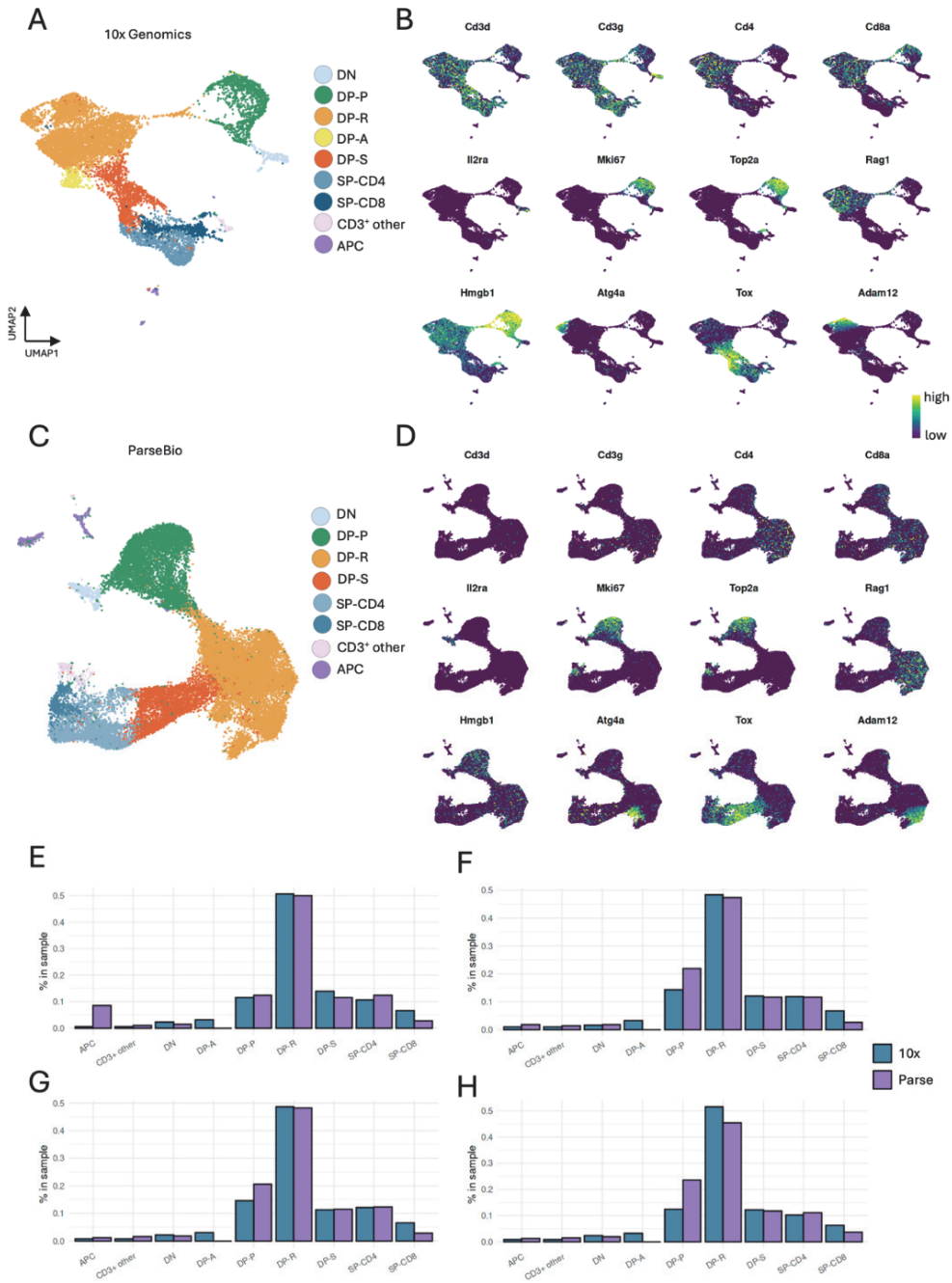


Figure 11. Cell type annotation. UMAP (A) and marker genes (B) of 10x Genomics data, UMAP (C) and marker genes (D) of Parse Bio data. Proportion of cells in 10x Genomics (blue) and Parse Bio (purple) in replicates (E) A_1, (F) A_2, (G) B_1, and (H) B_2.

6.1.7 Differences in cell type proportions between 10x Genomics and Parse Bio

Next, we studied the proportion of thymocyte cell types identified using each technology. The DP-R, DP-S, and SP-CD4 proportions were relatively consistent between the technologies and replicates (Figures 11E-H), indicating similar detection efficiency for these subpopulations. However, we noticed significantly fewer SP *Cd8*$^{+}$ cells in the Parse Bio data across all replicates compared to the 10x Genomics data. Conversely, Parse Bio detected a higher proportion of DP-P cells than 10x Genomics. Taken together, these findings highlight differences in cell subpopulation detection between the 10x Genomics and Parse Bio kits, which may reflect platform-specific biases in capturing certain cell types.

6.1.8 RNA velocity shows agreement in thymocyte trajectory between the two datasets

Thymocyte turnover is characterized by a continuous influx of progenitor cells, their differentiation or apoptosis, and eventual exit from the thymus, providing an ideal setting for benchmarking computational tools to study developmental dynamics. We conducted RNA velocity analysis, leveraging the ratio of spliced to unspliced transcripts to assess whether both technologies capture this dynamic information. In 10x Genomics data, 58% of transcripts were spliced and 39% were unspliced, with 3% categorized as ambiguous (Figure 12A). In contrast, Parse Bio data exhibited an inverted ratio, with 47% spliced and 53% unspliced transcripts (Figure 12C). We further assessed whether the velocity vector fields and latent time reflected thymocyte developmental trajectories, finding that DP-R cells exhibited the highest latent time values in both datasets (Figure 12B, D). The velocity vectors were directed toward DP cells and away from SP and DP-P populations, consistent with expected developmental transitions.

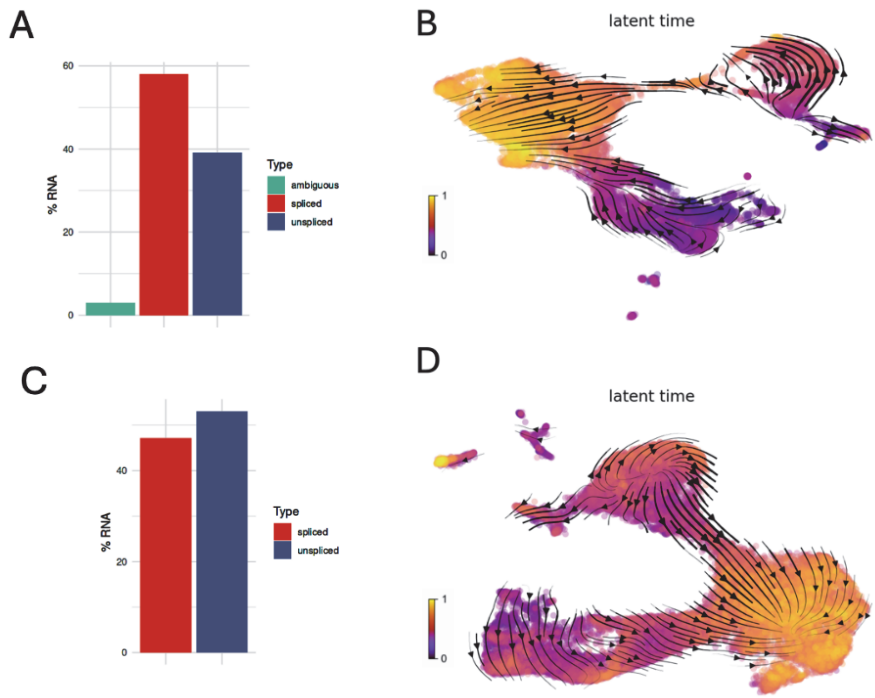


Figure 12. RNA velocity analysis. Detected RNA splicing in (A) 10x Genomics and (C) Parse Bio data. RNA velocity vectors and pseudotime in (B) 10x Genomics and (D) Parse Bio data.

6.2 PBMC ageing scRNA-seq data integration for immune biomarker validation (Paper II)

6.2.1 The PBMC aging atlas enables comprehensive immune cell type annotation

The scRNA-seq data from seven previously published aging studies was collected and processed in a standardized manner. To address technical variations across datasets, computational integration was conducted using *scVI*. Altogether, the combined dataset included gene expression profiles from 53 young individuals (22 females, 31 males) and 50 aged individuals (22 females, 28 males). After applying quality control measures, the cell atlas comprised transcriptomic data for 14,461 genes from 1,047,793 high-quality PBMCs (Figure 13A).

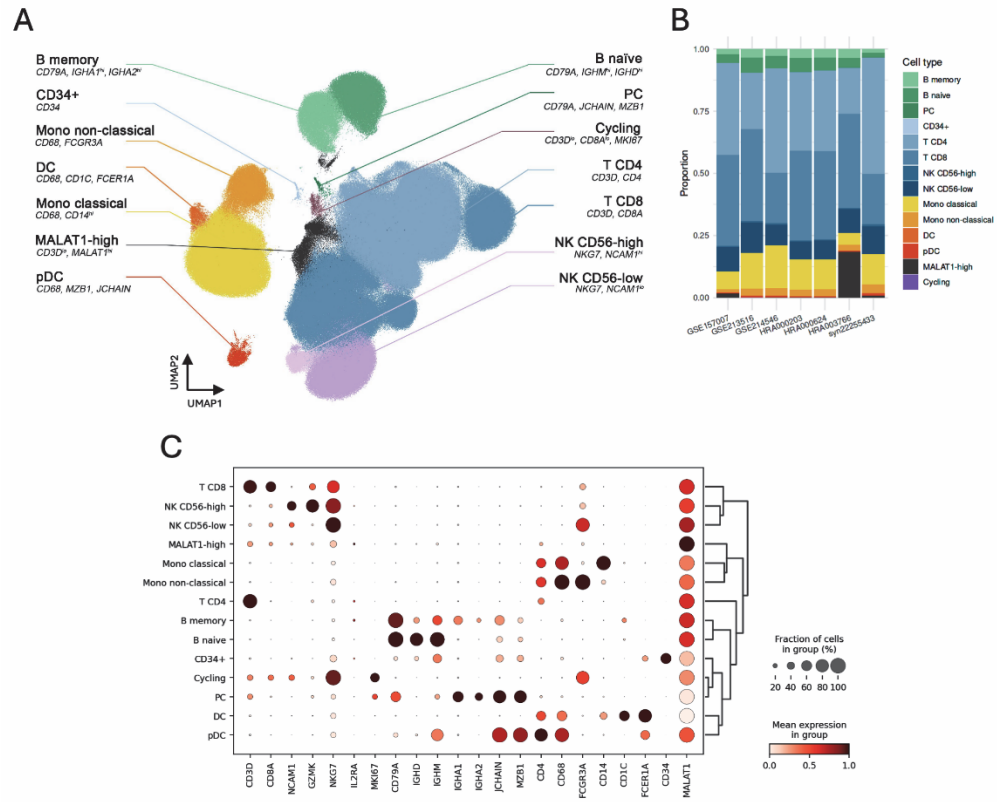


Figure 13. Single-cell PBMC immune atlas. (A) UMAP representation of the 1,047,793 PBMCs from young and old subjects identified by integrating seven scRNA-seq datasets. (B) Bar plots showing proportions of the major PBMC subpopulations in each dataset. (C) Gene expression of selected marker genes used to identify the major PBMC populations.

Unlike flow cytometry, which relies on well-established cell surface markers to identify immune cell subsets in PBMCs, the original studies used inconsistent cell type definitions due to the lack of established transcriptomic markers in the scRNA-seq field. Our large-scale integrated PBMC dataset allowed us to harmonize cell type definitions across the studies using a uniform set of annotation markers (Figure 13C). We identified clusters of naïve (*IGMH, IGHD*) and memory (*IGHA1*, *IGHA2*) B cells, along with a small population of plasma cells (*CD79A, JCHAIN, MZB1*). Myeloid cells included classical (*CD68*, *CD14*) and non-classical (*CD68, FCGR3A*) monocytes, DCs (*CD68, CD1C, FCER1A*), and pDCs (*CD68, MZB1, JCHAIN*). The NK cell compartment comprised $CD56^{hi}$ and $CD56^{lo}$ NK cells. Among T cells, we identified large populations of $CD4^{+}$ and $CD8^{+}$ T cells. Other populations included $CD34^{+}$ circulating progenitors, cycling cells (*MKI67*), and cells with high lncRNA *MALAT1* expression.

6.2.2 Pronounced cell abundance differences between the datasets

With the consistent cell type markers across all datasets in our age-related PBMC atlas, we investigated whether the annotated cell subsets were equally represented in each study. We found significant variations in major immune compartments across the seven datasets (Figure 13B). The proportion of $CD8^{+}$ T cells was lower in GSE214546 and syn22255433, while syn22255433 and GSE157007 had fewer B cells compared to other datasets. Additionally, GSE157007 and HRA003766 had fewer monocytes than the other five datasets. Interestingly, HRA000203 and HRA000624 showed only minor differences in cell type proportions despite having different age group definitions. Finally, the $MALAT1^{hi}$ cell population was unevenly distributed across the studies, with HRA003766 showing the highest abundance of these cells.

6.2.3 B cell subsets show stability with aging

The 77,378 B cells were subsetted and re-analyzed for more detailed characterization. This analysis identified three B cell subsets: atypical B cells, as well as naive and memory B cells (Figures 14A, B). ABCs, also known as aging-associated B cells, are a distinct subset of B cells that increase in frequency with age and in certain chronic inflammatory conditions. The proportion of ABCs was compared between young and older individuals in each dataset to evaluate whether their levels consistently change with aging. Notably, no significant difference in ABC proportions was observed between the two age groups in our analysis (Figure 14C).

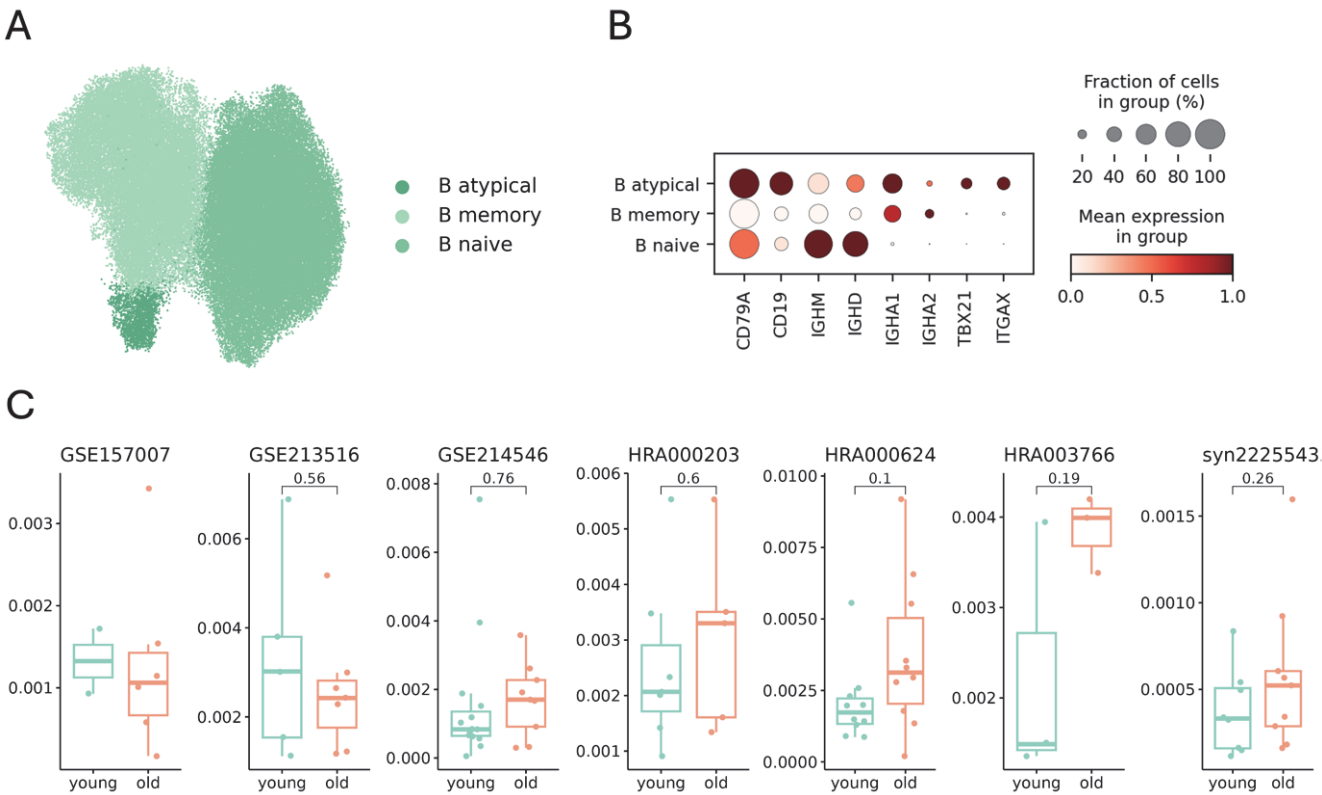


Figure 14. Integrated analysis of young and old B cells. (A) UMAP representation of the 77,378 B cells from young and old subjects identified by reanalysis of B cell subsets. (B) Dotplot of selected marker genes used to identify the B cell subsets. (C) ABC proportions in young and old groups in each dataset. The numbers denote a p-value from a t-test. The test was not performed for GSE157007 due to the small young group size.

6.2.4 *MALAT1^{hi}* cells are highly heterogeneous

The analysis of PBMCs identified a distinct *MALAT1^{hi}* cell population located outside the main PBMC clusters (Figures 13A,C). *MALAT1* has been suggested as a marker of exhausted T cells, and this *MALAT1^{hi}* population has been linked to aging and frailty (Luo et al., 2022). To conduct a deeper analysis of these cells, the 14,007 cells from all datasets were subsetted and reanalyzed (Figure 15A). The HRA003766 contributed the majority of the cells (79.59%); other datasets had lower proportions of *MALAT1^{hi}* cells (GSE157007, 7.22%; syn22255433, 6.43%; GSE214546, 4.6%; HRA000624, 0.78%; HRA000203, 0.52%; GSE213516, 0.86%).

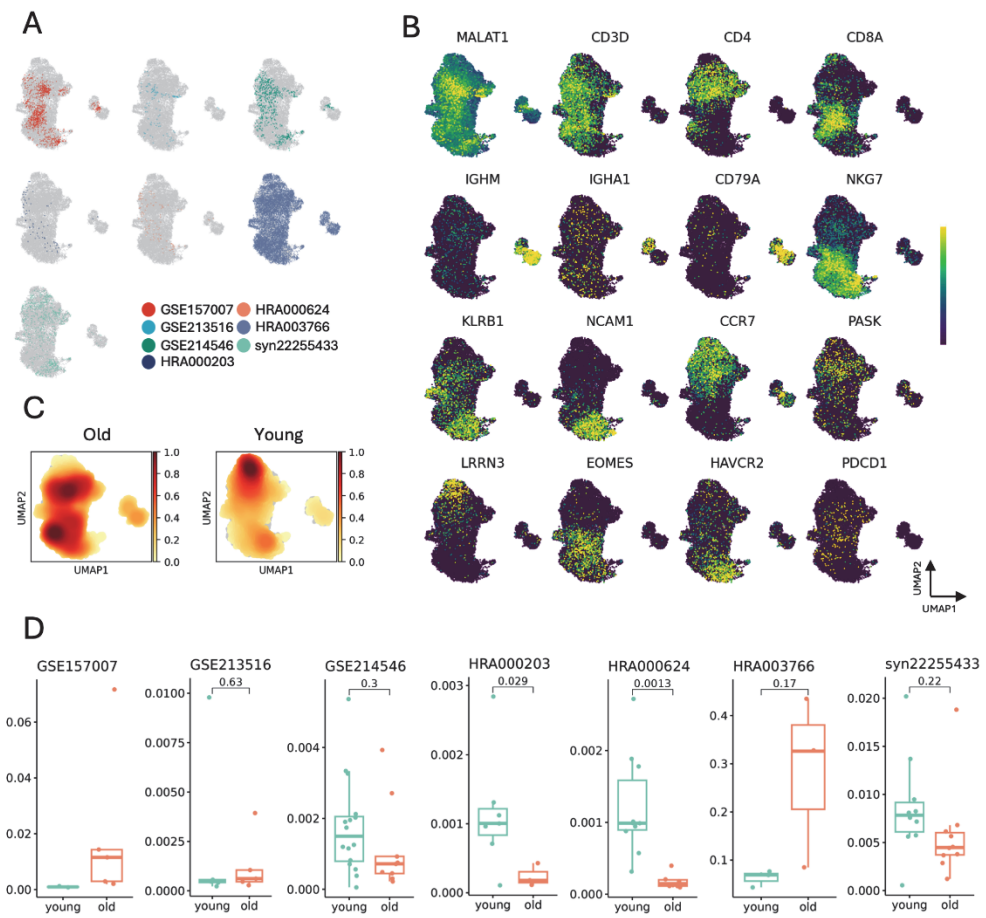


Figure 15. Reanalysis of *MALAT1^{hi}* cells. (A) UMAP representation of 14,007 *MALAT1^{hi}* cells colored by the dataset of origin. (B) Gene expression of selected marker genes. (C) Gaussian kernel density estimation of young and old groups density in the UMAP space. (D) Cell type abundance comparison between young and old groups. The numbers denote a p-value from a t-test. The test was not performed for GSE157007 due to the small young group size.

MALAT1^{hi} subset appeared highly heterogeneous in the expression of common T cell type markers (Figure 15B). A distinct *CD4*$^{+}$ T cell population co-expressed the Tn markers *CCR7* and *LEF1*, while other cells expressed *CD8A*, along with *EOMES* and *NKG7*, typical of terminal CD8^{+} Tem cells. An NK-like population with low *CD3D* expression but expressing cytotoxicity genes (*NKG7*, *KLRB1, NCAM1*) was also identified, along with a minor population expressing B cell markers. This suggests the *MALAT1*^{hi} cells contain both naïve-like and memory-like T cell subsets, NK (or NKT) cells, and B cells.

The abundance of *MALAT1*^{hi} cells was then examined in both young and old groups (Figure 15C). Interestingly, *CD8A*$^{+}$ cells were predominantly found in older individuals, while *CD4*$^{+}$ cells were mainly derived from younger donors, indicating differential *MALAT1* expression in *CD4*$^{+}$ and *CD8*$^{+}$ T cell subsets across age groups. Finally, the abundance of *MALAT1*^{hi} cells was compared between age groups (Figure 15D). Significantly more of these cells were found in young individuals in datasets HRA000203 and HRA000624, while no statistically significant difference was observed in the other datasets. However, a higher number of *MALAT1*^{hi} cells was noted in GSE157007 and HRA003766.

6.2.5 Reference CITE-seq atlas enables T cell marker identification

We next focused on the T cell compartment in the integrated atlas due to its significant shifts with aging, as both CD4^{+} and CD8^{+} T cell subsets undergo major rearrangements, including a decline in Tn numbers and an increase in Tem populations. A common challenge in single-cell T cell annotations is distinguishing between CD4^{+} and CD8^{+} Tn and central memory (Tcm) cells, given their similar gene expression profiles. For instance, while Zheng et al. annotated CD4^{+} Tn and Tcm cells based on *CCR7* and *CD69* expression, they did not differentiate CD8^{+} Tcm cells. Conversely, Luo et al. identified CD4^{+} and CD8^{+} Tcm populations but were unable to separate CD4^{+} and CD8^{+} Tn cells. These discrepancies in Tn and Tcm classifications, along with broader challenges in T cell identification, prompted us to search for more reliable transcriptomic markers to clearly differentiate T cell populations.

To resolve the challenges in distinguishing Tn and Tcm subsets, we searched for transcriptomic markers capable of reliably separating these populations. We utilized a CITE-seq reference dataset comprising 73,060 T cells from eight donors involved in an HIV vaccine trial (Elizaga et al., 2018; S. S. Li et al., 2017). This dataset included surface protein antibodies detecting CD45RA and CD45RO, markers commonly used in flow cytometry to differentiate between Tn and Tcm subsets. To reliably separate major *CD4*$^{+}$ and *CD8*$^{+}$ T cell subsets, including Tn and Tcm, we applied the *totalVI* probabilistic model, which integrates transcriptomic and surface protein data for dimensionality reduction and clustering (Gayoso et al., 2021). Following UMAP analysis, we found that the resulting cell clusters aligned with their original cell type annotations (Figures 16A,B).

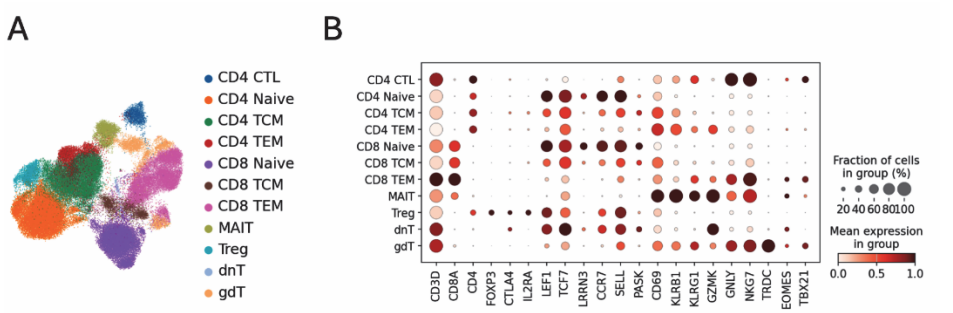


Figure 16. T cell markers from CITE-seq reference. (A) UMAP of T cells from the reference CITE-seq dataset used for T marker validation. (B) Reference T subset marker genes.

We observed that *CD4*$^+$ and *CD8*$^+$ Tn cells expressed the naïve cell markers *LEF1*, *TCF7*, *CCR7*, and *SELL* at higher levels than Tcm cells. Notably, *LRRN3* was expressed in *CD4*$^+$ and *CD8*$^+$ Tn cells but found in only a small fraction of Tcm cells, while the *CD69* gene showed higher expression in Tcm cells. Additionally, *PASK* was more highly expressed in *CD8*$^+$ Tn cells than in *CD4*$^+$ Tn cells. We also found that *KLRG1, GZMK, GNLY,* and *NKG7* were more abundantly expressed in *CD8*$^+$ Tem cells compared to Tcm cells. The transcription factors *EOMES* and *TBX21* were linked to T cell cytotoxicity, being present in both *CD4*$^+$ CTLs and *CD8*$^+$ Tem cells.

In summary, our re-analysis of marker genes demonstrated that *CD4*$^+$ Tn cells can be distinguished by high expression levels of *LEF1*, *TCF7*, *CCR7*, *SELL*, and *LRRN3*. While *CD4*$^+$ Tcm cells also express these markers, they show lower levels of *LRRN3* and are positive for the *PASK* gene. The same markers are applicable for identifying *CD8*$^+$ Tn and Tcm cells, with the key distinction that *PASK* is expressed at significantly higher levels in the *CD8A*$^+$ Tn population compared to *CD4*$^+$ Tn cells. Additionally, both *CD4*$^+$ and *CD8*$^+$ Tcm cells are characterized by the presence of the *CD69* marker.

6.2.6 T cell re-annotation reveals between-study differences

Using this comprehensive set of T cell markers, we reannotated the T cell subsets, which represent the largest populations in the aging atlas. We subset 672,765 T cells and performed the *scVI* data integration workflow, followed by dimensionality reduction and clustering. For the *CD4*$^+$ T cell populations, we identified *CD4*$^+$ Tn (*LRRN3*$^+$, *CCR7*^{hi}), Tcm (*LRRN3*$^-$, *PASK*$^+$), Tem (*KLRG1*$^+$, *PASK*^{lo}), and *CD4*$^+$ cytotoxic lymphocytes (CD4 CTLs; *GNLY*^{hi}, *NKG7*^{hi}). The *CD4*$^+$ T cell population also included regulatory T cells (Treg), which were characterized by high expression of *FOXP3, CTLA4*, and *IL2RA* (Figure 17A,C). Within the *CD8*$^+$ T cells, we identified *CD8*$^+$ Tn (*LRRN3*$^+$, *CCR7*^{hi}, *PASK*^{hi}), Tcm (*CCR7*^{lo}, *PASK*^{lo}), and Tem (*GNLY*^{hi}, *NKG7*^{hi}). Additionally, we annotated populations of gamma-delta T cells (Tgd; *TRDC*^{hi}), mucosal-associated invariant T cells (MAIT;

SLC4A10^{hi}), and natural killer T cells (NKT; *CD3D*^{lo}, *CD8*^{lo}, *NCAM1*^{lo}). We also discovered a minor subset of double-negative T cells (Tdn), which lack expression of both *CD4* and *CD8A*. Intriguingly, we identified a distinct cluster of T cells (Tribo), defined by the enrichment of ribosomal genes (*RPL10, RPL8, RPL11,* and *RPS13*) as well as the *FAU* gene (Ubiquitin-Like and Ribosomal Protein S30 Fusion). This Tribo population was virtually absent in 3' datasets, including GSE213516 and HRA003766, and was only a minor fraction (0.14%) of T cells in the GSE214546 dataset.

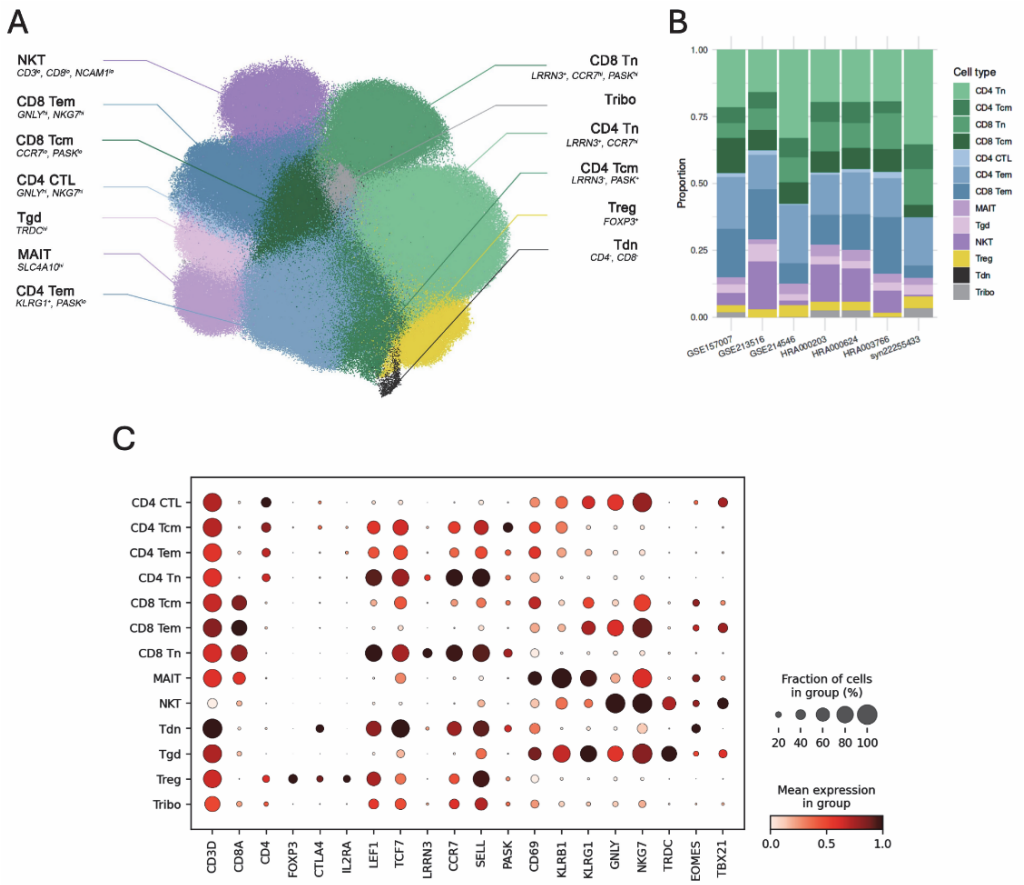


Figure 17. Integrated analysis of young and old T cells. (A) UMAP representation of the 672,765 T cells from young and old subjects identified by reanalysis of T cell subsets. (B) Bar plots showing proportions of T subpopulations in each dataset. (C) Dotplot of selected marker genes used to identify the T cell subsets.

We then examined whether the annotated T cell populations were evenly distributed across the datasets (Figure 17B). Notably, the GSE214546 and syn22255433 datasets contained a higher proportion of *CD4*$^+$ Tn cells compared to the others. We also found that the syn22255433, GSE157007, and GSE214546 datasets had few NKT and *CD4*$^+$ CTL cells, while the proportions in the other four datasets were more comparable. As in the case of the total PBMCs, HRA000203 and HRA000624 were similar in their T cell subset proportions.

6.2.7 Multi-cohort significant changes in abundance are detected in few cell types

Using harmonized cell type annotations, we examined age-related changes in cellular abundance for each cell type across datasets (Figure 18; Figures S19-S15 in Paper II). We conducted statistical tests to compare the proportion of cells between young and old individuals in all datasets except GSE157007, which included only two samples in the young group. Our analysis reaffirmed the major shifts seen in immune cell aging. Notably, *CD8*$^+$ Tn cells consistently declined with age across all seven datasets, with statistically significant reductions in five (GSE213516, GSE214546, HRA000203, HRA000624, and syn22255433; $p < 0.05$). A similar reduction was seen in the GSE157007 and HRA003766 datasets, although the latter did not reach statistical significance ($p = 0.089$). *CD4*$^+$ Tn cells were also significantly lower in three datasets (HRA000203, HRA000624, and HRA003766; $p < 0.05$). Additionally, we observed a consistent decline in MAIT cells among older donors, with significant changes in three datasets (GSE214546, HRA000624, and syn22255433; $p < 0.05$). Non-classical monocytes were more abundant in older donors across datasets, though this increase was borderline significant except in HRA000624, where it reached statistical significance ($p < 0.05$). Other cell populations did not show consistent trends. Overall, considerable variability in cell type prevalence was observed across the datasets.

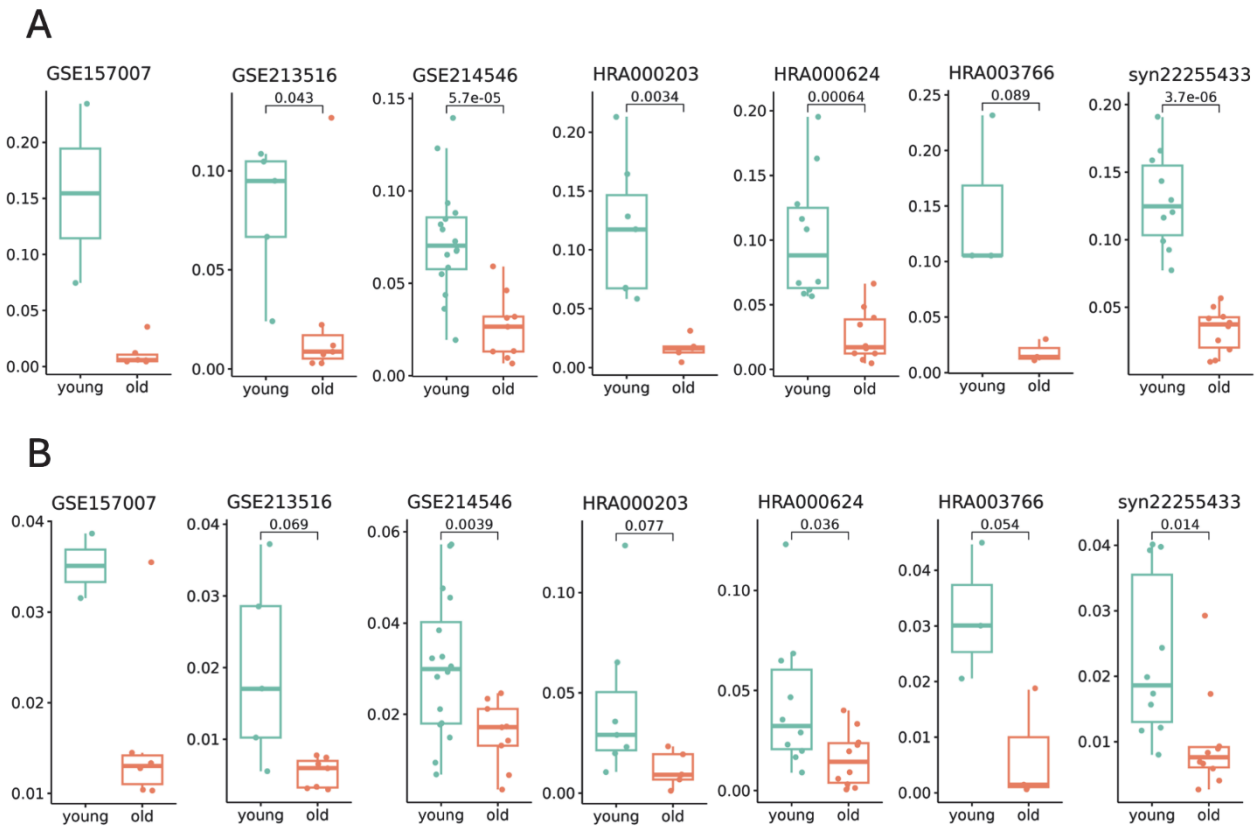


Figure 18. The cell type abundance differences in ageing. The proportion of (A) *CD8*$^{+}$ Tn and (B) MAIT cells in age groups and datasets. The numbers denote a p-value from a t-test. The test was not performed for GSE157007 due to the small young group size.

6.2.8 Differential expression changes are mostly cohort-specific

We then compared gene expression in immune cells between young and old groups within each dataset to evaluate age-related transcriptomic changes. Sets of differentially expressed genes and their associated log fold changes were obtained, followed by hierarchical clustering to determine whether cell types from different datasets would group together. Interestingly, clustering was primarily organized by dataset rather than by cell type, with the exception of HRA000203 and HRA000624 (Figure 19). These findings suggest that aging-related transcriptomic differences were less distinct than the variations observed between datasets.

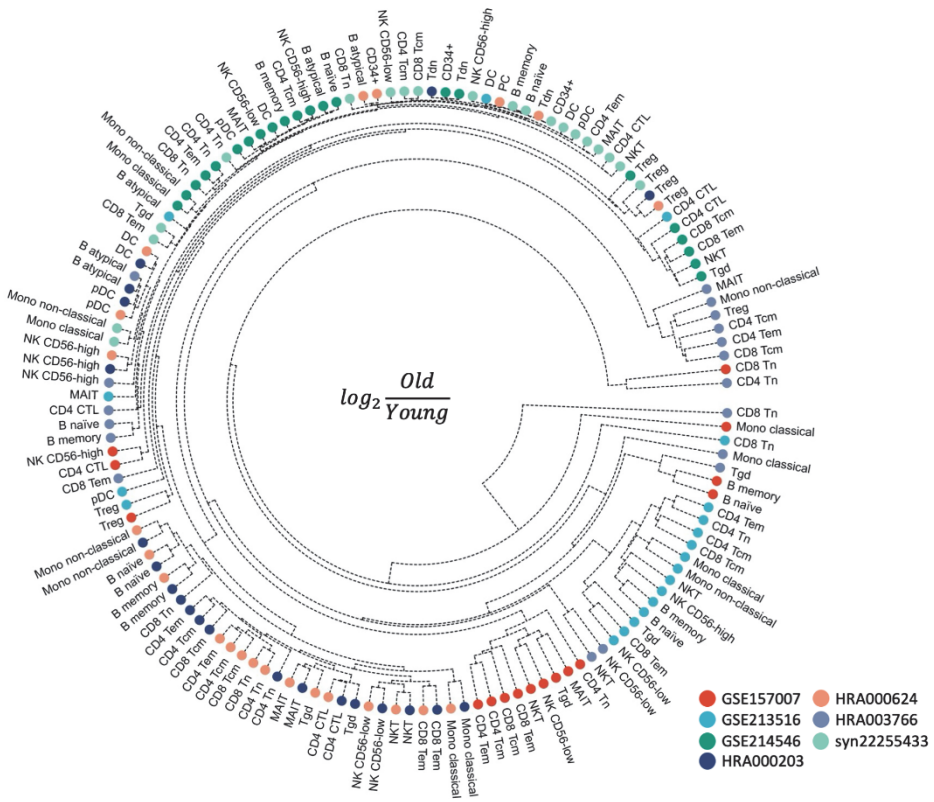


Figure 19. Differential expression in aging. The DE tests between young and old groups were performed for each cell type-dataset combination. The resulting fold change values were used for hierarchical clustering. The leaf nodes (cell types) are colored by the dataset.

We then sought to identify genes that consistently show differential expression with age across multiple datasets using a Wilcoxon signed-rank test. Each dataset was analyzed separately to pinpoint differentially expressed genes within each cell population. To generate high-confidence gene lists and minimize potential

false positives, we focused on genes differentially expressed in at least four datasets. Concentrating on the *CD8*$^+$ Tn population, which showed the most consistent age-related changes, we identified a gene signature of 43 genes most consistently upregulated in older individuals (Table S16 in Paper II). Enrichment analysis yielded GO terms via g:Profiler, revealing high-confidence driver terms such as "immune system process" (GO:0002376), "negative regulation of viral process" (GO:0048525), "response to external stimulus" (GO:0009605), "cell activation" (GO:0001775), and "cell death" (GO:0008219).

To address the Wilcoxon signed-rank test's susceptibility to false positives, we validated the differential expression findings using a pseudobulk approach. While this method yielded fewer significant results across cell types, findings for *CD8*$^+$ Tn cells remained consistent (Table S16 in Paper II). The g:Profiler results included terms such as the "granzyme-mediated programmed cell death signaling pathway" (GO:0140507), "cell activation" (GO:0001775), "cell killing" (GO:0001906), "defense response" (GO:0006952), and "apoptotic process" (GO:0006915). Shared genes from both differential expression analyses included inflammation and effector function markers such as *PTGER2*, *CLIC1*, *S100A4*, *PRF1*, *GZMA*, and *CCL5*. Together, these findings highlighted a consistent upregulation of multiple genes and associated biological processes in *CD8*$^+$ Tn cells with aging.

6.3 CD8$^+$ Temra cells in CMV and aging (Paper III)

6.3.1 Sorted CCR7^{lo} CD45RA^{hi} Temra cells are heterogenous

The late differentiated CD8$^+$ T cells are known to be a heterogeneous population. We conducted scRNA-seq analysis on these cells to gain deeper insights into their heterogeneity. The CD8$^+$ Temra population, defined as CCR7^{lo}CD45RA^{hi} cells, was sorted from a total of eleven individuals, including four young and seven old, with four having high and seven having low CMV antibody levels. These cells were then sequenced using the 10x Genomics scRNA-seq platform. Following quality control, transcriptomic profiles from 20,945 high-quality cells were analyzed. After clustering, ten distinct clusters were identified based on gene expression patterns and annotated using marker genes (Figures 20A,B). Single-cell clustering revealed high cellular heterogeneity, with clusters displaying intermediate expression levels of various marker genes. Four cell clusters (0, 1, 2, and 6), which included the three largest subpopulations, expressed *CD3*, *CD8A*, *CCL5*, *KLRD1*, and *NKG7*, and likely corresponded to Temra cells (Temra 1-4). Interestingly, T-box transcription factor Eomesodermin (*EOMES*) was weakly expressed in Temra 1, Temra 2 and Temra 4, but not in Temra 3 population. We conducted a differential expression analysis to identify other genes that differentiate these cell subsets. *CMC1* was predominantly expressed in the Temra 1 and Temra 2 subsets, whereas Temra 3 lacked *CMC1* expression but showed high levels of granzymes.

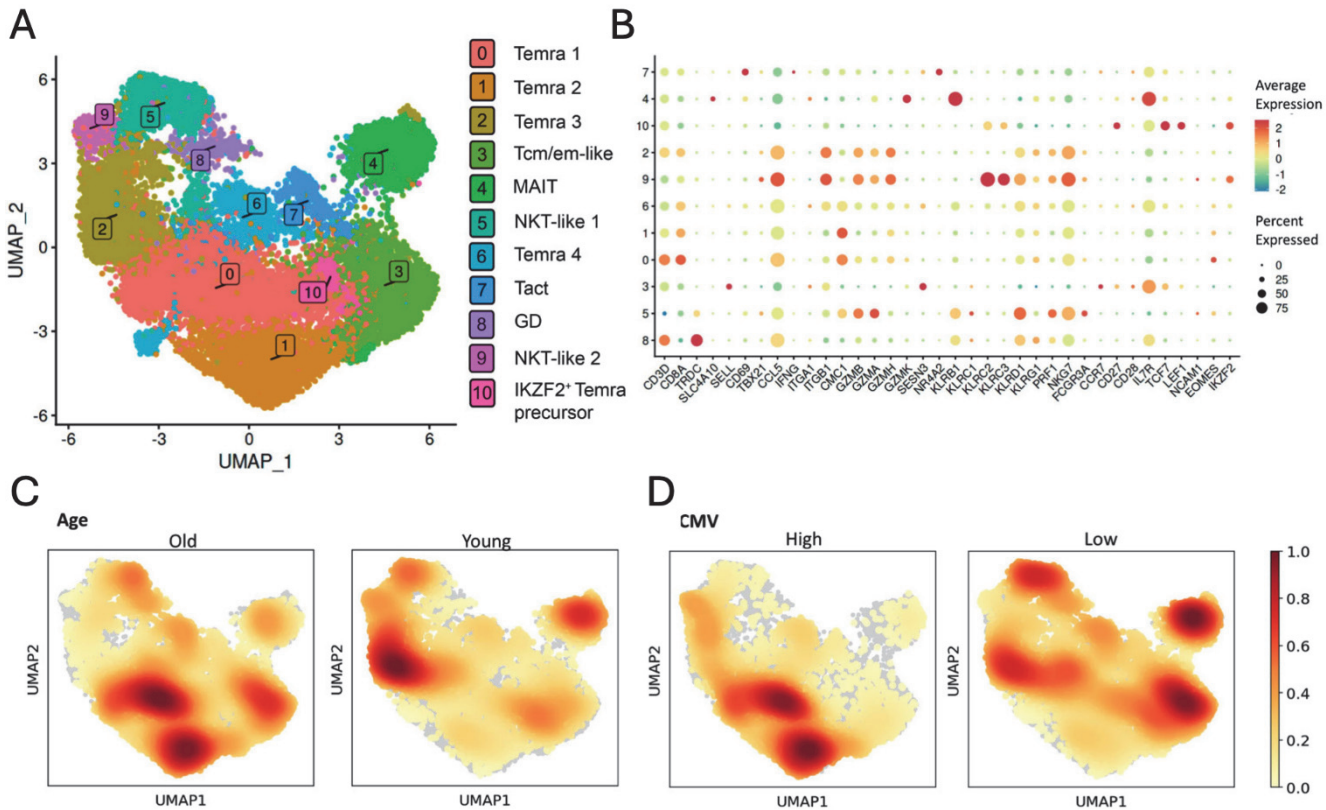


Figure 20. Single-cell analysis of CD8⁺ Temra cells. (A) Two-dimensional UMAP visualization of the sorted CD8⁺ Temra cells and (B) expression of top DEGs. The size of the dot shows the proportion of cells with expression in each cluster. Color of the dot shows average expression of all cells in each cluster. Temra cells colored by age (C) and CMV status (D) of the individuals.

Notably, several of the identified clusters did not directly align with recognized Temra phenotypes. Clusters 5 (*FCGR3A*$^+$*, KLRD1*$^+$*, PRF1*$^+$*)* and 9 (*KLRC2*$^+$*, KLRC3*$^+$*, NKG7*$^+$) expressed low levels of *CD3*, *CD8A*, *CD8B*, and *TYROBP* (DAP12) and were assigned to NKT-like populations. Cluster 3 was classified as Tcm/em-like due to its expression of genes such as *SELL*, *CCR7*, *CD28*, and *IL7R*. Cluster 4 was identified as *KLRB1*$^+$*SLC4A10*$^+$ and was thus classified as MAIT cells. Cluster 7 was identified as activated T cells (Tact) due to its expression of early activation genes *CD69* and *NR4A2*. Cluster 10 represented a small population of *IKZF2*$^+$ Temra precursors (*CD27*$^+$*, LEF1*$^+$*, TCF1*$^+$*, IKZF2*$^+$). Lastly, cluster 8 was comprised of gamma-delta T cells (GD), characterized by *TRDC* and *TRGC1* expression.

6.3.2 Abundance of single-cell clusters among age and CMV groups

We next compared cell abundance across clusters to determine whether age and CMV altered the subpopulation distributions within the sorted CCR7^{lo}CD45RA^{hi}CD8$^+$ T cells. For this, we calculated the density of cells from different conditions and overlaid it on the UMAP plot (Figures 20C,D). We also compared each cluster separately between the age and CMV groups (Figures 21A,B). However, the studied individuals had substantial variation in their subset proportions and most of the differences were not statistically significant. Among main clusters, we observed altered frequencies of Temra 1, 2 and 3, and MAIT cells in aged individuals. *CMC1*$^+$ Temra 1 and 2 were overrepresented among older whereas Temra 3 population was more enriched in young individuals. We also found statistically significant decrease of *IL7R*-expressing Tcm/em-like and MAIT cells in CMV^{hi} individuals, however, this association was not present in comparison between old and young individuals. Our pseudobulk analysis of differentially expressed genes in age and CMV group did not reveal substantial differences, suggesting that the cell subsets and not individual genes are altered among the sorted cell population. The results suggest that Temra populations in old and CMV^{hi} individuals are altered and shifted towards *CMC1*$^+$ Temra population whereas younger individuals tend to have more Tcm/em-like and MAIT cells.

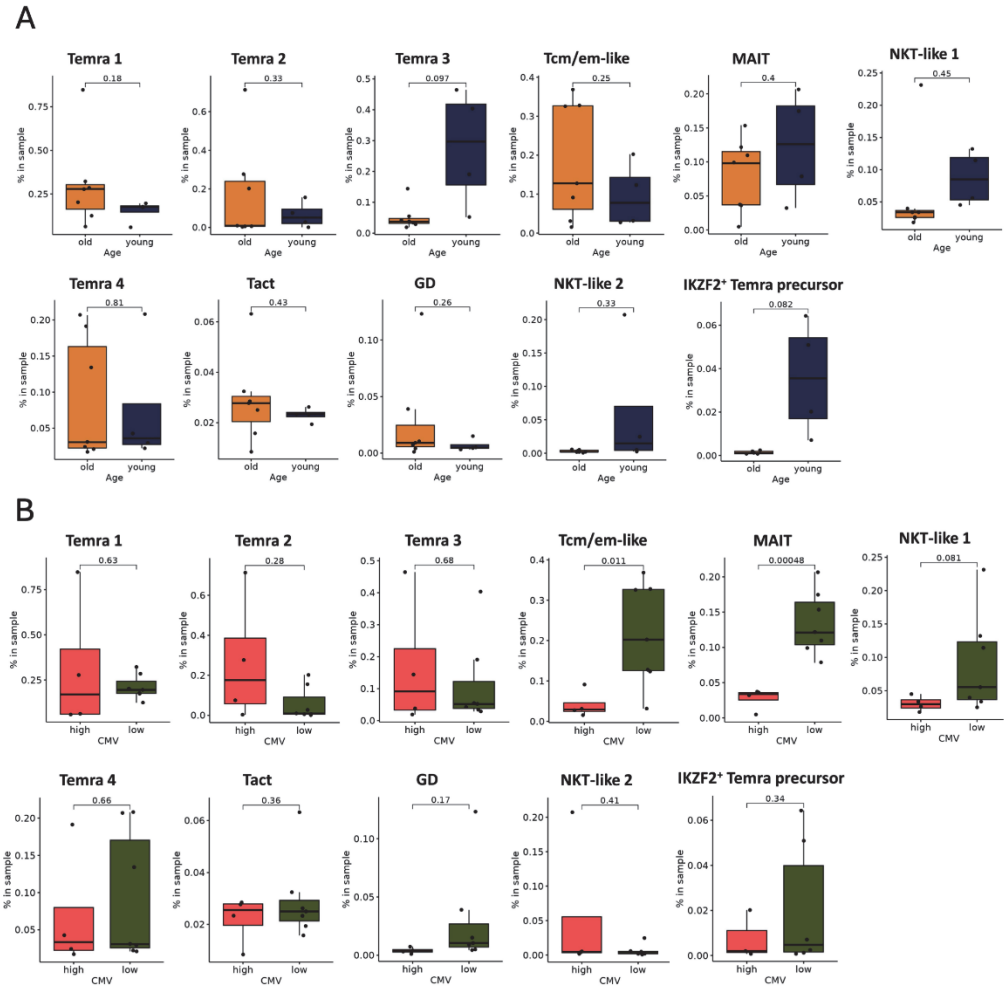
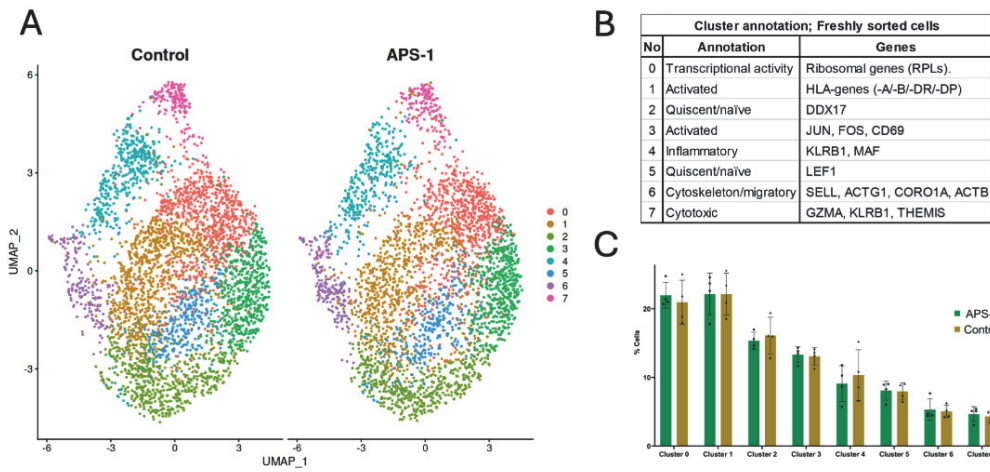


Figure 21. Changes in cell abundance. (A) Cell type proportions compared between young and old groups. (B) Cell type proportions compared between CMV^{hi} and CMV^{lo} groups. Statistical testing was performed using t-test.

6.4 Treg transcriptomes in APS-1 (Paper IV)

6.4.1 Treg subsets in APS-1 mirror those in healthy controls

We aimed to compare freshly sorted APS-1 Tregs with matched healthy controls from a Finnish cohort. To this end, the Tregs were collected from four APS-1 patients and four healthy age-matched controls. The cells were sorted as $CD4^{+}CD25^{+}CD127^{low}$ and subjected to 10x Genomics scRNA-seq protocol. After quality control, high-quality transcriptomic profiles from 7,380 cells were analyzed, resulting in eight distinct clusters (Figure 22A). These clusters were annotated according to the expression of marker genes (Figure 22B). Cluster 0 had a high level of transcriptional activity marked by expression of ribosomal genes. Clusters 1 and 3 consisted of activated cells, showing upregulation of HLA genes along with *JUN*, *FOS*, and *CD69*. Clusters 2 and 5 represented quiescent/naïve cells, marked by *DDX17* and *LEF1* expression. Cluster 4 displayed a pro-inflammatory profile with upregulation of *KLRB1* and *MAF*. Cluster 7 was enriched with genes related to cytotoxicity, such as *GZMA*, *KLRB1*, and *THEMIS*. Lastly, cluster 6 showed increased expression of genes linked to the cytoskeleton and cell migration, including *SELL*, *ACTG1*, *CORO1A*, and *ACTB*. Next, we compared the proportions of identified clusters between APS-1 patients and controls to assess any differences in cluster distribution between the two groups. However, the clusters showed no significant differences in distribution between APS-1 patients and controls (Figure 22C).



Cluster annotation; Freshly sorted cells		
No	Annotation	Genes
0	Transcriptional activity	Ribosomal genes (RPLs).
1	Activated	HLA-genes (-A/-B/-DR/-DP)
2	Quiscent/naïve	DDX17
3	Activated	JUN, FOS, CD69
4	Inflammatory	KLRB1, MAF
5	Quiscent/naïve	LEF1
6	Cytoskeleton/migratory	SELL, ACTG1, CORO1A, ACTB
7	Cytotoxic	GZMA, KLRB1, THEMIS

Figure 22. Single-cell transcriptomic profiles of sorted Tregs from four APS-1 patients and four healthy controls. (A) UMAP plot of global single-cell sequencing data for patients and controls, identifying eight distinct clusters of Tregs. (B) Cluster annotations with some of the most expressed genes in each cluster. (C) Bar plots of cluster frequencies per individual. Statistical testing was performed using t-test.

6.4.2 Treg transcriptomes in APS-1 exhibit mild disturbances

We sought to compare the gene expression profiles of Tregs between APS-1 patients and healthy controls. To this end, the Wilcoxon rank sum test was performed on all cells from both groups. We identified 277 differentially expressed genes, with 152 upregulated and 125 downregulated in APS-1 (Figure 23A; Table S5 in Paper IV). Among the top upregulated genes in APS-1 were *PMAIP1, TNFAIP3, ZFP36L2,* and *NFKBIA*, while the most downregulated genes included *TMSB4X, FYB1, TRAC,* and *IL32*. We compared our DE gene lists with results from single-cell transcriptome sequencing using a targeted Human Immunology Panel, which analyzed the expanded Treg population from an independent Norwegian cohort of eight APS-1 patients and eight healthy controls. To this end, we generated a graph in which the log2FC-values from each of the experiments represented one dimension (Figure 23B). Applying a filter of $|\log2FC| > 0.2$, only two genes were consistently down-regulated (*CD52* and *LTB*) and one gene upregulated (*TXNIP*) in APS-1 patients compared to healthy controls. Together, these results indicate that Treg transcriptomes in APS-1 exhibit mild changes at the transcriptomic level.

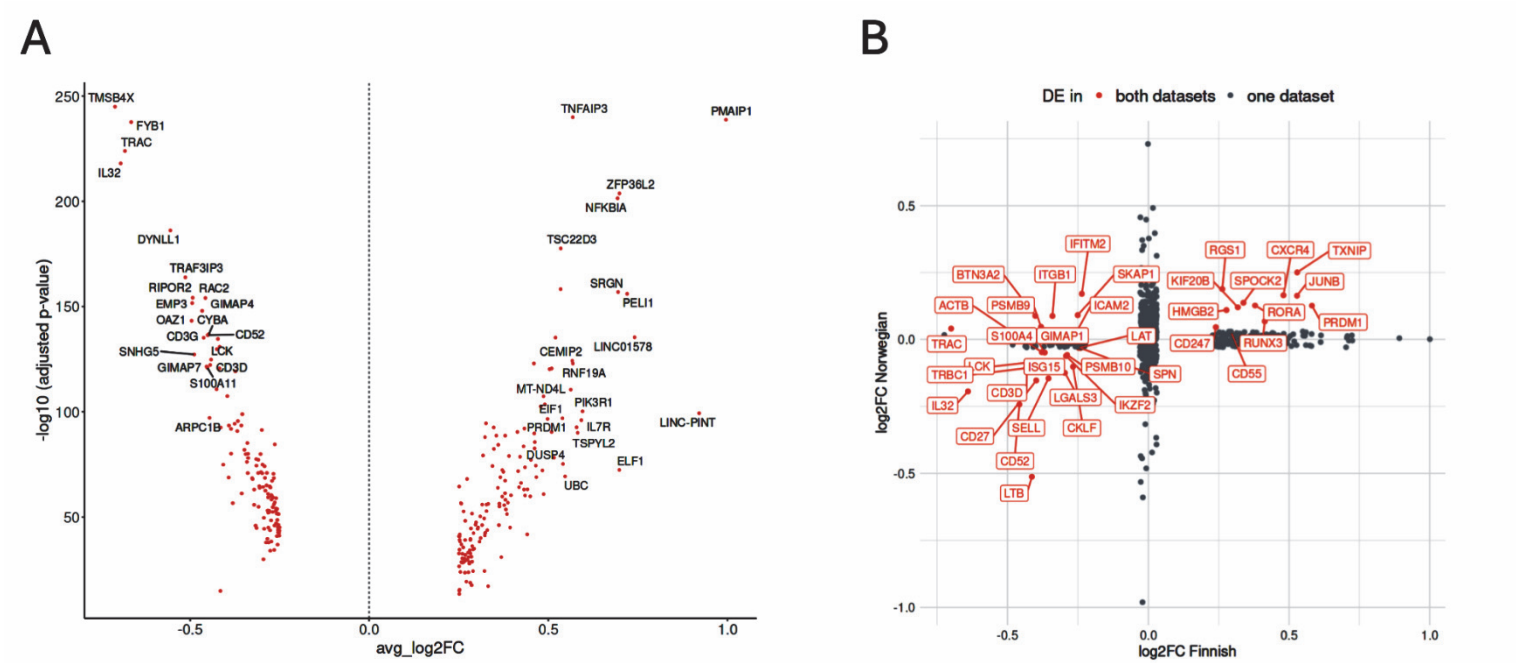


Figure 23. Differential expression analysis of Tregs in APS-1. (A) Differentially expressed genes in Tregs from Finnish APS-1 patients. (B) Differentially expressed genes shared between the Finnish and Norwegian cohorts.

7 DISCUSSION

7.1 Comparative transcriptomic analyses of thymocytes using 10x and Parse scRNA-seq technologies (Paper I)

7.1.1 Parse offers an instrument-free workflow at a cost of additional time

In Paper I, we evaluated the applicability of 10x Genomics and Parse Bio scRNA-seq kits to study the dynamics of T cell differentiation in mouse thymus and focused on cell type annotation as a central task in scRNA-seq data analysis. Our study reports the differences between 10x Genomics and Parse Bio kits, as summarized in Table 3, providing a reference for selecting the optimal approach for single-cell analysis. Considering the set-up costs, the Parse Bio protocol does not require the purchase of the microfluidic device (Chromium Controller) but involves more hands-on steps and reagents compared to 10x Genomics. However, 10x requires additional reagents and steps for sample multiplexing, which are integrated into the Parse Bio workflow. Technical issues with 10x Genomics, such as wetting failures and clogging, are associated with generating gel beds-in-emulsion in the Chromium Controller and can be avoided when using Parse Bio.

Table 3. Key differences between Parse Bio and 10x Genomics kits.

	Parse	**10x**
Inter-sample variability	High	Negligible
Intra-sample variability	High	Negligible
Sensitivity in gene detection	Higher	Lower
Minimum number of cells to load	100,000	800
Minimum number of recovered cells	200	500
Pause points	7	7
Benchwork	~11 h	~7.5 h
Maximum number of samples that can be analyzed in one kit with multiplexing	40 samples with 2500 cells/ sample (1 WT kit)	16 samples with 2500 cells/ sample (1 kit with 4 reactions)
Cell type annotation	Difficult	Easier
Doublet detection	Based on genetic diversity or computational	With hashtags

The time required for the Parse kit workflow, including sample processing and library preparation, is nearly twice as long as that for the 10x Genomics kits with more extensive benchwork (Figure 24). Both kits have the same pause points, but 10x Genomics offers a more flexible three-day workflow, while Parse Bio requires a minimum of a four-day schedule, with one day having more than six hours of workflow without any pause points. Since the cells can be fixed and stored for the Parse Bio kit, it allows samples to be processed from different time

points. Although 10x Genomics has introduced kits in which cells can be fixed and stored, these require the purchase of a different instrument.

		Day 1	Day 2	Day 3	Day 4	Day 5	
10x							Total
Cell prep	Benchwork	1 h					7 h 30 min benchwork
Labelling	Benchwork	1.5 h					
GEM generation and Barcoding	Benchwork	1 h					
	PCR	55 min					12 h 40 min protocol
cDNA Amplification	Benchwork		1.5 h				
	PCR		45 min				
	QC		1 h				
Gene expression library prep	Benchwork			2 h			
	PCR			50 min			
Cell surface protein library prep	Benchwork			30 min			
	PCR			40 min			
QC for both libraries	QC				1 h		
Parse							
Cell prep	Benchwork	1 h					12 h benchwork
Fixation	Benchwork	2 h					
Barcoding	Benchwork		3 h				
	PCR		3.5 h				23 h 30 min protocol
cDNA Amplification	Benchwork			3 h			
	PCR/Vortex			4.5 h			
	QC			1 h			
Library prep	Benchwork				3 h		
	PCR				1.5 h		
	QC					1 h	

Figure 24. Workflow and time frame required for each kit. Key differences between the steps and time spent on 10x Genomics and Parse Bio protocols have been highlighted.

7.1.2 10x Genomics features more consistent cell recovery rates

For cell sample loading, both kits require single-cell suspensions with high viability. Parse Bio requires a minimum of 100,000 cells per sample as starting material, whereas 10x Genomics requires only 800 cells per sample. We observed slightly higher cell recovery rates with 10x Genomics compared to Parse Bio. Notably, three out of four samples processed with Parse Bio showed nearly double the expected cell recovery rate. Notably, 10x Genomics has consistently demonstrated higher cell recovery rates (Xie et al., 2024; Yamawaki et al., 2021). Therefore, 10x Genomics offers an advantage over Parse Bio for studying rare populations, such as thymic epithelial cells, which exist in low abundance in the thymus.

7.1.3 Pronounced differences in QC metrics between 10x Genomics and Parse Bio

Quality metrics of cells in scRNA-seq datasets are crucial for reliable downstream analysis. Common indicators of high-quality cells include the percentage of mitochondrial-coding genes, doublet detection, and the counts of UMIs and genes detected (X. Wang et al., 2021). In our data, Parse showed a slightly higher percentage of mitochondrial-coding genes compared to 10x Genomics. Conversely, 10x Genomics has been reported to capture a higher proportion of ribosomal-coding genes than other platforms (Haque et al., 2017). Our results also showed over a 10-fold increase in ribosomal-coding gene percentages and double the lncRNA-coding gene percentages in 10x Genomics compared to Parse Bio. While 10x Genomics achieved higher cell recovery rates and lower mitochondrial-coding gene percentages, Parse Bio had reduced ribosomal and lncRNA transcript levels, likely due to ambient RNA loss from multiple washing steps in the Parse Bio protocol (Gezelius et al., 2024). The proportion of different gene biotypes affects downstream analysis and should be selected based on specific application needs.

Doublets or multiplets occur when two or more cells are captured in a single droplet, resulting in hybrid transcriptomic profiles. Including these aggregates in downstream analysis can lead to the incorrect identification of cell types, intermediate states, or specific transcriptomic signatures (B. Sun et al., 2021). In our dataset, we applied cell hashing to the 10x samples, allowing us to distinguish doublets from single cells and exclude them from further analysis. Although Parse Bio uses three rounds of barcoding to minimize doublet formation, we identified 31% doublets in the Parse Bio library using the *scDblFinder* package, compared to 14% in 10x Genomics using hashtags. A previous study benchmarking scRNA-seq kits with mouse and human cells reported a high doublet rate of up to 49% for Parse (Hornung et al., 2023), while other studies with human PBMCs and tumor nuclei have shown lower doublet rates in Parse Bio compared to 10x Genomics (Brown et al., 2024; Xie et al., 2024). These results might indicate that analyzing different types of tissues or cells in the same Parse Bio kit may aid in better detection of doublets in the libraries and improve downstream analysis.

Finally, Parse Bio kit identified nearly double the genes detected by 10x Genomics at similar sequencing depth. However, the considerable variability in the detected UMIs and genes between the technical replicates in Parse data suggested poor reproducibility compared to 10x Genomics. The strong batch effect between samples in Parse Bio data may be due to the differences in protocol introduced during the fixation, freezing, thawing, or filtration steps. While the integration of computational batch correction tools into the scRNA-seq data workflow complicates the data analysis pipeline, it becomes essential to utilize these tools while analyzing samples from multiple time-points or large number of samples in different runs.

7.1.4 Challenges in minor thymocyte population annotation with Parse

We found 10x efficiently differentiated between all major thymocyte populations. Further analysis distinguished between various subsets of DP cell populations as published previously (Y. Li et al., 2021). In contrast, Parse did not show the expected *Cd3d* and *Cd3g* gene expression levels. This resulted in difficulty in accurately annotating the minor thymocyte populations. We also could not identify the DP-A cluster in Parse. Nevertheless, we found comparable cell trajectories in the thymic populations between the two methods despite differences in the ratio of spliced and unspliced transcripts.

7.2 PBMC ageing scRNA-seq data integration for immune biomarker validation (Paper II)

7.2.1 PBMC aging atlas enables comprehensive cell type annotation

In Paper II, we developed an integrated atlas of peripheral immune cells to represent healthy human aging. Although previous studies have analyzed transcriptomic changes in human PBMCs across different ages, direct comparisons were limited by inconsistent cell type annotations across datasets. Our aging atlas addresses this gap by enabling a more granular exploration of distinct cell populations, thanks to the large sample size of over a million annotated cells from 103 individuals. To our knowledge, this is the most extensive single-cell RNA-seq reanalysis of the aging human immune system, comparable only to a recent large-scale study that profiled approximately 2 million cells across various age groups in healthy human blood (Terekhova et al., 2023).

7.2.2 Heterogeneity of *MALAT1^{hi}* cells

Our atlas enabled the detailed investigation of the *MALAT1^{hi}* population previously associated with T cell exhaustion aging (Luo et al., 2022). *MALAT1* has recently been implicated in T cell development and associated with COVID-19 immune response (Dey et al., 2023; Kanbar et al., 2022). Our analysis showed that *MALAT1^{hi}* cells were not equally distributed in seven datasets, with higher enrichment in GSE157007, HRA003766, and syn22255433. Interestingly, this cell population was heterogeneous in the expression of *CD3D* and T cell lineage markers (*CD4*, *CD8A*). Moreover, these cells varied considerably in their expression of naïve (*CCR7*, *PASK*, *LRRN3*) and exhaustion (*EOMES*, *PDCD1*, *HAVCR2*) markers. Taken together, our findings suggest that the *MALAT1^{hi}* cell population is highly heterogeneous. However, *MALAT1* is consistently the top expressed gene in 10x Genomics datasets, and its enrichment could suggest dying or damaged cells. Recent analyses of *MALAT1* expression in scRNA-seq datasets

suggest its potential utility as a quality control marker in the data processing workflow (Montserrat-Ayuso & Esteve-Codina, 2024). In line with this, we observed that the *MALAT1*^{hi} population in our data exhibited fewer UMIs per cell and a higher mitochondrial gene percentage than other PBMCs, indicating these cells may be under stress. Consequently, caution is warranted when interpreting the role of the *MALAT1*-enriched population in scRNA-seq data, as emerging evidence suggests it may serve as a quality measure in contrast to its biological relevance in T cell exhaustion.

7.2.3 Refining T cell annotations in the aging data

We here propose the scRNA-seq annotation to distinguish Tn and Tcm subsets from PBMC data. Previously, Cano-Gamez and colleagues showed that *LRRN3* is expressed in *CD4*$^+$ Tn but not *CD4*$^+$ Tcm cells, while *PASK* is expressed at higher levels in *CD4*$^+$ Tcm (Cano-Gamez et al., 2020). Our re-analysis of Tn and Tcm subsets demonstrated *LRRN3* as a subset marker not only for *CD4*$^+$ but also *CD8*$^+$ Tn cells, whereas the expression of PASK was specific for *CD8*$^+$ Tn cells. In contrast, *CD69* was expressed at higher level in *CD4*$^+$ and *CD8*$^+$ Tcm populations. The employment of consistent annotation markers is critical for proper cell identification and the prerequisite for the comparative analysis of scRNA-seq datasets.

7.2.4 *CD8*$^+$ Tn cells are most consistently changed in aging

The *CD8*$^+$ Tn subset showed the most consistent change in abundance during aging, confirmed by our analysis of individual cohorts and the integrated dataset. The decrease in CD8$^+$ Tn numbers has been associated with thymic involution, their increasingly defective regeneration capacity and conversion into effector memory phenotype cells over years (Sprent & Surh, 2011; Van Den Broek et al., 2018). Despite the high experimental variability, the scRNA-seq datasets shared multiple up- or downregulated genes in young and old individual samples. We found an increased expression of multiple effector and inflammation-associated genes in old individuals, including *PTGER2*, *CLIC1*, *S100A4*, *PRF1*, *GZMA*, and *CCL5*. We noted the upregulation of the same genes in a large dataset that studied the transcriptional landscape of age in PBMCs and our transcriptome analysis on aging CD8$^+$ T cells (Peters et al., 2015; Tserel et al., 2015). PTGER2 is a receptor for prostaglandin E2, a key mediator of inflammation. It has been shown to induce senescence markers, loss of CD28 expression, increased p16 cell cycle inhibitor expression, and diminished IL-2 and IFN-γ production in CD8$^+$ T cells (Chou et al., 2014). The *CLIC1* gene encodes a chloride intracellular channel 1 and is a sensor of oxidative stress in cell nucleus. *S100A4*, is a potent trigger of inflammatory processes and is secreted in response to stress situations, stimulating the release of several cytokines and growth factors in the lymphoid and myeloid cells (Ambartsumian et al., 2019). *PRF1* encodes perforin, a protein essential for the cytotoxic function of CD8$^+$ T cells, enabling them to lyse target cells by forming

pores in their membranes (Migueles et al., 2002; Vemulawada et al., 2024). *GZMA* encodes granzyme A, a serine protease that, in addition to its role in cell killing, can stimulate the release of pro-inflammatory cytokines (Arias et al., 2014; Metkar et al., 2008). Finally, CCL5, also known as RANTES, is a chemokine critically involved in the inflammatory response by mediating the recruitment of leukocytes to sites of infection (Marra & Tacke, 2014). Thus, the upregulation of these genes may reflect the higher inflammatory environment in the blood of old individuals. Inflammageing, used to describe age-related increase in the levels of pro-inflammatory markers is considered a strong risk factor for multiple chronic age-related diseases (Mittelbrunn & Kroemer, 2021). Nevertheless, the functional role of these genes in naïve $CD8^+$ T cell homeostatic proliferation, maintenance and priming capacity, and their pathophysiological correlations remain to be studied.

7.2.5 Limitations and future directions

Despite collating an extensive collection of cells from many donors and multiple studies, comprehensive phenotype-metadata associations remain challenging. As the age group definitions vary between the studies, the results of cell type abundance comparisons and differential expression should be interpreted cautiously. Recent scRNA-seq studies have demonstrated cell type-specific effects of genetic variation on gene expression, which may contribute to the variability seen in our analysis (Schmiedel et al., 2022). Our study also included datasets from different ethnic origins. In addition, the metadata provided by the authors lacks standardization. For example, the HRA000624 metadata did not specify the exact age of donors but rather the age range. Factors like genetic background, ethnicity and socioeconomic status could influence immune phenotypes, highlighting the importance of collecting additional metadata for the analysis.

7.3 $CD8^+$ Temra cells in CMV and aging (Paper III)

7.3.1 Heterogeneity of $CD8^+$ Temra cells

The senescent $CD8^+$ Temra population has been shown to exhibit transcriptional heterogeneity with multiple subsets identified previously (Callender et al., 2018; J. Lu et al., 2022). In this study, we focused on exploring heterogeneity specifically within $CD8^+CCR7^{lo}CD45RA^{hi}$ sorted cells and identified a total of ten distinct cell populations. We classified the three largest clusters (0, 1, and 2), along with cluster 6, as Temra cells. These clusters expressed genes such as *NKG7*, *GZMH*, *KLRD1*, and *CMC1* and appeared likely to be TCR-responsive, as indicated by *CD3D* and *CD8A* expression. Temra 1 and 2 were notably increased in older individuals and showed distinct *CMC1* expression. *CMC1*, a mitochondrial protein, functions as a chaperone for the cytochrome c oxidase complex, which plays a key role in the final steps of electron transfer and ATP production (Bourens & Barrientos, 2017). The enrichment of $CMC1^+$ Temra

subsets in older individuals aligns with findings by Dong et al., who observed increased *CMC1*$^{+}$ and *GZMB*$^{+}$ cytotoxic T cell clusters in centenarian PBMCs (Dong et al., 2022).

In contrast, the Temra 3 population was positive for both *GZMB* and *GZMH*, potentially aligning with the *GZMB*$^{+}$ cytotoxic cells previously described (Dong et al., 2022). *GZMB* and *GZMH* also served as markers for two Temra subpopulations identified previously (J. Lu et al., 2022). A recent analysis similarly reported two *GZMB*$^{+}$ clusters within late-differentiated *CD8*$^{+}$ T cells, categorized as *GZMB*$^{+}$ Tem and Temra (Terekhova et al., 2023). In that study, no differences in late-differentiated *CD8*$^{+}$ T cells were observed between younger and older individuals, though an age-related increase in *GZMK*$^{+}$ Tem (CD45RA^{-}) was noted, which likely fell outside our gating parameters. Interestingly, unlike *CMC1*$^{+}$ Temra cells, we observed a lower frequency of Temra 3 cells in older individuals. However, this trend was not statistically significant due to our limited sample size. Confirming the association of higher Temra 3 levels with younger individuals would require a larger cohort and functional cytotoxicity assays.

7.3.2 Non-Temra subsets in CMV and aging

We also identified two clusters with low *CD3D* and *CD8A* expression, which we termed NKT-like cells. These cells expressed cytotoxic markers *GZMB*, *GZMA*, and *NKG7*, suggesting a resemblance to Temra cells. These cells had a reduced *CD8A* expression and potentially represented a transition from Temra to NK-like cells with low TCR signaling and high cytotoxicity (Pereira et al., 2020). However, we did not observe an increase in NKT-like cells in older individuals, contrary to the expected age-related accumulation. Interestingly, we found a decrease in Tcm/em-like and MAIT cell populations in CMV^{hi} individuals. Given that these cells express high levels of *IL7R*, this suggests lower *IL7R* expression in CMV^{hi} hosts. Previous studies have linked higher numbers of CMV-specific CD8^{+} T cells with reduced *IL7R* expression on memory T cells (Sauce et al., 2007), suggesting that persistent CMV infection may slow *IL7R* recovery on circulating CD8^{+} T cells. Our findings further indicate that chronic CMV infection might reduce *IL7R*$^{+}$ CD8^{+} T cell populations, hinting that MAIT cells may play a role in the T-cell response to CMV – a connection previously observed in CMV-seropositive individuals.

7.4 Treg transcriptomes in APS-1 (Paper IV)

Tregs and their suppressive functions are critical for maintaining peripheral tolerance, with proper Treg activity essential to prevent autoimmune diseases. Understanding Treg function in APS-1 is necessary to develop novel therapies for patients. It has long been proposed that autoimmune pathology in APS-1 stems from defects in both central and peripheral tolerance, as AIRE is key for thymic expression of tissue-restricted antigens and may also play a role in Treg gene-

ration that remains incompletely understood (Anderson et al., 2005; Aschenbrenner et al., 2007; Malchow et al., 2016). Notably, APS-1 patients consistently display reduced levels of circulating Tregs (Sng et al., 2019; Wolff et al., 2010). Previous bulk RNA sequencing of freshly isolated Tregs suggested impairments in lipid metabolism and gut homing in APS-1 patients (Berger et al., 2021). In this study, we reveal subtle transcriptomic alterations in both freshly sorted and expanded Tregs from APS-1 patients.

Both expanded and freshly sorted Tregs showed some transcriptomic differences between cohorts, though log-fold change values were small, with only three genes consistently differentially expressed across both freshly sorted and expanded cells: *CD52* and *LTB* (downregulated in patients), and *TXNIP* (upregulated in patients). Reduced *CD52* expression may indicate decreased Treg effector capabilities, while lower *LTB* expression suggests impaired lipid metabolism, previously observed via bulk sequencing through altered *FASN* expression (Berger et al., 2021; Samten, 2013). This may also relate to Treg instability and aberrant migration (Saxena et al., 2022). *TXNIP* encodes thioredoxin-interacting protein, an important regulator of glucose metabolism and redox state; the MondoA-TXNIP axis is a crucial metabolic regulator of Treg identity and function, and altered *TXNIP* expression may impact overall Treg function.

8 CONCLUSIONS

1. In comparing the 10x Genomics and Parse Bio scRNA-seq technologies using murine thymocytes, we found that 10x Genomics demonstrated more consistent cell capture efficiency and uniformity across samples, while Parse Bio showed higher variability and detected more unique genes, though with notable batch effects and a higher doublet rate. Both platforms had distinct gene enrichment patterns, with 10x Genomics showing greater sensitivity to lineage-defining markers such as *Cd3d*. RNA velocity analysis further showed that both platforms captured developmental trajectories. Lastly, while 10x Genomics requires specialized instrumentation for processing, Parse Bio does not; however, the time required for processing with each platform should be carefully considered. In summary, these findings highlight distinct strengths and limitations for each platform in capturing the transcriptional diversity of thymocyte populations.
2. The integration and analysis of scRNA-seq data from multiple studies provided a unified PBMC atlas that enabled consistent immune cell type annotation across aging datasets. This atlas revealed notable age-related changes in cell type abundance, particularly a consistent decline in $CD8^+$ Tn cells and MAIT cells. *MALAT1*-expressing cells exhibited distinct heterogeneity, with their abundance varying significantly across datasets. While some cell populations showed consistent trends with age, significant variability in cell abundance and gene expression patterns underscored the complexity of immune aging. The *CD8*$^+$ Tn transcriptome showed an upregulation of molecules with pro-inflammatory and effector functions, emphasizing conserved biological processes linked to immune cell activation. This work underscores the value of multi-cohort integrative analyses in understanding immune aging, while also addressing challenges associated with inter-study variability in single-cell transcriptomics.
3. In this study, we used scRNA-seq to explore the heterogeneity of $CD8^+$ $CCR7^{lo}CD45RA^{hi}$ Temra cells in relation to aging and CMV infection. Our analysis identified ten distinct clusters, with the largest four clusters corresponding to Temra subpopulations, characterized by genes such as *NKG7*, *KLRD1*, and *CMC1*. Notably, *CMC1*$^+$ Temra 1 and Temra 2 subsets were more abundant in older individuals, whereas Temra 3, which expressed high levels of granzymes *GZMB* and *GZMH*, was more enriched in younger individuals. Additionally, we identified non-Temra clusters, including NKT-like, MAIT, and Tcm/em-like cells, showing distinctive marker profiles. CMV^{hi} individuals displayed a reduction in *IL7R*$^+$ Tcm/em-like and MAIT cells, supporting a potential role for CMV in modulating memory T cell composition. Together, these findings underscore how aging and CMV infection shape $CD8^+$ Temra cell diversity.
4. We conducted an in-depth single-cell transcriptomic analysis of freshly sorted and expanded Tregs from APS-1 patients. We found that AIRE deficiency had

minimal impact on the Treg transcriptome, suggesting APS-1 is not primarily a functional Tregopathy. Rather, the reduced Treg count may result from impaired generation in the thymus or periphery. However, our study has limitations. The sample size was limited, and caution is warranted when comparing freshly sorted and expanded Tregs, as they differ in origin, state, and activation

9 SUMMARY IN ESTONIAN

Üherakuline andmeanalüüs immunoloogias: tehnoloogiast rakendusteni

Geenide avaldumise analüüs üksiku raku tasandil ehk üksikraku sekveneerimine (scRNA-seq) on võimas tehnoloogia, mis võimaldab uurida geenide tööd samaaegselt erinevates rakutüüpides ning analüüsida ühes katses tuhandeid kuni miljoneid rakke (Jovic jt, 2022). Erinevalt traditsioonilistest transkriptoomi analüüsi meetoditest, mis keskmistavad geenide avaldumise kogu rakupopulatsioonis, annab üksikraku sekveneerimine unikaalse molekulaarse profiili iga raku kohta. scRNA-seq meetod toob esile rakkude heterogeensuse ja ka haruldased või seni tundmatud rakutüübid (Miao jt, 2020). Keerukate bioloogiliste süsteemide mõistmisel on üksikraku tehnoloogia oluline kuna erinevad rakutüübid koostoimes mõjutavad mitmeid füsioloogilisi ning patoloogilisi muutusi (Ding jt, 2020; X. Wu jt, 2021). Rakulise mitmekesisuse kaardistamise abil on võimalik tuvastada füsioloogiliste seisundite või haigustega seotud molekulaarseid mustreid, jälgida rakkudes dünaamilisi muutusi ja avastada võimalikke biomarkereid (Brazovskaja jt, 2024; Matson jt, 2022). Sellised teadmised on olulised haiguste mehhanismide mõistmiseks ja sihitud ravimeetodite väljatöötamiseks, mis lõppkokkuvõttes tõhustavad diagnostika täpsust ja ravistrateegiaid (Jagadeesh jt, 2022; Lim jt, 2023; Van De Sande jt, 2023).

Vananemine mõjutab immuunrakkude tasakaalu ja funktsiooni, ning selle tagajärjel langeb nii kaasasündinud kui ka omandatud immuunvastuse võimekus. Kaasasündinud immuunsuse osas on vananemisel leitud funktsionaalseid häireid erinevates rakutüüpides nagu monotsüütides, makrofaagides, dendriitrakkudes ja loomulikes tapjarakkudes (NK; *natural killers*). Näiteks langeb neis rakkudes fagotsütoosi aktiivsus ja muutub mediaatorvalkude, tsütokiinide, tootmine (Germic jt, 2019; Hearps jt, 2012; Reitsema jt, 2024). Muutub ka omandatud immuunsus. B- ja T-rakkude seas vähenevad naiivsete rakkude ja suurenevad mälurakkude osakaalud, mis nõrgendab T-rakkude võimet reageerida uutele antigeenidele (Jalali jt, 2022; Mittelbrunn & Kroemer, 2021). T-rakkude vananemine, mida iseloomustab langenud proliferatiivne võimekus ja mitokondrite aktiivsus, ning telomeeride lühenemine, mõjutavad lisaks immuunvastuse funktsiooni (Bektas jt, 2019; Martínez-Zamudio jt, 2021). Kroonilised infektsioonid, nagu tsütomegaloviirus (CMV), põhjustavad viirus-spetsiifiliste T-rakkude klonaalset laienemist, vähendades immuunrepertuaari mitmekesisust ja lisades immunosenesentsust (Ripa jt, 2017; Zuhair jt, 2019). Selle tagajärjel langeb immuunsüsteemi kaitsevõime, suurendades eakate vastuvõtlikkust infektsioonidele ja kroonilistele põletikulistele haigustele.

APS-1 on haruldane autoimmuunhaigus, mida iseloomustab lai sümptomite spekter. Peamised haigused, mis APS-1 haigetel esinevad on krooniline limaskestade ja naha kandidoos, neerupealiste puudulikkus, hüpoparatüreoidism ja mitmed erinevad autoimmuunsed haigused (Borchers jt, 2020; Kisand & Peterson, 2011b). APS-1 haiguse põhjuseks on mutatsioonid AIRE-geenis. AIRE

geenil on oluline ülesanne inimese immuuntolerantsuse kujunemises, kuna see mõjutab koespetsiifiliste antigeenide avaldumist tüümuses ning AIRE puudulikkuse korral väljuvad tüümusest autoreaktiivsed T-rakud. Autoreaktiivsete T rakkude olemasolu haigete organismis viib ulatuslikule autoimmuunsele reaktsioonile ja autoantikehade tootmisele, millest osa on suunatud interferoonide ja tsütokiinide vastu (Fishman jt, 2017; Peterson & Peltonen, 2005). Tüümuse puudulikkuse tõttu on APS-1 haigetel sageli vähem regulatoorseid T-rakke (Treg) või on nende funktsioon häiritud (Kekäläinen jt, 2007). Tregid on eelkõige olulised perifeerse tolerantsuse säilitamisel ja autoimmuunsuse vältimisel. APS-1 haigetes on Treg rakkude puudulik funktsioon tõenäoliselt üheks peamiseks põhjuseks, mis neis haigetes tekivad autimmuunsed reaktsioonid.

Doktoritöö üldisem eesmärk oli analüüsida üksikraku sekveneerimise tehnoloogia võimalusi ja rakendusi immunoloogilistes uuringutes, et avastada immuunrakkude mitmekesisust ja neis toimuvaid molekulaarseid mehhanisme.

Konkreetsed eesmärgid olid järgmised:

1. Võrrelda *10x Genomics* ja *Parse* scRNA-seq tehnoloogiaid, kasutades selleks tüümust kui primaarset immuunfunktsiooni organit (Artikkel I).
2. Leida immuunsüsteemi vananemisega seonduvaid muutusi koondades seni avaldatud scRNA-seq andmestikud, kus on kirjeldatud vananemisega toimuvaid muutusi perifeerse vere mononukleaarsetes rakkudes (Artikkel II).
3. Uurida $CD8^+$ Temra rakkude heterogeensust vananemise ja CMV infektsiooni korral, kasutades selleks uut scRNA-seq andmestikku (Artikkel III).
4. Kirjeldada transkriptoomi muutusi APS-1 haigete Treg rakkudes ja võrrelda neid tervete kontrollidega kasutades perifeerse vere mononukleaarsete rakkude scRNA-seq andmeid (Artikkel IV).

10x Genomics ja *Parse* üksikraku sekveneerimise meetodite võrdlemiseks analüüsisime paralleelselt diferentseeruvaid T-rakke hiire tüümuses. *Parse* meetodi andmed näitasid tehnilistes kordustes suuremat varieeruvust ning katseseeria (nn *batch*) efekte, mis tõenäoliselt tekkisid rakkude käsitlemise erinevatel etappidel. Mõned T-raku spetsiifilised geenid, nagu *Cd3d*, olid *Parse* meetodi abil halvemini tuvastatavad, mis raskendas väiksemate tümotsüüdi populatsioonide detekteerimist. Suuremad rakutüübid tuvastati edukalt mõlemas andmestikus, ka leiti võrreldavad rakkude diferentseerumise trajektoorid mõlema meetodiga. Kuigi *Parse* tehnoloogia ei nõua spetsiaalset mikrofluidika instrumenti on selle laboratoorne osa pikem ja vajab rohkem käsitsitööd võrreldes *10x Genomics*-ga, mis tagab ühtlasema rakkude märgistamise ja väiksema proovidevahelise varieeruvuse. Seega on konkreetse platvormi valimisel oluline kaaluda üheraku sekveneerimise tehnoloogiate erinevaid külgi, sealhulgas seadistuskulusid ja töövoo sobivust antud katsele.

Inimese vananemise uuringus analüüsisime perifeerse vere mononukleaarseid rakke. Leidsime, et vanemates inimestes olid kõige enam muutunud $CD8^+$ naiivsete T-rakkude populatsioon, mille osakaal oli märkimisväärselt vähenenud. Rühmadevaheline analüüs tuvastas, et vanemates inimestes on ülesreguleeritud põletiku, oksüdatiivse stressi ja tsütotoksiliste funktsioonidega seotud geenid

PTGER2, CLIC1, S100A4, PRF1, GZMA ja *CCL5*. Nende geenide muutused peegeldavad vananemisega seotud põletikulist keskkonda organismis, mida tuntakse *inflammaging*-una. Lisaks uurisime *MALAT1*-geeni ekspresseerivate T-rakkude heterogeensust, mis on varasemalt seotud nende kurnatuse ja stressivastustega. Meie analüüs näitas *MALAT1*$^{+}$ rakkude varieeruvat osakaalu andmestikes ja nende seondumise kvaliteedikontrolli mõõdikutega. Seega annab *MALAT1* geeni marker võimaluse kasutada teda pigem kvaliteedikontrollina kui otsese bioloogilise tunnusena T-rakkude kurnatusele. Saadud tulemused annavad teavet immuunsüsteemi vananemise kohta ja näitavad kui oluline on rohkemate andmestike integreerimine, et tuvastada ühiseid molekulaarseid signatuure.

Vananedes suureneb inimeses CD8^{+} T rakkude seas terminaalselt diferentseerunud rakkude alampopulatsioon, mida voolutsütomeetrias sorditakse markeritega CCR7 ja CD45RA. Kolmandas töös tuvastasime vananemisega seotud CD8^{+} Temra rakkude seas märkimisväärse heterogeensuse. Leidsime scRNA-seq meetodit kasutades sorditud CD8^{+} CCR7^{-} CD45RA^{+} rakkude hulgas kümme eristatavat alampopulatsiooni. CD8^{+} Temra 1 ja 2 rakupopulatsioonid olid vanemates inimestes rohkem esindatud ja näitasid mitokondriaalse funktsiooniga CMC1 geeni kõrget ekspressiooni. Nooremates inimestes oli rohkem esindatud Temra 3 alampopulatsioon, mida iseloomustas *GZMB* ja *GZMH* koekspressioon. Lisaks tuvastati sorditud rakkude seas NKT-laadsed alampopulatsioonid. Neil oli madalam *CD3D* ja *CD8A*, ja kõrgem tsütotoksiliste markerite ekspressioon, mis võib viidata Temra rakkude üleminekule vananedes NK-laadseteks rakkudeks. Krooniline CMV infektsioon oli seotud *IL7R*$^{+}$ T-rakkude populatsioonide vähenemisega, sealhulgas Tcm/em-sarnaste ja MAIT rakkudega. Seega võib järeldada, et püsiv CMV infektsioon mõjutab CD8^{+} T-rakkude alampopulatsioone ja nende tsütokiiniretseptorite ekspressiooni. Kokkuvõttes näitavad meie tulemused Temra rakkude heterogeensust ja seost vananemise ja CMV infektsiooniga, mis omakorda mõjutab vanemate inimeste immuunsüsteemi vastupanuvõimet.

APS-1 Treg-rakkude scRNA-seq analüüs näitas mõningaid muutusi geenide avaldumises. Näiteks olid *CD52* ja *LTB* geenid madalamalt ekspresseerunud APS-1 haigete Tregides, mis võib mõjutada nende rakkude funktsionaalset võimekust alla suruda autoimmuunseid reaktsioone. Geeni *TXNIP* tase oli tõusnud, mis omakorda võib olla seotud Tregide metaboolsetele muutustega. Tulemused on kooskõlas varasemalt näidatud APS-1 Treg-rakkude funktsionaalsete puudustega ning seonduvad Treg-rakkude muutunud glükoosi- ja lipiidide ainevahetuse mõjuga nende funktsioonile. Antud muutuste mõistmine annab informatsiooni APS-1 immuunregulatsiooni häirete kohta ja võib olla abiks terapeutiliste strateegiate väljatöötamisel.

REFERENCES

Aaltonen, J., Björses, P., Perheentupa, J., Horelli–Kuitunen, N., Palotie, A., Peltonen, L., Lee, Y. S., Francis, F., Henning, S., Thiel, C., Leharach, H., & Yaspo, M. (1997). An autoimmune disease, APECED, caused by mutations in a novel gene featuring two PHD-type zinc-finger domains. *Nature Genetics*, *17*(4), 399–403. https://doi.org/10.1038/ng1297-399

Abdill, R. J., Talarico, E., & Grieneisen, L. (2024). A how-to guide for code sharing in biology. *PLOS Biology*, *22*(9), e3002815. https://doi.org/10.1371/journal.pbio.3002815

Ahlmann-Eltze, C., & Huber, W. (2021). glmGamPoi: Fitting Gamma-Poisson generalized linear models on single cell count data. *Bioinformatics*, *36*(24), 5701–5702. https://doi.org/10.1093/bioinformatics/btaa1009

Akbar, A. N., Henson, S. M., & Lanna, A. (2016). Senescence of T Lymphocytes: Implications for Enhancing Human Immunity. *Trends in Immunology*, *37*(12), 866–876. https://doi.org/10.1016/j.it.2016.09.002

Alder, J. K., & Armanios, M. (2022). Telomere-mediated lung disease. *Physiological Reviews*, *102*(4), 1703–1720. https://doi.org/10.1152/physrev.00046.2021

Alter-Wolf, S., Blomberg, B. B., & Riley, R. L. (2009). Deviation of the B Cell Pathway in Senescent Mice Is Associated with Reduced Surrogate Light Chain Expression and Altered Immature B Cell Generation, Phenotype, and Light Chain Expression. *The Journal of Immunology*, *182*(1), 138–147. https://doi.org/10.4049/jimmunol.182.1.138

Ambartsumian, N., Klingelhöfer, J., & Grigorian, M. (2019). The Multifaceted S100A4 Protein in Cancer and Inflammation. In C. W. Heizmann (Ed.), *Calcium-Binding Proteins of the EF-Hand Superfamily* (Vol. 1929, pp. 339–365). Springer New York. https://doi.org/10.1007/978-1-4939-9030-6_22

Amezquita, R. A., Lun, A. T. L., Becht, E., Carey, V. J., Carpp, L. N., Geistlinger, L., Marini, F., Rue-Albrecht, K., Risso, D., Soneson, C., Waldron, L., Pagès, H., Smith, M. L., Huber, W., Morgan, M., Gottardo, R., & Hicks, S. C. (2020). Orchestrating single-cell analysis with Bioconductor. *Nature Methods*, *17*(2), 137–145. https://doi.org/10.1038/s41592-019-0654-x

Anderson, M. S., Venanzi, E. S., Chen, Z., Berzins, S. P., Benoist, C., & Mathis, D. (2005). The Cellular Mechanism of Aire Control of T Cell Tolerance. *Immunity*, *23*(2), 227–239. https://doi.org/10.1016/j.immuni.2005.07.005

Andrews, T. S., & Hemberg, M. (2018). Identifying cell populations with scRNASeq. *Molecular Aspects of Medicine*, *59*, 114–122. https://doi.org/10.1016/j.mam.2017.07.002

Andrews, T. S., Kiselev, V. Y., McCarthy, D., & Hemberg, M. (2021). Tutorial: Guidelines for the computational analysis of single-cell RNA sequencing data. *Nature Protocols*, *16*(1), 1–9. https://doi.org/10.1038/s41596-020-00409-w

Aran, D., Looney, A. P., Liu, L., Wu, E., Fong, V., Hsu, A., Chak, S., Naikawadi, R. P., Wolters, P. J., Abate, A. R., Butte, A. J., & Bhattacharya, M. (2019). Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. *Nature Immunology*, *20*(2), 163–172. https://doi.org/10.1038/s41590-018-0276-y

Arceneaux, D., Chen, Z., Simmons, A. J., Heiser, C. N., Southard-Smith, A. N., Brenan, M. J., Yang, Y., Chen, B., Xu, Y., Choi, E., Campbell, J. D., Liu, Q., & Lau, K. S. (2023). A contamination focused approach for optimizing the single-cell RNA-seq experiment. *iScience*, *26*(7), 107242. https://doi.org/10.1016/j.isci.2023.107242

Arendt, D. (2008). The evolution of cell types in animals: Emerging principles from molecular studies. *Nature Reviews Genetics*, *9*(11), 868–882. https://doi.org/10.1038/nrg2416

Arendt, D., Musser, J. M., Baker, C. V. H., Bergman, A., Cepko, C., Erwin, D. H., Pavlicev, M., Schlosser, G., Widder, S., Laubichler, M. D., & Wagner, G. P. (2016). The origin and evolution of cell types. *Nature Reviews Genetics*, *17*(12), 744–757. https://doi.org/10.1038/nrg.2016.127

Arias, M. A., Jiménez de Bagües, M. P., Aguiló, N., Menao, S., Hervás-Stubbs, S., de Martino, A., Alcaraz, A., Simon, M. M., Froelich, C. J., & Pardo, J. (2014). Elucidating Sources and Roles of Granzymes A and B during Bacterial Infection and Sepsis. *Cell Reports*, *8*(2), 420–429. https://doi.org/10.1016/j.celrep.2014.06.012

Aschenbrenner, K., D'Cruz, L. M., Vollmann, E. H., Hinterberger, M., Emmerich, J., Swee, L. K., Rolink, A., & Klein, L. (2007). Selection of Foxp3+ regulatory T cells specific for self antigen expressed and presented by Aire+ medullary thymic epithelial cells. *Nature Immunology*, *8*(4), 351–358. https://doi.org/10.1038/ni1444

Bains, I., Thiébaut, R., Yates, A. J., & Callard, R. (2009). Quantifying Thymic Export: Combining Models of Naive T Cell Proliferation and TCR Excision Circle Dynamics Gives an Explicit Measure of Thymic Output. *The Journal of Immunology*, *183*(7), 4329–4336. https://doi.org/10.4049/jimmunol.0900743

Baran-Gale, J., Morgan, M. D., Maio, S., Dhalla, F., Calvo-Asensio, I., Deadman, M. E., Handel, A. E., Maynard, A., Chen, S., Green, F., Sit, R. V., Neff, N. F., Darmanis, S., Tan, W., May, A. P., Marioni, J. C., Ponting, C. P., & Holländer, G. A. (2020). Ageing compromises mouse thymus function and remodels epithelial cell differentiation. *eLife*, *9*, e56221. https://doi.org/10.7554/eLife.56221

Bektas, A., Schurman, S. H., Gonzalez-Freire, M., Dunn, C. A., Singh, A. K., Macian, F., Cuervo, A. M., Sen, R., & Ferrucci, L. (2019). Age-associated changes in human CD4+ T cells point to mitochondrial dysfunction consequent to impaired autophagy. *Aging*, *11*(21), 9234–9263. https://doi.org/10.18632/aging.102438

Berger, A. H., Bratland, E., Sjøgren, T., Heimli, M., Tyssedal, T., Bruserud, Ø., Johansson, S., Husebye, E. S., Oftedal, B. E., & Wolff, A. S. B. (2021). Transcriptional Changes in Regulatory T Cells From Patients With Autoimmune Polyendocrine Syndrome Type 1 Suggest Functional Impairment of Lipid Metabolism and Gut Homing. *Frontiers in Immunology*, *12*, 722860. https://doi.org/10.3389/fimmu.2021.722860

Berry, R., Watson, G. M., Jonjic, S., Degli-Esposti, M. A., & Rossjohn, J. (2020). Modulation of innate and adaptive immunity by cytomegaloviruses. *Nature Reviews Immunology*, *20*(2), 113–127. https://doi.org/10.1038/s41577-019-0225-5

Blagosklonny, M. V. (2013). Aging is not programmed: Genetic pseudo-program is a shadow of developmental growth. *Cell Cycle*, *12*(24), 3736–3742. https://doi.org/10.4161/cc.27188

Blank, T. O., & Bellizzi, K. M. (2008). A gerontologic perspective on cancer and aging. *Cancer*, *112*(S11), 2569–2576. https://doi.org/10.1002/cncr.23444

Bou Sleiman, M., Roy, S., Gao, A. W., Sadler, M. C., Von Alvensleben, G. V. G., Li, H., Sen, S., Harrison, D. E., Nelson, J. F., Strong, R., Miller, R. A., Kutalik, Z., Williams, R. W., & Auwerx, J. (2022). Sex- and age-dependent genetics of longevity in a heterogeneous mouse population. *Science*, *377*(6614), eabo3191. https://doi.org/10.1126/science.abo3191

Bourens, M., & Barrientos, A. (2017). A *CMC 1* -knockout reveals translation-independent control of human mitochondrial complex IV biogenesis. *EMBO Reports*, *18*(3), 477–494. https://doi.org/10.15252/embr.201643103

Brown, D. V., Anttila, C. J. A., Ling, L., Grave, P., Baldwin, T. M., Munnings, R., Farchione, A. J., Bryant, V. L., Dunstone, A., Biben, C., Taoudi, S., Weber, T. S., Naik, S. H., Hadla, A., Barker, H. E., Vandenberg, C. J., Dall, G., Scott, C. L., Moore, Z., … Bowden, R. (2024). A risk-reward examination of sample multiplexing reagents for single cell RNA-Seq. *Genomics*, *116*(2), 110793. https://doi.org/10.1016/j.ygeno.2024.110793

Brüning, R. S., Tombor, L., Schulz, M. H., Dimmeler, S., & John, D. (2022). Comparative analysis of common alignment tools for single-cell RNA sequencing. *GigaScience*, *11*, giac001. https://doi.org/10.1093/gigascience/giac001

Butler, A., Hoffman, P., Smibert, P., Papalexi, E., & Satija, R. (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nature Biotechnology*, *36*(5), 411–420. https://doi.org/10.1038/nbt.4096

Byrne, A. J., Powell, J. E., O'Sullivan, B. J., Ogger, P. P., Hoffland, A., Cook, J., Bonner, K. L., Hewitt, R. J., Wolf, S., Ghai, P., Walker, S. A., Lukowski, S. W., Molyneaux, P. L., Saglani, S., Chambers, D. C., Maher, T. M., & Lloyd, C. M. (2020). Dynamics of human monocytes and airway macrophages during healthy aging and after transplant. *Journal of Experimental Medicine*, *217*(3), e20191236. https://doi.org/10.1084/jem.20191236

Cabeza-Cabrerizo, M., Cardoso, A., Minutti, C. M., Pereira Da Costa, M., & Reis E Sousa, C. (2021). Dendritic Cells Revisited. *Annual Review of Immunology*, *39*(1), 131–166. https://doi.org/10.1146/annurev-immunol-061020-053707

Cadwallader, L., Mac Gabhann, F., Papin, J., & Pitzer, V. E. (2022). Advancing code sharing in the computational biology community. *PLOS Computational Biology*, *18*(6), e1010193. https://doi.org/10.1371/journal.pcbi.1010193

Cakir, B., Prete, M., Huang, N., van Dongen, S., Pir, P., & Kiselev, V. Y. (2020). Comparison of visualization tools for single-cell RNAseq data. *NAR Genomics and Bioinformatics*, *2*(3), lqaa052. https://doi.org/10.1093/nargab/lqaa052

Callender, L. A., Carroll, E. C., Beal, R. W. J., Chambers, E. S., Nourshargh, S., Akbar, A. N., & Henson, S. M. (2018). Human CD 8^{+} EMRA T cells display a senescence-associated secretory phenotype regulated by p38 MAPK. *Aging Cell*, *17*(1), e12675. https://doi.org/10.1111/acel.12675

Callender, L. A., Carroll, E. C., Bober, E. A., Akbar, A. N., Solito, E., & Henson, S. M. (2020). Mitochondrial mass governs the extent of human T cell senescence. *Aging Cell*, *19*(2), e13067. https://doi.org/10.1111/acel.13067

Campos, C., Pera, A., Sanchez-Correa, B., Alonso, C., Lopez-Fernandez, I., Morgado, S., Tarazona, R., & Solana, R. (2014). Effect of age and CMV on NK cell subpopulations. *Experimental Gerontology*, *54*, 130–137. https://doi.org/10.1016/j.exger.2014.01.008

Cancro, M. P. (2020). Age-Associated B Cells. *Annual Review of Immunology*, *38*(1), 315–340. https://doi.org/10.1146/annurev-immunol-092419-031130

Cano-Gamez, E., Soskic, B., Roumeliotis, T. I., So, E., Smyth, D. J., Baldrighi, M., Willé, D., Nakic, N., Esparza-Gordillo, J., Larminie, C. G. C., Bronson, P. G., Tough, D. F., Rowan, W. C., Choudhary, J. S., & Trynka, G. (2020). Single-cell transcriptomics identifies an effectorness gradient shaping the response of CD4+ T cells to cytokines. *Nature Communications*, *11*(1), 1801. https://doi.org/10.1038/s41467-020-15543-y

Cao, J., Packer, J. S., Ramani, V., Cusanovich, D. A., Huynh, C., Daza, R., Qiu, X., Lee, C., Furlan, S. N., Steemers, F. J., Adey, A., Waterston, R. H., Trapnell, C., & Shendure, J. (2017). Comprehensive single-cell transcriptional profiling of a multicellular organism. *Science*, *357*(6352), 661–667. https://doi.org/10.1126/science.aam8940

Chen, J., Yang, W., Yu, C., & Li, Y. (2008). Autoimmune Regulator Initiates the Expression of Promiscuous Genes in Thymic Epithelial Cells. *Immunological Investigations*, *37*(3), 203–214. https://doi.org/10.1080/08820130801967841

Cherrier, M., Ramachandran, G., & Golub, R. (2020). The interplay between innate lymphoid cells and T cells. *Mucosal Immunology*, *13*(5), 732–742. https://doi.org/10.1038/s41385-020-0320-8

Chidrawar, S. M., Khan, N., Chan, Y. L. T., Nayak, L., & Moss, P. A. (2006). Ageing is associated with a decline in peripheral blood CD56bright NK cells. *Immunity & Ageing*, *3*(1), 10. https://doi.org/10.1186/1742-4933-3-10

Chou, J. P., Ramirez, C. M., Ryba, D. M., Koduri, M. P., & Effros, R. B. (2014). Prostaglandin E2 Promotes Features of Replicative Senescence in Chronically Activated Human CD8+ T Cells. *PLoS ONE*, *9*(6), e99432. https://doi.org/10.1371/journal.pone.0099432

Choudhary, S., & Satija, R. (2022). Comparison and evaluation of statistical error models for scRNA-seq. *Genome Biology*, *23*(1), 27. https://doi.org/10.1186/s13059-021-02584-9

Chung, S. S., & Park, C. Y. (2017). Aging, hematopoiesis, and the myelodysplastic syndromes. *Hematology*, *2017*(1), 73–78. https://doi.org/10.1182/asheducation-2017.1.73

Clarke, A. J., & Simon, A. K. (2019). Autophagy in the renewal, differentiation and homeostasis of immune cells. *Nature Reviews Immunology*, *19*(3), 170–183. https://doi.org/10.1038/s41577-018-0095-2

Clarke, Z. A., Andrews, T. S., Atif, J., Pouyabahar, D., Innes, B. T., MacParland, S. A., & Bader, G. D. (2021). Tutorial: Guidelines for annotating single-cell transcriptomic maps using automated and manual methods. *Nature Protocols*, *16*(6), 2749–2764. https://doi.org/10.1038/s41596-021-00534-0

Cobey, S. (2024). Vaccination against rapidly evolving pathogens and the entanglements of memory. *Nature Immunology*, *25*(11), 2015–2023. https://doi.org/10.1038/s41590-024-01970-2

Cokelaer, T., Cohen-Boulakia, S., & Lemoine, F. (2023). Reprohackathons: Promoting reproducibility in bioinformatics through training. *Bioinformatics*, *39*(Supplement_1), i11–i20. https://doi.org/10.1093/bioinformatics/btad227

Davidson, S., Efremova, M., Riedel, A., Mahata, B., Pramanik, J., Huuhtanen, J., Kar, G., Vento-Tormo, R., Hagai, T., Chen, X., Haniffa, M. A., Shields, J. D., & Teichmann, S. A. (2020). Single-Cell RNA Sequencing Reveals a Dynamic Stromal Niche That Supports Tumor Growth. *Cell Reports*, *31*(7), 107628. https://doi.org/10.1016/j.celrep.2020.107628

Davis, A., Gao, R., & Navin, N. E. (2019). SCOPIT: Sample size calculations for single-cell sequencing experiments. *BMC Bioinformatics*, *20*(1), 566. https://doi.org/10.1186/s12859-019-3167-9

De Haan, G., & Lazare, S. S. (2018). Aging of hematopoietic stem cells. *Blood*, *131*(5), 479–487. https://doi.org/10.1182/blood-2017-06-746412

De Magalhães, J. P. (2023). Ageing as a software design flaw. *Genome Biology*, *24*(1), 51. https://doi.org/10.1186/s13059-023-02888-y

Denisenko, E., Guo, B. B., Jones, M., Hou, R., De Kock, L., Lassmann, T., Poppe, D., Clément, O., Simmons, R. K., Lister, R., & Forrest, A. R. R. (2020). Systematic assessment of tissue dissociation and storage biases in single-cell and single-nucleus RNA-seq workflows. *Genome Biology*, *21*(1), 130. https://doi.org/10.1186/s13059-020-02048-6

Desdín-Micó, G., Soto-Heredero, G., Aranda, J. F., Oller, J., Carrasco, E., Gabandé-Rodríguez, E., Blanco, E. M., Alfranca, A., Cussó, L., Desco, M., Ibañez, B., Gortazar, A. R., Fernández-Marcos, P., Navarro, M. N., Hernaez, B., Alcamí, A., Baixauli, F., & Mittelbrunn, M. (2020). T cells with dysfunctional mitochondria induce multimorbidity and premature senescence. *Science*, *368*(6497), 1371–1376. https://doi.org/10.1126/science.aax0860

Dey, S., Ashwin, H., Milross, L., Hunter, B., Majo, J., Filby, A. J., Fisher, A. J., Kaye, P. M., & Lagos, D. (2023). Downregulation of *MALAT1* is a hallmark of tissue and peripheral proliferative T cells in COVID-19. *Clinical and Experimental Immunology*, *212*(3), 262–275. https://doi.org/10.1093/cei/uxad034

Domínguez Conde, C., Xu, C., Jarvis, L. B., Rainbow, D. B., Wells, S. B., Gomes, T., Howlett, S. K., Suchanek, O., Polanski, K., King, H. W., Mamanova, L., Huang, N., Szabo, P. A., Richardson, L., Bolt, L., Fasouli, E. S., Mahbubani, K. T., Prete, M., Tuck, L., … Teichmann, S. A. (2022). Cross-tissue immune cell analysis reveals tissue-specific features in humans. *Science*, *376*(6594), eabl5197. https://doi.org/10.1126/science.abl5197

Donato, A. J., Machin, D. R., & Lesniewski, L. A. (2018). Mechanisms of Dysfunction in the Aging Vasculature and Role in Age-Related Disease. *Circulation Research*, *123*(7), 825–848. https://doi.org/10.1161/CIRCRESAHA.118.312563

Dong, C. (2021). Cytokine Regulation and Function in T Cells. *Annual Review of Immunology*, *39*(1), 51–76. https://doi.org/10.1146/annurev-immunol-061020-053702

Dong, C., Miao, Y., Zhao, R., Yang, M., Guo, A., Xue, Z., Li, T., Zhang, Q., Bao, Y., Shen, C., Sun, C., Yang, Y., Gu, X., Jin, Y., Li, R., Xu, M., Guo, J., Zong, Z., Zhou, W., … Gu, Z. (2022). Single-Cell Transcriptomics Reveals Longevity Immune Remodeling Features Shared by Centenarians and Their Offspring. *Advanced Science*, *9*(36), 2204849. https://doi.org/10.1002/advs.202204849

Dorfman, D. M., Hwang, E. S., Shahsafaei, A., & Glimcher, L. H. (2004). T-bet, a T-Cell–Associated Transcription Factor, Is Expressed in a Subset of B-Cell Lymphoproliferative Disorders. *American Journal of Clinical Pathology*, *122*(2), 292–297. https://doi.org/10.1309/AQQ2DVM75DVY0PWP

Duggal, N. A., Pollock, R. D., Lazarus, N. R., Harridge, S., & Lord, J. M. (2018). Major features of immunesenescence, including reduced thymic output, are ameliorated by high levels of physical activity in adulthood. *Aging Cell*, *17*(2), e12750. https://doi.org/10.1111/acel.12750

Eberwine, J., Yeh, H., Miyashiro, K., Cao, Y., Nair, S., Finnell, R., Zettel, M., & Coleman, P. (1992). Analysis of gene expression in single live neurons. *Proceedings of the National Academy of Sciences*, *89*(7), 3010–3014. https://doi.org/10.1073/pnas.89.7.3010

Eling, N., Richard, A. C., Richardson, S., Marioni, J. C., & Vallejos, C. A. (2018). Correcting the Mean-Variance Dependency for Differential Variability Testing Using Single-Cell RNA Sequencing Data. *Cell Systems*, *7*(3), 284-294.e12. https://doi.org/10.1016/j.cels.2018.06.011

Elizaga, M. L., Li, S. S., Kochar, N. K., Wilson, G. J., Allen, M. A., Tieu, H. V. N., Frank, I., Sobieszczyk, M. E., Cohen, K. W., Sanchez, B., Latham, T. E., Clarke, D. K., Egan, M. A., Eldridge, J. H., Hannaman, D., Xu, R., Ota-Setlik, A., McElrath, M. J., Hay, C. M., & on behalf of the NIAID HIV Vaccine Trials Network (HVTN) 087 Study Team. (2018). Safety and tolerability of HIV-1 multiantigen pDNA vaccine given with IL-12 plasmid DNA via electroporation, boosted with a recombinant vesicular stomatitis

virus HIV Gag vaccine in healthy volunteers in a randomized, controlled clinical trial. *PLOS ONE*, *13*(9), e0202753. https://doi.org/10.1371/journal.pone.0202753

Ferrucci, L., & Fabbri, E. (2018). Inflammageing: Chronic inflammation in ageing, cardiovascular disease, and frailty. *Nature Reviews Cardiology*, *15*(9), 505–522. https://doi.org/10.1038/s41569-018-0064-2

Fishman, D., Kisand, K., Hertel, C., Rothe, M., Remm, A., Pihlap, M., Adler, P., Vilo, J., Peet, A., Meloni, A., Podkrajsek, K. T., Battelino, T., Bruserud, Ø., Wolff, A. S. B., Husebye, E. S., Kluger, N., Krohn, K., Ranki, A., Peterson, H., … Peterson, P. (2017). Autoantibody Repertoire in APECED Patients Targets Two Distinct Subgroups of Proteins. *Frontiers in Immunology*, *8*, 976. https://doi.org/10.3389/fimmu.2017.00976

Fleck, J. S., Camp, J. G., & Treutlein, B. (2023). What is a cell type? *Science*, *381*(6659), 733–734. https://doi.org/10.1126/science.adf6162

Fleming, S. J., Chaffin, M. D., Arduini, A., Akkad, A.-D., Banks, E., Marioni, J. C., Philippakis, A. A., Ellinor, P. T., & Babadi, M. (2023). Unsupervised removal of systematic background noise from droplet-based single-cell experiments using CellBender. *Nature Methods*, *20*(9), 1323–1335. https://doi.org/10.1038/s41592-023-01943-7

Franceschi, C., Bonafè, M., Valensin, S., Olivieri, F., De Luca, M., Ottaviani, E., & De Benedictis, G. (2000). Inflamm-aging: An Evolutionary Perspective on Immunosenescence. *Annals of the New York Academy of Sciences*, *908*(1), 244–254. https://doi.org/10.1111/j.1749-6632.2000.tb06651.x

Frasca, D., Landin, A. M., Lechner, S. C., Ryan, J. G., Schwartz, R., Riley, R. L., & Blomberg, B. B. (2008). Aging Down-Regulates the Transcription Factor E2A, Activation-Induced Cytidine Deaminase, and Ig Class Switch in Human B Cells. *The Journal of Immunology*, *180*(8), 5283–5290. https://doi.org/10.4049/jimmunol.180.8.5283

Frasca, D., Van Der Put, E., Riley, R. L., & Blomberg, B. B. (2004). Reduced Ig Class Switch in Aged Mice Correlates with Decreased E47 and Activation-Induced Cytidine Deaminase. *The Journal of Immunology*, *172*(4), 2155–2162. https://doi.org/10.4049/jimmunol.172.4.2155

Füllgrabe, A., George, N., Green, M., Nejad, P., Aronow, B., Fexova, S. K., Fischer, C., Freeberg, M. A., Huerta, L., Morrison, N., Scheuermann, R. H., Taylor, D., Vasilevsky, N., Clarke, L., Gehlenborg, N., Kent, J., Marioni, J., Teichmann, S., Brazma, A., & Papatheodorou, I. (2020). Guidelines for reporting single-cell RNA-seq experiments. *Nature Biotechnology*, *38*(12), 1384–1386. https://doi.org/10.1038/s41587-020-00744-z

Fülöp, T., Larbi, A., & Witkowski, J. M. (2019). Human Inflammaging. *Gerontology*, *65*(5), 495–504. https://doi.org/10.1159/000497375

García-Torre, A., Bueno-García, E., López-Martínez, R., Rioseras, B., Díaz-Molina, B., Lambert, J. L., Quirós, C., Alonso-Álvarez, S., Alonso-Arias, R., & Moro-García, M. A. (2021). CMV Infection Is Directly Related to the Inflammatory Status in Chronic Heart Failure Patients. *Frontiers in Immunology*, *12*, 687582. https://doi.org/10.3389/fimmu.2021.687582

Gayoso, A., Steier, Z., Lopez, R., Regier, J., Nazor, K. L., Streets, A., & Yosef, N. (2021). Joint probabilistic modeling of single-cell multi-omic data with totalVI. *Nature Methods*, *18*(3), 272–282. https://doi.org/10.1038/s41592-020-01050-x

Germain, P.-L., Lun, A., Garcia Meixide, C., Macnair, W., & Robinson, M. D. (2022). Doublet identification in single-cell sequencing data using scDblFinder. *F1000-Research*, *10*, 979. https://doi.org/10.12688/f1000research.73600.2

Gezelius, H., Enblad, A. P., Lundmark, A., Åberg, M., Blom, K., Rudfeldt, J., Raine, A., Harila, A., Rendo, V., Heinäniemi, M., Andersson, C., & Nordlund, J. (2024). Comparison of high-throughput single-cell RNA-seq methods for ex vivo drug screening. *NAR Genomics and Bioinformatics*, *6*(1), lqae001. https://doi.org/10.1093/nargab/lqae001

Golstein, P., & Griffiths, G. M. (2018). An early history of T cell-mediated cytotoxicity. *Nature Reviews Immunology*, *18*(8), 527–535. https://doi.org/10.1038/s41577-018-0009-3

Goronzy, J. J., & Weyand, C. M. (2019). Mechanisms underlying T cell ageing. *Nature Reviews Immunology*, *19*(9), 573–583. https://doi.org/10.1038/s41577-019-0180-1

Guilliams, M., Bonnardel, J., Haest, B., Vanderborght, B., Wagner, C., Remmerie, A., Bujko, A., Martens, L., Thoné, T., Browaeys, R., De Ponti, F. F., Vanneste, B., Zwicker, C., Svedberg, F. R., Vanhalewyn, T., Gonçalves, A., Lippens, S., Devriendt, B., Cox, E., … Scott, C. L. (2022). Spatial proteogenomics reveals distinct and evolutionarily conserved hepatic macrophage niches. *Cell*, *185*(2), 379-396.e38. https://doi.org/10.1016/j.cell.2021.12.018

Guilliams, M., Mildner, A., & Yona, S. (2018). Developmental and Functional Heterogeneity of Monocytes. *Immunity*, *49*(4), 595–613. https://doi.org/10.1016/j.immuni.2018.10.005

Hadeiba, H., Lahl, K., Edalati, A., Oderup, C., Habtezion, A., Pachynski, R., Nguyen, L., Ghodsi, A., Adler, S., & Butcher, E. C. (2012). Plasmacytoid Dendritic Cells Transport Peripheral Antigens to the Thymus to Promote Central Tolerance. *Immunity*, *36*(3), 438–450. https://doi.org/10.1016/j.immuni.2012.01.017

Haghverdi, L., Lun, A. T. L., Morgan, M. D., & Marioni, J. C. (2018). Batch effects in single-cell RNA-sequencing data are corrected by matching mutual nearest neighbors. *Nature Biotechnology*, *36*(5), 421–427. https://doi.org/10.1038/nbt.4091

Hao, Y., Hao, S., Andersen-Nissen, E., Mauck, W. M., Zheng, S., Butler, A., Lee, M. J., Wilk, A. J., Darby, C., Zager, M., Hoffman, P., Stoeckius, M., Papalexi, E., Mimitou, E. P., Jain, J., Srivastava, A., Stuart, T., Fleming, L. M., Yeung, B., … Satija, R. (2021). Integrated analysis of multimodal single-cell data. *Cell*, *184*(13), 3573-3587.e29. https://doi.org/10.1016/j.cell.2021.04.048

Haque, A., Engel, J., Teichmann, S. A., & Lönnberg, T. (2017). A practical guide to single-cell RNA-sequencing for biomedical research and clinical applications. *Genome Medicine*, *9*(1), 75. https://doi.org/10.1186/s13073-017-0467-4

Harman, D. (1992). Free radical theory of aging. *Mutation Research/DNAging*, *275*(3–6), 257–266. https://doi.org/10.1016/0921-8734(92)90030-S

Hayflick, L. (2007). Biological Aging Is No Longer an Unsolved Problem. *Annals of the New York Academy of Sciences*, *1100*(1), 1–13. https://doi.org/10.1196/annals.1395.001

He, D., Zakeri, M., Sarkar, H., Soneson, C., Srivastava, A., & Patro, R. (2022). Alevin-fry unlocks rapid, accurate and memory-frugal quantification of single-cell RNA-seq data. *Nature Methods*, *19*(3), 316–322. https://doi.org/10.1038/s41592-022-01408-3

Hearps, A. C., Martin, G. E., Angelovich, T. A., Cheng, W., Maisa, A., Landay, A. L., Jaworowski, A., & Crowe, S. M. (2012). Aging is associated with chronic innate immune activation and dysregulation of monocyte phenotype and function. *Aging Cell*, *11*(5), 867–875. https://doi.org/10.1111/j.1474-9726.2012.00851.x

Heikkilä, N., Laakso, S. M., Mannerström, H., Kekäläinen, E., Saavalainen, P., Jarva, H., & Arstila, T. P. (2016). Expanded CD4+ Effector/Memory T Cell Subset in APECED Produces Predominantly Interferon Gamma. *Journal of Clinical Immunology*, *36*(6), 555–563. https://doi.org/10.1007/s10875-016-0302-5

Heumos, L., Schaar, A. C., Lance, C., Litinetskaya, A., Drost, F., Zappia, L., Lücken, M. D., Strobl, D. C., Henao, J., Curion, F., Single-cell Best Practices Consortium, Aliee, H., Ansari, M., Badia-i-Mompel, P., Büttner, M., Dann, E., Dimitrov, D., Dony, L., Frishberg, A., … Theis, F. J. (2023). Best practices for single-cell analysis across modalities. *Nature Reviews Genetics*, *24*(8), 550–572. https://doi.org/10.1038/s41576-023-00586-w

Hie, B., Peters, J., Nyquist, S. K., Shalek, A. K., Berger, B., & Bryson, B. D. (2020). Computational Methods for Single-Cell RNA Sequencing. *Annual Review of Biomedical Data Science*, *3*(1), 339–364. https://doi.org/10.1146/annurev-biodatasci-012220-100601

Ho, Y.-H., & Méndez-Ferrer, S. (2020). Microenvironmental contributions to hematopoietic stem cell aging. *Haematologica*, *105*(1), 38–46. https://doi.org/10.3324/haematol.2018.211334

Holla, P., Dizon, B., Ambegaonkar, A. A., Rogel, N., Goldschmidt, E., Boddapati, A. K., Sohn, H., Sturdevant, D., Austin, J. W., Kardava, L., Yuesheng, L., Liu, P., Moir, S., Pierce, S. K., & Madi, A. (2021). Shared transcriptional profiles of atypical B cells suggest common drivers of expansion and function in malaria, HIV, and autoimmunity. *Science Advances*, *7*(22), eabg8384. https://doi.org/10.1126/sciadv.abg8384

Hong, R., Koga, Y., Bandyadka, S., Leshchyk, A., Wang, Y., Akavoor, V., Cao, X., Sarfraz, I., Wang, Z., Alabdullatif, S., Jansen, F., Yajima, M., Johnson, W. E., & Campbell, J. D. (2022). Comprehensive generation, visualization, and reporting of quality control metrics for single-cell RNA sequencing data. *Nature Communications*, *13*(1), 1688. https://doi.org/10.1038/s41467-022-29212-9

Hornung, B. V. H., Azmani, Z., Den Dekker, A. T., Oole, E., Ozgur, Z., Brouwer, R. W. W., Van Den Hout, M. C. G. N., & Van IJcken, W. F. J. (2023). Comparison of Single Cell Transcriptome Sequencing Methods: Of Mice and Men. *Genes*, *14*(12), 2226. https://doi.org/10.3390/genes14122226

Hou, Y., Dan, X., Babbar, M., Wei, Y., Hasselbalch, S. G., Croteau, D. L., & Bohr, V. A. (2019). Ageing as a risk factor for neurodegenerative disease. *Nature Reviews Neurology*, *15*(10), 565–581. https://doi.org/10.1038/s41582-019-0244-7

Huang, Z., Chen, B., Liu, X., Li, H., Xie, L., Gao, Y., Duan, R., Li, Z., Zhang, J., Zheng, Y., & Su, W. (2021). Effects of sex and aging on the immune cell landscape as assessed by single-cell transcriptomic analysis. *Proceedings of the National Academy of Sciences*, *118*(33), e2023216118. https://doi.org/10.1073/pnas.2023216118

Hubert, F.-X., Kinkel, S. A., Crewther, P. E., Cannon, P. Z. F., Webster, K. E., Link, M., Uibo, R., O'Bryan, M. K., Meager, A., Forehan, S. P., Smyth, G. K., Mittaz, L., Antonarakis, S. E., Peterson, P., Heath, W. R., & Scott, H. S. (2009). Aire-Deficient C57BL/6 Mice Mimicking the Common Human 13-Base Pair Deletion Mutation Present with Only a Mild Autoimmune Phenotype. *The Journal of Immunology*, *182*(6), 3902–3918. https://doi.org/10.4049/jimmunol.0802124

Husebye, E. S., Perheentupa, J., Rautemaa, R., & Kämpe, O. (2009). Clinical manifestations and management of patients with autoimmune polyendocrine syndrome type I. *Journal of Internal Medicine*, *265*(5), 514–529. https://doi.org/10.1111/j.1365-2796.2009.02090.x

Hwang, B., Lee, J. H., & Bang, D. (2018). Single-cell RNA sequencing technologies and bioinformatics pipelines. *Experimental & Molecular Medicine*, *50*(8), 1–14. https://doi.org/10.1038/s12276-018-0071-8

Islam, S., Zeisel, A., Joost, S., La Manno, G., Zajac, P., Kasper, M., Lönnerberg, P., & Linnarsson, S. (2014). Quantitative single-cell RNA-seq with unique molecular identifiers. *Nature Methods*, *11*(2), 163–166. https://doi.org/10.1038/nmeth.2772

Jalali, S., Harpur, C. M., Piers, A. T., Auladell, M., Perriman, L., Li, S., An, K., Anderson, J., Berzins, S. P., Licciardi, P. V., Ashhurst, T. M., Konstantinov, I. E., & Pellicci, D. G. (2022). A high-dimensional cytometry atlas of peripheral blood over the human life span. *Immunology & Cell Biology*, *100*(10), 805–821. https://doi.org/10.1111/imcb.12594

Johnson, W. E., Li, C., & Rabinovic, A. (2007). Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics*, *8*(1), 118–127. https://doi.org/10.1093/biostatistics/kxj037

Jöhrens, K., Anagnostopoulos, I., Dürkop, H., & Stein, H. (2006). Different T-bet expression patterns characterize particular reactive lymphoid tissue lesions. *Histopathology*, *48*(4), 343–352. https://doi.org/10.1111/j.1365-2559.2005.02305.x

Kaminow, B., Yunusov, D., & Dobin, A. (2021). *STARsolo: Accurate, fast and versatile mapping/quantification of single-cell and single-nucleus RNA-seq data*. https://doi.org/10.1101/2021.05.05.442755

Kanbar, J. N., Ma, S., Kim, E. S., Kurd, N. S., Tsai, M. S., Tysl, T., Widjaja, C. E., Limary, A. E., Yee, B., He, Z., Hao, Y., Fu, X.-D., Yeo, G. W., Huang, W. J., & Chang, J. T. (2022). The long noncoding RNA Malat1 regulates CD8+ T cell differentiation by mediating epigenetic repression. *Journal of Experimental Medicine*, *219*(6), e20211756. https://doi.org/10.1084/jem.20211756

Kekäläinen, E., Tuovinen, H., Joensuu, J., Gylling, M., Franssila, R., Pöntynen, N., Talvensaari, K., Perheentupa, J., Miettinen, A., & Arstila, T. P. (2007). A Defect of Regulatory T Cells in Patients with Autoimmune Polyendocrinopathy-Candidiasis-Ectodermal Dystrophy. *The Journal of Immunology*, *178*(2), 1208–1215. https://doi.org/10.4049/jimmunol.178.2.1208

Kelley, J. B., Datta, S., Snow, C. J., Chatterjee, M., Ni, L., Spencer, A., Yang, C.-S., Cubeñas-Potts, C., Matunis, M. J., & Paschal, B. M. (2011). The Defective Nuclear Lamina in Hutchinson-Gilford Progeria Syndrome Disrupts the Nucleocytoplasmic Ran Gradient and Inhibits Nuclear Localization of Ubc9. *Molecular and Cellular Biology*, *31*(16), 3378–3395. https://doi.org/10.1128/MCB.05087-11

Kelly, J., Khan, A. A., Yin, J., Ferguson, T. A., & Apte, R. S. (2007). Senescence regulates macrophage activation and angiogenic fate at sites of tissue injury in mice. *Journal of Clinical Investigation*, *117*(11), 3421–3426. https://doi.org/10.1172/JCI32430

Kennedy, R. B., Ovsyannikova, I. G., Haralambieva, I. H., Oberg, A. L., Zimmermann, M. T., Grill, D. E., & Poland, G. A. (2016). Immunosenescence-Related Transcriptomic and Immunologic Changes in Older Individuals Following Influenza Vaccination. *Frontiers in Immunology*, *7*. https://doi.org/10.3389/fimmu.2016.00450

Kim, J., Kim, A.-R., & Shin, E.-C. (2015). Cytomegalovirus Infection and Memory T Cell Inflation. *Immune Network*, *15*(4), 186. https://doi.org/10.4110/in.2015.15.4.186

Kimmel, J. C., Penland, L., Rubinstein, N. D., Hendrickson, D. G., Kelley, D. R., & Rosenthal, A. Z. (2019). Murine single-cell RNA-seq reveals cell-identity- and tissue-specific trajectories of aging. *Genome Research*, *29*(12), 2088–2103. https://doi.org/10.1101/gr.253880.119

Kisand, K., & Peterson, P. (2011). Autoimmune polyendocrinopathy candidiasis ectodermal dystrophy: Known and novel aspects of the syndrome. *Annals of the New York Academy of Sciences*, *1246*(1), 77–91. https://doi.org/10.1111/j.1749-6632.2011.06308.x

Kiselev, V. Y., Andrews, T. S., & Hemberg, M. (2019). Challenges in unsupervised clustering of single-cell RNA-seq data. *Nature Reviews Genetics*, *20*(5), 273–282. https://doi.org/10.1038/s41576-018-0088-9

Kiselev, V. Y., Yiu, A., & Hemberg, M. (2018). scmap: Projection of single-cell RNA-seq data across data sets. *Nature Methods*, *15*(5), 359–362. https://doi.org/10.1038/nmeth.4644

Klein, A. M., Mazutis, L., Akartuna, I., Tallapragada, N., Veres, A., Li, V., Peshkin, L., Weitz, D. A., & Kirschner, M. W. (2015). Droplet Barcoding for Single-Cell Transcriptomics Applied to Embryonic Stem Cells. *Cell*, *161*(5), 1187–1201. https://doi.org/10.1016/j.cell.2015.04.044

Klein, L., & Petrozziello, E. (2024). Antigen presentation for central tolerance induction. *Nature Reviews Immunology*. https://doi.org/10.1038/s41577-024-01076-8

Klein, L., Robey, E. A., & Hsieh, C.-S. (2019). Central CD4+ T cell tolerance: Deletion versus regulatory T cell differentiation. *Nature Reviews Immunology*, *19*(1), 7–18. https://doi.org/10.1038/s41577-018-0083-6

Kolberg, L., Raudvere, U., Kuzmin, I., Adler, P., Vilo, J., & Peterson, H. (2023). g:Profiler—Interoperable web service for functional enrichment analysis and gene identifier mapping (2023 update). *Nucleic Acids Research*, *51*(W1), W207–W212. https://doi.org/10.1093/nar/gkad347

Konturek-Ciesla, A., Dhapola, P., Zhang, Q., Säwén, P., Wan, H., Karlsson, G., & Bryder, D. (2023). Temporal multimodal single-cell profiling of native hematopoiesis illuminates altered differentiation trajectories with age. *Cell Reports*, *42*(4), 112304. https://doi.org/10.1016/j.celrep.2023.112304

Korsunsky, I., Millard, N., Fan, J., Slowikowski, K., Zhang, F., Wei, K., Baglaenko, Y., Brenner, M., Loh, P., & Raychaudhuri, S. (2019). Fast, sensitive and accurate integration of single-cell data with Harmony. *Nature Methods*, *16*(12), 1289–1296. https://doi.org/10.1038/s41592-019-0619-0

Kuijpers, L., Hornung, B., Van Den Hout - Van Vroonhoven, M. C. G. N., Van IJcken, W. F. J., Grosveld, F., & Mulugeta, E. (2024). Split Pool Ligation-based Single-cell Transcriptome sequencing (SPLiT-seq) data processing pipeline comparison. *BMC Genomics*, *25*(1), 361. https://doi.org/10.1186/s12864-024-10285-3

Kumar, P. G., Laloraya, M., Wang, C.-Y., Ruan, Q.-G., Davoodi-Semiromi, A., Kao, K.-J., & She, J.-X. (2001). The Autoimmune Regulator (AIRE) Is a DNA-binding Protein. *Journal of Biological Chemistry*, *276*(44), 41357–41364. https://doi.org/10.1074/jbc.M104898200

L. Lun, A. T., Bach, K., & Marioni, J. C. (2016). Pooling across cells to normalize single-cell RNA sequencing data with many zero counts. *Genome Biology*, *17*(1), 75. https://doi.org/10.1186/s13059-016-0947-7

Laakso, S. M., Kekäläinen, E., Rossi, L. H., Laurinolli, T.-T., Mannerström, H., Heikkilä, N., Lehtoviita, A., Perheentupa, J., Jarva, H., & Arstila, T. P. (2011). IL-7 Dysregulation and Loss of CD8+ T Cell Homeostasis in the Monogenic Human Disease Autoimmune Polyendocrinopathy–Candidiasis–Ectodermal Dystrophy. *The Journal of Immunology*, *187*(4), 2023–2030. https://doi.org/10.4049/jimmunol.1100212

Laakso, S. M., Laurinolli, T.-T., Rossi, L. H., Lehtoviita, A., Sairanen, H., Perheentupa, J., Kekäläinen, E., & Arstila, T. P. (2010). Regulatory T cell defect in APECED

patients is associated with loss of naive FOXP3+ precursors and impaired activated population. *Journal of Autoimmunity*, *35*(4), 351–357. https://doi.org/10.1016/j.jaut.2010.07.008

Lam, V. C., & Lanier, L. L. (2017). NK cells in host responses to viral infections. *Current Opinion in Immunology*, *44*, 43–51. https://doi.org/10.1016/j.coi.2016.11.003

Lambolez, B., Audinat, E., Bochet, P., Crépel, F., & Rossier, J. (1992). AMPA receptor subunits expressed by single purkinje cells. *Neuron*, *9*(2), 247–258. https://doi.org/10.1016/0896-6273(92)90164-9

Lever, M., Maini, P. K., Van Der Merwe, P. A., & Dushek, O. (2014). Phenotypic models of T cell activation. *Nature Reviews Immunology*, *14*(9), 619–629. https://doi.org/10.1038/nri3728

Li, J., Xiong, M., Fu, X.-H., Fan, Y., Dong, C., Sun, X., Zheng, F., Wang, S.-W., Liu, L., Xu, M., Wang, C., Ping, J., Che, S., Wang, Q., Yang, K., Zuo, Y., Lu, X., Zheng, Z., Lan, T., … Liu, G.-H. (2023). Determining a multimodal aging clock in a cohort of Chinese women. *Med*, *4*(11), 825–848.e13. https://doi.org/10.1016/j.medj.2023.06.010

Li, S. S., Kochar, N. K., Elizaga, M., Hay, C. M., Wilson, G. J., Cohen, K. W., De Rosa, S. C., Xu, R., Ota-Setlik, A., Morris, D., Finak, G., Allen, M., Tieu, H.-V., Frank, I., Sobieszczyk, M. E., Hannaman, D., Gottardo, R., Gilbert, P. B., Tomaras, G. D., … NIAID HIV Vaccine Trials Network. (2017). DNA Priming Increases Frequency of T-Cell Responses to a Vesicular Stomatitis Virus HIV Vaccine with Specific Enhancement of CD8 $^{+}$ T-Cell Responses by Interleukin-12 Plasmid DNA. *Clinical and Vaccine Immunology*, *24*(11). https://doi.org/10.1128/CVI.00263-17

Li, Y., Li, K., Zhu, L., Li, B., Zong, D., Cai, P., Jiang, C., Du, P., Lin, J., & Qu, K. (2021). Development of double-positive thymocytes at single-cell resolution. *Genome Medicine*, *13*(1), 49. https://doi.org/10.1186/s13073-021-00861-7

Liiv, I., Rebane, A., Org, T., Saare, M., Maslovskaja, J., Kisand, K., Juronen, E., Valmu, L., Bottomley, M. J., Kalkkinen, N., & Peterson, P. (2008). DNA-PK contributes to the phosphorylation of AIRE: Importance in transcriptional activity. *Biochimica et Biophysica Acta (BBA) – Molecular Cell Research*, *1783*(1), 74–83. https://doi.org/10.1016/j.bbamcr.2007.09.003

Lipsky, M. S., & King, M. (2015). Biological theories of aging. *Disease-a-Month*, *61*(11), 460–466. https://doi.org/10.1016/j.disamonth.2015.09.005

Liu, Z., Liang, Q., Ren, Y., Guo, C., Ge, X., Wang, L., Cheng, Q., Luo, P., Zhang, Y., & Han, X. (2023). Immunosenescence: Molecular mechanisms and diseases. *Signal Transduction and Targeted Therapy*, *8*(1), 200. https://doi.org/10.1038/s41392-023-01451-2

Lopez, R., Regier, J., Cole, M. B., Jordan, M. I., & Yosef, N. (2018). Deep generative modeling for single-cell transcriptomics. *Nature Methods*, *15*(12), 1053–1058. https://doi.org/10.1038/s41592-018-0229-2

López-Otín, C., Blasco, M. A., Partridge, L., Serrano, M., & Kroemer, G. (2013). The Hallmarks of Aging. *Cell*, *153*(6), 1194–1217. https://doi.org/10.1016/j.cell.2013.05.039

López-Otín, C., Blasco, M. A., Partridge, L., Serrano, M., & Kroemer, G. (2023). Hallmarks of aging: An expanding universe. *Cell*, *186*(2), 243–278. https://doi.org/10.1016/j.cell.2022.11.001

López-Otín, C., & Kroemer, G. (2021). Hallmarks of Health. *Cell*, *184*(1), 33–63. https://doi.org/10.1016/j.cell.2020.11.034

Lotfollahi, M., Wolf, F. A., & Theis, F. J. (2019). scGen predicts single-cell perturbation responses. *Nature Methods*, *16*(8), 715–721. https://doi.org/10.1038/s41592-019-0494-8

Lu, A. T., Fei, Z., Haghani, A., Robeck, T. R., Zoller, J. A., Li, C. Z., Lowe, R., Yan, Q., Zhang, J., Vu, H., Ablaeva, J., Acosta-Rodriguez, V. A., Adams, D. M., Almunia, J., Aloysius, A., Ardehali, R., Arneson, A., Baker, C. S., Banks, G., … Horvath, S. (2023). Universal DNA methylation age across mammalian tissues. *Nature Aging*, *3*(9), 1144–1166. https://doi.org/10.1038/s43587-023-00462-6

Lu, A. T., Quach, A., Wilson, J. G., Reiner, A. P., Aviv, A., Raj, K., Hou, L., Baccarelli, A. A., Li, Y., Stewart, J. D., Whitsel, E. A., Assimes, T. L., Ferrucci, L., & Horvath, S. (2019). DNA methylation GrimAge strongly predicts lifespan and healthspan. *Aging*, *11*(2), 303–327. https://doi.org/10.18632/aging.101684

Lu, J., Ahmad, R., Nguyen, T., Cifello, J., Hemani, H., Li, J., Chen, J., Li, S., Wang, J., Achour, A., Chen, J., Colie, M., Lustig, A., Dunn, C., Zukley, L., Chia, C. W., Burd, I., Zhu, J., Ferrucci, L., & Weng, N. (2022). Heterogeneity and transcriptome changes of human CD8+ T cells across nine decades of life. *Nature Communications*, *13*(1), 5128. https://doi.org/10.1038/s41467-022-32869-x

Luecken, M. D., Büttner, M., Chaichoompu, K., Danese, A., Interlandi, M., Mueller, M. F., Strobl, D. C., Zappia, L., Dugas, M., Colomé-Tatché, M., & Theis, F. J. (2022). Benchmarking atlas-level data integration in single-cell genomics. *Nature Methods*, *19*(1), 41–50. https://doi.org/10.1038/s41592-021-01336-8

Lun, A. T. L., participants in the 1st Human Cell Atlas Jamboree, Riesenfeld, S., Andrews, T., Dao, T. P., Gomes, T., & Marioni, J. C. (2019). EmptyDrops: Distinguishing cells from empty droplets in droplet-based single-cell RNA sequencing data. *Genome Biology*, *20*(1), 63. https://doi.org/10.1186/s13059-019-1662-y

Luo, O. J., Lei, W., Zhu, G., Ren, Z., Xu, Y., Xiao, C., Zhang, H., Cai, J., Luo, Z., Gao, L., Su, J., Tang, L., Guo, W., Su, H., Zhang, Z.-J., Fang, E. F., Ruan, Y., Leng, S. X., Ju, Z., … Chen, G. (2022). Multidimensional single-cell analysis of human peripheral blood reveals characteristic features of the immune system landscape in aging and frailty. *Nature Aging*, *2*(4), 348–364. https://doi.org/10.1038/s43587-022-00198-9

Macosko, E. Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., Tirosh, I., Bialas, A. R., Kamitaki, N., Martersteck, E. M., Trombetta, J. J., Weitz, D. A., Sanes, J. R., Shalek, A. K., Regev, A., & McCarroll, S. A. (2015). Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell*, *161*(5), 1202–1214. https://doi.org/10.1016/j.cell.2015.05.002

Madissoon, E., Wilbrey-Clark, A., Miragaia, R. J., Saeb-Parsy, K., Mahbubani, K. T., Georgakopoulos, N., Harding, P., Polanski, K., Huang, N., Nowicki-Osuch, K., Fitzgerald, R. C., Loudon, K. W., Ferdinand, J. R., Clatworthy, M. R., Tsingene, A., Van Dongen, S., Dabrowska, M., Patel, M., Stubbington, M. J. T., … Meyer, K. B. (2020). scRNA-seq assessment of the human lung, spleen, and esophagus tissue stability after cold preservation. *Genome Biology*, *21*(1), 1. https://doi.org/10.1186/s13059-019-1906-x

Malchow, S., Leventhal, D. S., Lee, V., Nishi, S., Socci, N. D., & Savage, P. A. (2016). Aire Enforces Immune Tolerance by Directing Autoreactive T Cells into the Regulatory T Cell Lineage. *Immunity*, *44*(5), 1102–1113. https://doi.org/10.1016/j.immuni.2016.02.009

Marra, F., & Tacke, F. (2014). Roles for Chemokines in Liver Disease. *Gastroenterology*, *147*(3), 577-594.e1. https://doi.org/10.1053/j.gastro.2014.06.043

Marsh, S. E., Walker, A. J., Kamath, T., Dissing-Olesen, L., Hammond, T. R., De Soysa, T. Y., Young, A. M. H., Murphy, S., Abdulraouf, A., Nadaf, N., Dufort, C., Walker, A. C., Lucca, L. E., Kozareva, V., Vanderburg, C., Hong, S., Bulstrode, H., Hutchinson, P. J., Gaffney, D. J., … Stevens, B. (2022). Dissection of artifactual and confounding glial signatures by single-cell sequencing of mouse and human brain. *Nature Neuroscience*, *25*(3), 306–316. https://doi.org/10.1038/s41593-022-01022-8

McGinnis, C. S., Patterson, D. M., Winkler, J., Conrad, D. N., Hein, M. Y., Srivastava, V., Hu, J. L., Murrow, L. M., Weissman, J. S., Werb, Z., Chow, E. D., & Gartner, Z. J. (2019). MULTI-seq: Sample multiplexing for single-cell RNA sequencing using lipid-tagged indices. *Nature Methods*, *16*(7), 619–626. https://doi.org/10.1038/s41592-019-0433-8

McGrath, S., Grimstad, K., Thorarinsdottir, K., Forslind, K., Glinatsi, D., Leu Agelii, M., Aranburu, A., Sundell, T., Jonsson, C. A., Camponeschi, A., Hultgård Ekwall, A., Tilevik, A., Gjertsson, I., & Mårtensson, I. (2024). Correlation of Professional Antigen-Presenting Tbet^{+} CD11c^{+} B Cells With Bone Destruction in Untreated Rheumatoid Arthritis. *Arthritis & Rheumatology*, *76*(8), 1263–1277. https://doi.org/10.1002/art.42857

McInnes, L., Healy, J., & Melville, J. (2018). *UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction* (Version 3). arXiv. https://doi.org/10.48550/ARXIV.1802.03426

McIntyre, L. M., Lopiano, K. K., Morse, A. M., Amin, V., Oberg, A. L., Young, L. J., & Nuzhdin, S. V. (2011). RNA-seq: Technical variability and sampling. *BMC Genomics*, *12*(1), 293. https://doi.org/10.1186/1471-2164-12-293

McKenna, R. W., Washington, L. T., Aquino, D. B., Picker, L. J., & Kroft, S. H. (2001). Immunophenotypic analysis of hematogones (B-lymphocyte precursors) in 662 consecutive bone marrow specimens by 4-color flow cytometry. *Blood*, *98*(8), 2498–2507. https://doi.org/10.1182/blood.V98.8.2498

McQuattie-Pimentel, A. C., Ren, Z., Joshi, N., Watanabe, S., Stoeger, T., Chi, M., Lu, Z., Sichizya, L., Aillon, R. P., Chen, C.-I., Soberanes, S., Chen, Z., Reyfman, P. A., Walter, J. M., Anekalla, K. R., Davis, J. M., Helmin, K. A., Runyan, C. E., Abdala-Valencia, H., … Budinger, G. R. S. (2021). The lung microenvironment shapes a dysfunctional response of alveolar macrophages in aging. *Journal of Clinical Investigation*, *131*(4), e140299. https://doi.org/10.1172/JCI140299

Medvedev, Zh. A. (1972). Repetition of molecular-genetic information as a possible factor in evolutionary changes of life span. *Experimental Gerontology*, *7*(4), 227–238. https://doi.org/10.1016/0531-5565(72)90012-5

Metkar, S. S., Menaa, C., Pardo, J., Wang, B., Wallich, R., Freudenberg, M., Kim, S., Raja, S. M., Shi, L., Simon, M. M., & Froelich, C. J. (2008). Human and Mouse Granzyme A Induce a Proinflammatory Cytokine Response. *Immunity*, *29*(5), 720–733. https://doi.org/10.1016/j.immuni.2008.08.014

Miao, Z., Moreno, P., Huang, N., Papatheodorou, I., Brazma, A., & Teichmann, S. A. (2020). Putative cell type discovery from single-cell gene expression data. *Nature Methods*, *17*(6), 621–628. https://doi.org/10.1038/s41592-020-0825-9

Migueles, S. A., Laborico, A. C., Shupert, W. L., Sabbaghian, M. S., Rabin, R., Hallahan, C. W., Baarle, D. V., Kostense, S., Miedema, F., McLaughlin, M., Ehler, L., Metcalf, J., Liu, S., & Connors, M. (2002). HIV-specific CD8+ T cell proliferation is coupled to perforin expression and is maintained in nonprogressors. *Nature Immunology*, *3*(11), 1061–1068. https://doi.org/10.1038/ni845

Mittelbrunn, M., & Kroemer, G. (2021). Hallmarks of T cell aging. *Nature Immunology*, *22*(6), 687–698. https://doi.org/10.1038/s41590-021-00927-z

Mogilenko, D. A., Shchukina, I., & Artyomov, M. N. (2022). Immune ageing at single-cell resolution. *Nature Reviews Immunology*, *22*(8), 484–498. https://doi.org/10.1038/s41577-021-00646-4

Mogilenko, D. A., Shpynov, O., Andhey, P. S., Arthur, L., Swain, A., Esaulova, E., Brioschi, S., Shchukina, I., Kerndl, M., Bambouskova, M., Yao, Z., Laha, A., Zaitsev, K., Burdess, S., Gillfilan, S., Stewart, S. A., Colonna, M., & Artyomov, M. N. (2021). Comprehensive Profiling of an Aging Immune System Reveals Clonal GZMK+ CD8+ T Cells as Conserved Hallmark of Inflammaging. *Immunity*, *54*(1), 99-115.e12. https://doi.org/10.1016/j.immuni.2020.11.005

Mohr, A., Malhotra, R., Mayer, G., Gorochov, G., & Miyara, M. (2018). Human FOXP 3^+ T regulatory cell heterogeneity. *Clinical & Translational Immunology*, *7*(1), e1005. https://doi.org/10.1002/cti2.1005

Montecino-Rodriguez, E., Berent-Maoz, B., & Dorshkind, K. (2013). Causes, consequences, and reversal of immune system aging. *Journal of Clinical Investigation*, *123*(3), 958–965. https://doi.org/10.1172/JCI64096

Montserrat-Ayuso, T., & Esteve-Codina, A. (2024). *Revealing the Prevalence of Suboptimal Cells and Organs in Reference Cell Atlases: An Imperative for Enhanced Quality Control*. https://doi.org/10.1101/2024.04.18.590104

Morris, S. A. (2019). The evolving concept of cell identity in the single cell era. *Development*, *146*(12), dev169748. https://doi.org/10.1242/dev.169748

Murray, P. J., & Wynn, T. A. (2011). Protective and pathogenic functions of macrophage subsets. *Nature Reviews Immunology*, *11*(11), 723–737. https://doi.org/10.1038/nri3073

Muskovic, W., & Powell, J. E. (2021). DropletQC: Improved identification of empty droplets and damaged cells in single-cell RNA-seq data. *Genome Biology*, *22*(1), 329. https://doi.org/10.1186/s13059-021-02547-0

Nagamine, K., Peterson, P., Scott, H. S., Kudoh, J., Minoshima, S., Heino, M., Krohn, K. J. E., Lalioti, M. D., Mullis, P. E., Antonarakis, S. E., Kawasaki, K., Asakawa, S., Ito, F., & Shimizu, N. (1997). Positional cloning of the APECED gene. *Nature Genetics*, *17*(4), 393–398. https://doi.org/10.1038/ng1297-393

Naradikian, M. S., Myles, A., Beiting, D. P., Roberts, K. J., Dawson, L., Herati, R. S., Bengsch, B., Linderman, S. L., Stelekati, E., Spolski, R., Wherry, E. J., Hunter, C., Hensley, S. E., Leonard, W. J., & Cancro, M. P. (2016). Cutting Edge: IL-4, IL-21, and IFN-γ Interact To Govern T-bet and CD11c Expression in TLR-Activated B Cells. *The Journal of Immunology*, *197*(4), 1023–1028. https://doi.org/10.4049/jimmunol.1600522

Nickerson, K. M., Smita, S., Hoehn, K. B., Marinov, A. D., Thomas, K. B., Kos, J. T., Yang, Y., Bastacky, S. I., Watson, C. T., Kleinstein, S. H., & Shlomchik, M. J. (2023). Age-associated B cells are heterogeneous and dynamic drivers of autoimmunity in mice. *Journal of Experimental Medicine*, *220*(5), e20221346. https://doi.org/10.1084/jem.20221346

Nothelfer, K., Sansonetti, P. J., & Phalipon, A. (2015). Pathogen manipulation of B cells: The best defence is a good offence. *Nature Reviews Microbiology*, *13*(3), 173–184. https://doi.org/10.1038/nrmicro3415

Novella-Rausell, C., Grudniewska, M., Peters, D. J. M., & Mahfouz, A. (2023). A comprehensive mouse kidney atlas enables rare cell population characterization and

robust marker discovery. *iScience*, *26*(6), 106877. https://doi.org/10.1016/j.isci.2023.106877

Okin, D., & Medzhitov, R. (2012). Evolution of Inflammatory Diseases. *Current Biology*, *22*(17), R733–R740. https://doi.org/10.1016/j.cub.2012.07.029

Orgel, L. E. (1963). The maintenance of the accuracy of protein synthesis and its relevance to ageing. *Proceedings of the National Academy of Sciences*, *49*(4), 517–521. https://doi.org/10.1073/pnas.49.4.517

Orlova, E. M., Sozaeva, L. S., Kareva, M. A., Oftedal, B. E., Wolff, A. S. B., Breivik, L., Zakharova, E. Y., Ivanova, O. N., Kämpe, O., Dedov, I. I., Knappskog, P. M., Peterkova, V. A., & Husebye, E. S. (2017). Expanding the Phenotypic and Genotypic Landscape of Autoimmune Polyendocrine Syndrome Type 1. *The Journal of Clinical Endocrinology & Metabolism*, *102*(9), 3546–3556. https://doi.org/10.1210/jc.2017-00139

Ossart, J., Moreau, A., Autrusseau, E., Ménoret, S., Martin, J. C., Besnard, M., Ouisse, L.-H., Tesson, L., Flippe, L., Kisand, K., Peterson, P., Hubert, F.-X., Anegon, I., Josien, R., & Guillonneau, C. (2018). Breakdown of Immune Tolerance in AIRE-Deficient Rats Induces a Severe Autoimmune Polyendocrinopathy–Candidiasis–Ectodermal Dystrophy – like Autoimmune Disease. *The Journal of Immunology*, *201*(3), 874–887. https://doi.org/10.4049/jimmunol.1701318

Oven, I., Brdičková, N., Kohoutek, J., Vaupotič, T., Narat, M., & Peterlin, B. M. (2007). AIRE Recruits P-TEFb for Transcriptional Elongation of Target Genes in Medullary Thymic Epithelial Cells. *Molecular and Cellular Biology*, *27*(24), 8815–8823. https://doi.org/10.1128/MCB.01085-07

Palmer, S., Albergante, L., Blackburn, C. C., & Newman, T. J. (2018). Thymic involution and rising disease incidence with age. *Proceedings of the National Academy of Sciences*, *115*(8), 1883–1888. https://doi.org/10.1073/pnas.1714478115

Pang, W. W., Price, E. A., Sahoo, D., Beerman, I., Maloney, W. J., Rossi, D. J., Schrier, S. L., & Weissman, I. L. (2011). Human bone marrow hematopoietic stem cells are increased in frequency and myeloid-biased with age. *Proceedings of the National Academy of Sciences*, *108*(50), 20012–20017. https://doi.org/10.1073/pnas.1116110108

Pereira, B. I., De Maeyer, R. P. H., Covre, L. P., Nehar-Belaid, D., Lanna, A., Ward, S., Marches, R., Chambers, E. S., Gomes, D. C. O., Riddell, N. E., Maini, M. K., Teixeira, V. H., Janes, S. M., Gilroy, D. W., Larbi, A., Mabbott, N. A., Ucar, D., Kuchel, G. A., Henson, S. M., … Akbar, A. N. (2020). Sestrins induce natural killer function in senescent-like CD8+ T cells. *Nature Immunology*, *21*(6), 684–694. https://doi.org/10.1038/s41590-020-0643-3

Peters, M. J., Joehanes, R., Pilling, L. C., Schurmann, C., Conneely, K. N., Powell, J., Reinmaa, E., Sutphin, G. L., Zhernakova, A., Schramm, K., Wilson, Y. A., Kobes, S., Tukiainen, T., NABEC/UKBEC Consortium, Nalls, M. A., Hernandez, D. G., Cookson, M. R., Gibbs, R. J., Hardy, J., … Johnson, A. D. (2015). The transcriptional landscape of age in human peripheral blood. *Nature Communications*, *6*(1), 8570. https://doi.org/10.1038/ncomms9570

Peterson, P., & Peltonen, L. (2005). Autoimmune polyendocrinopathy syndrome type 1 (APS1) and AIRE gene: New views on molecular basis of autoimmunity. *Journal of Autoimmunity*, *25*, 49–55. https://doi.org/10.1016/j.jaut.2005.09.022

Pitkänen, J., Doucas, V., Sternsdorf, T., Nakajima, T., Aratani, S., Jensen, K., Will, H., Vähämurto, P., Ollila, J., Vihinen, M., Scott, H. S., Antonarakis, S. E., Kudoh, J., Shimizu, N., Krohn, K., & Peterson, P. (2000). The Autoimmune Regulator Protein Has Transcriptional Transactivating Properties and Interacts with the Common

Coactivator CREB-binding Protein. *Journal of Biological Chemistry*, *275*(22), 16802–16809. https://doi.org/10.1074/jbc.M908944199

Pitkänen, J., Rebane, A., Rowell, J., Murumägi, A., Ströbel, P., Möll, K., Saare, M., Heikkilä, J., Doucas, V., Marx, A., & Peterson, P. (2005). Cooperative activation of transcription by autoimmune regulator AIRE and CBP. *Biochemical and Biophysical Research Communications*, *333*(3), 944–953. https://doi.org/10.1016/j.bbrc.2005.05.187

Potter, S. S. (2018). Single-cell RNA sequencing for the study of development, physiology and disease. *Nature Reviews Nephrology*, *14*(8), 479–492. https://doi.org/10.1038/s41581-018-0021-7

Puntambekar, S., Hesselberth, J. R., Riemondy, K. A., & Fu, R. (2021). Cell-level metadata are indispensable for documenting single-cell sequencing datasets. *PLOS Biology*, *19*(5), e3001077. https://doi.org/10.1371/journal.pbio.3001077

Purohit, S., Kumar, P. G., Laloraya, M., & She, J.-X. (2005). Mapping DNA-binding domains of the autoimmune regulator protein. *Biochemical and Biophysical Research Communications*, *327*(3), 939–944. https://doi.org/10.1016/j.bbrc.2004.12.093

Quinn, K. M., Fox, A., Harland, K. L., Russ, B. E., Li, J., Nguyen, T. H. O., Loh, L., Olshanksy, M., Naeem, H., Tsyganov, K., Wiede, F., Webster, R., Blyth, C., Sng, X. Y. X., Tiganis, T., Powell, D., Doherty, P. C., Turner, S. J., Kedzierska, K., & La Gruta, N. L. (2018). Age-Related Decline in Primary CD8+ T Cell Responses Is Associated with the Development of Senescence in Virtual Memory CD8+ T Cells. *Cell Reports*, *23*(12), 3512–3524. https://doi.org/10.1016/j.celrep.2018.05.057

Regev, A., Teichmann, S. A., Lander, E. S., Amit, I., Benoist, C., Birney, E., Bodenmiller, B., Campbell, P., Carninci, P., Clatworthy, M., Clevers, H., Deplancke, B., Dunham, I., Eberwine, J., Eils, R., Enard, W., Farmer, A., Fugger, L., Göttgens, B., … the Human Cell Atlas Meeting Participants. (2017). *The Human Cell Atlas*. https://doi.org/10.1101/121202

Reitsema, R. D., Kumawat, A. K., Hesselink, B.-C., Van Baarle, D., & Van Sleen, Y. (2024). Effects of ageing and frailty on circulating monocyte and dendritic cell subsets. *Npj Aging*, *10*(1), 17. https://doi.org/10.1038/s41514-024-00144-6

Ripa, M., Chiappetta, S., & Tambussi, G. (2017). Immunosenescence and hurdles in the clinical management of older HIV-patients. *Virulence*, *8*(5), 508–528. https://doi.org/10.1080/21505594.2017.1292197

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., & Smyth, G. K. (2015). Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research*, *43*(7), e47–e47. https://doi.org/10.1093/nar/gkv007

Rood, J. E., Wynne, S., Robson, L., Hupalowska, A., Randell, J., Teichmann, S. A., & Regev, A. (2024). The Human Cell Atlas from a cell census to a unified foundation model. *Nature*. https://doi.org/10.1038/s41586-024-08338-4

Rubtsov, A. V., Rubtsova, K., Fischer, A., Meehan, R. T., Gillis, J. Z., Kappler, J. W., & Marrack, P. (2011). Toll-like receptor 7 (TLR7)–driven accumulation of a novel CD11c+ B-cell population is important for the development of autoimmunity. *Blood*, *118*(5), 1305–1315. https://doi.org/10.1182/blood-2011-01-331462

Russell Knode, L. M., Naradikian, M. S., Myles, A., Scholz, J. L., Hao, Y., Liu, D., Ford, M. L., Tobias, J. W., Cancro, M. P., & Gearhart, P. J. (2017). Age-Associated B Cells Express a Diverse Repertoire of VH and Vκ Genes with Somatic Hypermutation. *The Journal of Immunology*, *198*(5), 1921–1927. https://doi.org/10.4049/jimmunol.1601106

Saelens, W., Cannoodt, R., Todorov, H., & Saeys, Y. (2019). A comparison of single-cell trajectory inference methods. *Nature Biotechnology*, *37*(5), 547–554. https://doi.org/10.1038/s41587-019-0071-9

Samten, B. (2013). CD52 as both a marker and an effector molecule of T cells with regulatory action: Identification of novel regulatory T cells. *Cellular & Molecular Immunology*, *10*(6), 456–458. https://doi.org/10.1038/cmi.2013.38

Sanderson, S. L., & Simon, A. K. (2017). In aged primary T cells, mitochondrial stress contributes to telomere attrition measured by a novel imaging flow cytometry assay. *Aging Cell*, *16*(6), 1234–1243. https://doi.org/10.1111/acel.12640

Sarfraz, I., Wang, Y., Shastry, A., Teh, W. K., Sokolov, A., Herb, B. R., Creasy, H. H., Virshup, I., Dries, R., Degatano, K., Mahurkar, A., Schnell, D. J., Madrigal, P., Hilton, J., Gehlenborg, N., Tickle, T., & Campbell, J. D. (2024). MAMS: Matrix and analysis metadata standards to facilitate harmonization and reproducibility of single-cell data. *Genome Biology*, *25*(1), 205. https://doi.org/10.1186/s13059-024-03349-w

Sauce, D., Larsen, M., Leese, A. M., Millar, D., Khan, N., Hislop, A. D., & Rickinson, A. B. (2007). IL-7Rα versus CCR7 and CD45 as Markers of Virus-Specific $CD8^+$ T Cell Differentiation: Contrasting Pictures in Blood and Tonsillar Lymphoid Tissue. *The Journal of Infectious Diseases*, *195*(2), 268–278. https://doi.org/10.1086/510248

Saxena, V., Piao, W., Li, L., Paluskievicz, C., Xiong, Y., Simon, T., Lakhan, R., Brinkman, C. C., Walden, S., Hippen, K. L., WillsonShirkey, M., Lee, Y. S., Wagner, C., Blazar, B. R., & Bromberg, J. S. (2022). Treg tissue stability depends on lymphotoxin beta-receptor- and adenosine-receptor-driven lymphatic endothelial cell responses. *Cell Reports*, *39*(3), 110727. https://doi.org/10.1016/j.celrep.2022.110727

Schmiedel, B. J., Gonzalez-Colin, C., Fajardo, V., Rocha, J., Madrigal, A., Ramírez-Suástegui, C., Bhattacharyya, S., Simon, H., Greenbaum, J. A., Peters, B., Seumois, G., Ay, F., Chandra, V., & Vijayanand, P. (2022). Single-cell eQTL analysis of activated T cell subsets reveals activation and cell type–dependent effects of disease-risk variants. *Science Immunology*, *7*(68), eabm2508. https://doi.org/10.1126/sciimmunol.abm2508

Schober, K., Buchholz, V. R., & Busch, D. H. (2018). TCR repertoire evolution during maintenance of CMV -specific T-cell populations. *Immunological Reviews*, *283*(1), 113–128. https://doi.org/10.1111/imr.12654

See, P., Lum, J., Chen, J., & Ginhoux, F. (2018). A Single-Cell Sequencing Guide for Immunologists. *Frontiers in Immunology*, *9*, 2425. https://doi.org/10.3389/fimmu.2018.02425

Shimasaki, N., Jain, A., & Campana, D. (2020). NK cells for cancer immunotherapy. *Nature Reviews Drug Discovery*, *19*(3), 200–218. https://doi.org/10.1038/s41573-019-0052-1

Sielemann, K., Hafner, A., & Pucker, B. (2020). The reuse of public datasets in the life sciences: Potential risks and rewards. *PeerJ*, *8*, e9954. https://doi.org/10.7717/peerj.9954

Sikkema, L., Ramírez-Suástegui, C., Strobl, D. C., Gillett, T. E., Zappia, L., Madissoon, E., Markov, N. S., Zaragosi, L.-E., Ji, Y., Ansari, M., Arguel, M.-J., Apperloo, L., Banchero, M., Bécavin, C., Berg, M., Chichelnitskiy, E., Chung, M., Collin, A., Gay, A. C. A., … Theis, F. J. (2023). An integrated cell atlas of the lung in health and disease. *Nature Medicine*, *29*(6), 1563–1577. https://doi.org/10.1038/s41591-023-02327-2

Simmons, A. J., & Lau, K. S. (2022). Dissociation and inDrops microfluidic encapsulation of human gut tissues for single-cell atlasing studies. *STAR Protocols*, *3*(3), 101570. https://doi.org/10.1016/j.xpro.2022.101570

Simon, A. K., Hollander, G. A., & McMichael, A. (2015). Evolution of the immune system in humans from infancy to old age. *Proceedings of the Royal Society B: Biological Sciences*, *282*(1821), 20143085. https://doi.org/10.1098/rspb.2014.3085

Sin, J. H., Sucharov, J., Kashyap, S., Wang, Y., Proekt, I., Liu, X., Parent, A. V., Gupta, A., Kastner, P., Chan, S., Gardner, J. M., Ntranos, V., Miller, C. N., Anderson, M. S., Schjerven, H., & Waterfield, M. R. (2023). Ikaros is a principal regulator of Aire^{+} mTEC homeostasis, thymic mimetic cell diversity, and central tolerance. *Science Immunology*, *8*(88), eabq3109. https://doi.org/10.1126/sciimmunol.abq3109

Sng, J., Ayoglu, B., Chen, J. W., Schickel, J.-N., Ferre, E. M. N., Glauzy, S., Romberg, N., Hoenig, M., Cunningham-Rundles, C., Utz, P. J., Lionakis, M. S., & Meffre, E. (2019). AIRE expression controls the peripheral selection of autoreactive B cells. *Science Immunology*, *4*(34), eaav6778. https://doi.org/10.1126/sciimmunol.aav6778

Snyder, C. M., Cho, K. S., Bonnett, E. L., Allan, J. E., & Hill, A. B. (2011). Sustained CD8+ T Cell Memory Inflation after Infection with a Single-Cycle Cytomegalovirus. *PLoS Pathogens*, *7*(10), e1002295. https://doi.org/10.1371/journal.ppat.1002295

Solana, R., Campos, C., Pera, A., & Tarazona, R. (2014). Shaping of NK cell subsets by aging. *Current Opinion in Immunology*, *29*, 56–61. https://doi.org/10.1016/j.coi.2014.04.002

Sousa Cardoso, R., Magalhães, D. A. R., Baião, A. M. T., Junta, C. M., Macedo, C., Marques, M. M. C., Sakamoto-Hojo, E. T., Donadi, E. A., & Passos, G. A. S. (2006). Onset of promiscuous gene expression in murine fetal thymus organ culture. *Immunology*, *119*(3), 369–375. https://doi.org/10.1111/j.1365-2567.2006.02441.x

Sprent, J., & Surh, C. D. (2011). Normal T cell homeostasis: The conversion of naive cells into memory-phenotype cells. *Nature Immunology*, *12*(6), 478–484. https://doi.org/10.1038/ni.2018

Stebegg, M., Bignon, A., Hill, D. L., Silva-Cayetano, A., Krueger, C., Vanderleyden, I., Innocentin, S., Boon, L., Wang, J., Zand, M. S., Dooley, J., Clark, J., Liston, A., Carr, E., & Linterman, M. A. (2020). Rejuvenating conventional dendritic cells and T follicular helper cell formation after vaccination. *eLife*, *9*, e52473. https://doi.org/10.7554/eLife.52473

Stoeckius, M., Zheng, S., Houck-Loomis, B., Hao, S., Yeung, B. Z., Mauck, W. M., Smibert, P., & Satija, R. (2018). Cell Hashing with barcoded antibodies enables multiplexing and doublet detection for single cell genomics. *Genome Biology*, *19*(1), 224. https://doi.org/10.1186/s13059-018-1603-1

Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W. M., Hao, Y., Stoeckius, M., Smibert, P., & Satija, R. (2019). Comprehensive Integration of Single-Cell Data. *Cell*, *177*(7), 1888-1902.e21. https://doi.org/10.1016/j.cell.2019.05.031

Stubbington, M. J. T., Rozenblatt-Rosen, O., Regev, A., & Teichmann, S. A. (2017). Single-cell transcriptomics to explore the immune system in health and disease. *Science*, *358*(6359), 58–63. https://doi.org/10.1126/science.aan6828

Sullivan, D. K., Min, K. H., Hjörleifsson, K. E., Luebbert, L., Holley, G., Moses, L., Gustafsson, J., Bray, N. L., Pimentel, H., Booeshaghi, A. S., Melsted, P., & Pachter, L. (2024). Kallisto, bustools and kb-python for quantifying bulk, single-cell and single-nucleus RNA-seq. *Nature Protocols*. https://doi.org/10.1038/s41596-024-01057-0

Sun, B., Bugarin-Estrada, E., Overend, L. E., Walker, C. E., Tucci, F. A., & Bashford-Rogers, R. J. M. (2021). Double-jeopardy: scRNA-seq doublet/multiplet detection using multi-omic profiling. *Cell Reports Methods*, *1*(1), 100008. https://doi.org/10.1016/j.crmeth.2021.100008

Sun, K., Xu, R., Ma, F., Yang, N., Li, Y., Sun, X., Jin, P., Kang, W., Jia, L., Xiong, J., Hu, H., Tian, Y., & Lan, X. (2022). scRNA-seq of gastric tumor shows complex intercellular interaction with an alternative T cell exhaustion trajectory. *Nature Communications*, *13*(1), 4943. https://doi.org/10.1038/s41467-022-32627-z

Svensson, V., Vento-Tormo, R., & Teichmann, S. A. (2018). Exponential scaling of single-cell RNA-seq in the past decade. *Nature Protocols*, *13*(4), 599–604. https://doi.org/10.1038/nprot.2017.149

Szilard, L. (1959). ON THE NATURE OF THE AGING PROCESS. *Proceedings of the National Academy of Sciences*, *45*(1), 30–45. https://doi.org/10.1073/pnas.45.1.30

Tang, F., Barbacioru, C., Wang, Y., Nordman, E., Lee, C., Xu, N., Wang, X., Bodeau, J., Tuch, B. B., Siddiqui, A., Lao, K., & Surani, M. A. (2009). mRNA-Seq whole-transcriptome analysis of a single cell. *Nature Methods*, *6*(5), 377–382. https://doi.org/10.1038/nmeth.1315

Terekhova, M., Swain, A., Bohacova, P., Aladyeva, E., Arthur, L., Laha, A., Mogilenko, D. A., Burdess, S., Sukhov, V., Kleverov, D., Echalar, B., Tsurinov, P., Chernyatchik, R., Husarcikova, K., & Artyomov, M. N. (2023). Single-cell atlas of healthy human blood unveils age-related loss of NKG2C+GZMB−CD8+ memory T cells and accumulation of type 2 memory T cells. *Immunity*, *56*(12), 2836-2854.e9. https://doi.org/10.1016/j.immuni.2023.10.013

The Tabula Muris Consortium, Almanzar, N., Antony, J., Baghel, A. S., Bakerman, I., Bansal, I., Barres, B. A., Beachy, P. A., Berdnik, D., Bilen, B., Brownfield, D., Cain, C., Chan, C. K. F., Chen, M. B., Clarke, M. F., Conley, S. D., Darmanis, S., Demers, A., Demir, K., … Zou, J. (2020). A single-cell transcriptomic atlas characterizes ageing tissues in the mouse. *Nature*, *583*(7817), 590–595. https://doi.org/10.1038/s41586-020-2496-1

The Tabula Muris Consortium, Overall coordination, Logistical coordination, Organ collection and processing, Library preparation and sequencing, Computational data analysis, Cell type annotation, Writing group, Supplemental text writing group, & Principal investigators. (2018). Single-cell transcriptomics of 20 mouse organs creates a Tabula Muris. *Nature*, *562*(7727), 367–372. https://doi.org/10.1038/s41586-018-0590-4

Thome, J. J. C., Grinshpun, B., Kumar, B. V., Kubota, M., Ohmura, Y., Lerner, H., Sempowski, G. D., Shen, Y., & Farber, D. L. (2016). Long-term maintenance of human naïve T cells through in situ homeostasis in lymphoid tissue sites. *Science Immunology*, *1*(6), eaah6506. https://doi.org/10.1126/sciimmunol.aah6506

Thomson, Z., He, Z., Swanson, E., Henderson, K., Phalen, C., Zaim, S. R., Pebworth, M.-P., Okada, L. Y., Heubeck, A. T., Roll, C. R., Hernandez, V., Weiss, M., Genge, P. C., Reading, J., Giles, J. R., Manne, S., Dougherty, J., Jasen, C. J., Greenplate, A. R., … Gustafson, C. E. (2023). Trimodal single-cell profiling reveals a novel pediatric CD8αα+ T cell subset and broad age-related molecular reprogramming across the T cell compartment. *Nature Immunology*, *24*(11), 1947–1959. https://doi.org/10.1038/s41590-023-01641-8

Timmons, J. A., & Brenner, C. (2024). The information theory of aging has not been tested. *Cell*, *187*(5), 1101–1102. https://doi.org/10.1016/j.cell.2024.01.013

Traag, V. A., Waltman, L., & Van Eck, N. J. (2019). From Louvain to Leiden: Gua ranteeing well-connected communities. *Scientific Reports*, *9*(1), 5233. https://doi.org/10.1038/s41598-019-41695-z

Tserel, L., Kolde, R., Limbach, M., Tretyakov, K., Kasela, S., Kisand, K., Saare, M., Vilo, J., Metspalu, A., Milani, L., & Peterson, P. (2015). Age-related profiling of DNA

methylation in CD8+ T cells reveals changes in immune response and transcriptional regulator genes. *Scientific Reports*, *5*(1), 13107. https://doi.org/10.1038/srep13107

Tung, P.-Y., Blischak, J. D., Hsiao, C. J., Knowles, D. A., Burnett, J. E., Pritchard, J. K., & Gilad, Y. (2017). Batch effects and the effective design of single-cell gene expression studies. *Scientific Reports*, *7*(1), 39921. https://doi.org/10.1038/srep39921

Turvey, S. E., & Broide, D. H. (2010). Innate immunity. *Journal of Allergy and Clinical Immunology*, *125*(2), S24–S32. https://doi.org/10.1016/j.jaci.2009.07.016

Utzschneider, D. T., Gabriel, S. S., Chisanga, D., Gloury, R., Gubser, P. M., Vasanthakumar, A., Shi, W., & Kallies, A. (2020). Early precursor T cells establish and propagate T cell exhaustion in chronic infection. *Nature Immunology*, *21*(10), 1256–1266. https://doi.org/10.1038/s41590-020-0760-z

Vallejos, C. A., Risso, D., Scialdone, A., Dudoit, S., & Marioni, J. C. (2017). Normalizing single-cell RNA sequencing data: Challenges and opportunities. *Nature Methods*, *14*(6), 565–571. https://doi.org/10.1038/nmeth.4292

Van Den Broek, T., Borghans, J. A. M., & Van Wijk, F. (2018). The full spectrum of human naive T cells. *Nature Reviews Immunology*, *18*(6), 363–373. https://doi.org/10.1038/s41577-018-0001-y

Vemulawada, C., Renavikar, P. S., Crawford, M. P., Steward-Tharp, S., & Karandikar, N. J. (2024). Disruption of *IFN* γ, *GZM* B, *PRF* 1, or *LYST* Results in Reduced Suppressive Function in Human CD8+ T Cells. *The Journal of Immunology*, *212*(11), 1722–1732. https://doi.org/10.4049/jimmunol.2300388

Walford, R. L. (1964). The Immunologic Theory of Aging. *The Gerontologist*, *4*(4), 195–197. https://doi.org/10.1093/geront/4.4.195

Wang, X., He, Y., Zhang, Q., Ren, X., & Zhang, Z. (2021). Direct Comparative Analyses of 10X Genomics Chromium and Smart-Seq2. *Genomics, Proteomics & Bioinformatics*, *19*(2), 253–266. https://doi.org/10.1016/j.gpb.2020.02.005

Wang, Z., Ding, H., & Zou, Q. (2020). Identifying cell types to interpret scRNA-seq data: How, why and more possibilities. *Briefings in Functional Genomics*, *19*(4), 286–291. https://doi.org/10.1093/bfgp/elaa003

Wilbrey-Clark, A., Roberts, K., & Teichmann, S. A. (2020). Cell Atlas technologies and insights into tissue architecture. *Biochemical Journal*, *477*(8), 1427–1442. https://doi.org/10.1042/BCJ20190341

Williams, G. C. (1957). PLEIOTROPY, NATURAL SELECTION, AND THE EVOLUTION OF SENESCENCE. *Evolution*, *11*(4), 398–411. https://doi.org/10.1111/j.1558-5646.1957.tb02911.x

Wohnhaas, C. T., Leparc, G. G., Fernandez-Albert, F., Kind, D., Gantner, F., Viollet, C., Hildebrandt, T., & Baum, P. (2019). DMSO cryopreservation is the method of choice to preserve cells for droplet-based single-cell RNA sequencing. *Scientific Reports*, *9*(1), 10699. https://doi.org/10.1038/s41598-019-46932-z

Wolf, F. A., Angerer, P., & Theis, F. J. (2018). SCANPY: Large-scale single-cell gene expression data analysis. *Genome Biology*, *19*(1), 15. https://doi.org/10.1186/s13059-017-1382-0

Wolff, A. S. B., Oftedal, B. E. V., Kisand, K., Ersvaer, E., Lima, K., & Husebye, E. S. (2010). Flow Cytometry Study of Blood Cell Subtypes Reflects Autoimmune and Inflammatory Processes in Autoimmune Polyendocrine Syndrome Type I: Immune Phenotyping of Patients with APS I and Relatives. *Scandinavian Journal of Immunology*, *71*(6), 459–467. https://doi.org/10.1111/j.1365-3083.2010.02397.x

Wu, A. R., Wang, J., Streets, A. M., & Huang, Y. (2017). Single-Cell Transcriptional Analysis. *Annual Review of Analytical Chemistry*, *10*(1), 439–462. https://doi.org/10.1146/annurev-anchem-061516-045228

Xi, N. M., & Li, J. J. (2021). Benchmarking Computational Doublet-Detection Methods for Single-Cell RNA Sequencing Data. *Cell Systems*, *12*(2), 176-194.e6. https://doi.org/10.1016/j.cels.2020.11.008

Xie, Y., Chen, H., Chellamuthu, V. R., Lajam, A. B. M., Albani, S., Low, A. H. L., Petretto, E., & Behmoaras, J. (2024). Comparative Analysis of Single-Cell RNA Sequencing Methods with and without Sample Multiplexing. *International Journal of Molecular Sciences*, *25*(7), 3828. https://doi.org/10.3390/ijms25073828

Xu, C., Lopez, R., Mehlman, E., Regier, J., Jordan, M. I., & Yosef, N. (2021). Probabilistic harmonization and annotation of single-cell transcriptomics data with deep generative models. *Molecular Systems Biology*, *17*(1), e9620. https://doi.org/10.15252/msb.20209620

Yamano, T., Nedjic, J., Hinterberger, M., Steinert, M., Koser, S., Pinto, S., Gerdes, N., Lutgens, E., Ishimaru, N., Busslinger, M., Brors, B., Kyewski, B., & Klein, L. (2015). Thymic B Cells Are Licensed to Present Self Antigens for Central T Cell Tolerance Induction. *Immunity*, *42*(6), 1048–1061. https://doi.org/10.1016/j.immuni.2015.05.013

Yamawaki, T. M., Lu, D. R., Ellwanger, D. C., Bhatt, D., Manzanillo, P., Arias, V., Zhou, H., Yoon, O. K., Homann, O., Wang, S., & Li, C.-M. (2021). Systematic comparison of high-throughput single-cell RNA-seq methods for immune cell profiling. *BMC Genomics*, *22*(1), 66. https://doi.org/10.1186/s12864-020-07358-4

Yang, J.-H., Hayano, M., Griffin, P. T., Amorim, J. A., Bonkowski, M. S., Apostolides, J. K., Salfati, E. L., Blanchette, M., Munding, E. M., Bhakta, M., Chew, Y. C., Guo, W., Yang, X., Maybury-Lewis, S., Tian, X., Ross, J. M., Coppotelli, G., Meer, M. V., Rogers-Hammond, R., … Sinclair, D. A. (2023). Loss of epigenetic information as a cause of mammalian aging. *Cell*, *186*(2), 305-326.e27. https://doi.org/10.1016/j.cell.2022.12.027

Yang, S., Corbett, S. E., Koga, Y., Wang, Z., Johnson, W. E., Yajima, M., & Campbell, J. D. (2020). Decontamination of ambient RNA in single-cell RNA-seq with DecontX. *Genome Biology*, *21*(1), 57. https://doi.org/10.1186/s13059-020-1950-6

You, Y., Tian, L., Su, S., Dong, X., Jabbari, J. S., Hickey, P. F., & Ritchie, M. E. (2021). Benchmarking UMI-based single-cell RNA-seq preprocessing workflows. *Genome Biology*, *22*(1), 339. https://doi.org/10.1186/s13059-021-02552-3

Young, K., Eudy, E., Bell, R., Loberg, M. A., Stearns, T., Sharma, D., Velten, L., Haas, S., Filippi, M.-D., & Trowbridge, J. J. (2021). Decline in IGF1 in the bone marrow microenvironment initiates hematopoietic stem cell aging. *Cell Stem Cell*, *28*(8), 1473-1482.e7. https://doi.org/10.1016/j.stem.2021.03.017

Young, M. D., & Behjati, S. (2020). SoupX removes ambient RNA contamination from droplet-based single-cell RNA sequencing data. *GigaScience*, *9*(12), giaa151. https://doi.org/10.1093/gigascience/giaa151

Zappia, L., Phipson, B., & Oshlack, A. (2018). Exploring the single-cell RNA-seq analysis landscape with the scRNA-tools database. *PLOS Computational Biology*, *14*(6), e1006245. https://doi.org/10.1371/journal.pcbi.1006245

Zeng, H. (2022). What is a cell type and how to define it? *Cell*, *185*(15), 2739–2755. https://doi.org/10.1016/j.cell.2022.06.031

Zeng, L., Yang, K., Zhang, T., Zhu, X., Hao, W., Chen, H., & Ge, J. (2022). Research progress of single-cell transcriptome sequencing in autoimmune diseases and auto-

inflammatory disease: A review. *Journal of Autoimmunity*, *133*, 102919. https://doi.org/10.1016/j.jaut.2022.102919

Zhang, X., Shi, M., Zhao, X., Bin, E., Hu, Y., Tang, N., Dai, H., & Wang, C. (2022). Telomere shortening impairs alveolar regeneration. *Cell Proliferation*, *55*(4), e13211. https://doi.org/10.1111/cpr.13211

Zhang, Y., Tan, J., Yang, K., Fan, W., Yu, B., & Shi, W. (2023). Ambient RNAs removal of cortex-specific snRNA-seq reveals Apoe+ microglia/macrophage after deeper cerebral hypoperfusion in mice. *Journal of Neuroinflammation*, *20*(1), 152. https://doi.org/10.1186/s12974-023-02831-9

Zhang, Z., Tian, X., Lu, J. Y., Boit, K., Ablaeva, J., Zakusilo, F. T., Emmrich, S., Firsanov, D., Rydkina, E., Biashad, S. A., Lu, Q., Tyshkovskiy, A., Gladyshev, V. N., Horvath, S., Seluanov, A., & Gorbunova, V. (2023). Increased hyaluronan by naked mole-rat Has2 improves healthspan in mice. *Nature*, *621*(7977), 196–205. https://doi.org/10.1038/s41586-023-06463-0

Zheng, G. X. Y., Terry, J. M., Belgrader, P., Ryvkin, P., Bent, Z. W., Wilson, R., Ziraldo, S. B., Wheeler, T. D., McDermott, G. P., Zhu, J., Gregory, M. T., Shuga, J., Montesclaros, L., Underwood, J. G., Masquelier, D. A., Nishimura, S. Y., Schnall-Levin, M., Wyatt, P. W., Hindson, C. M., … Bielas, J. H. (2017). Massively parallel digital transcriptional profiling of single cells. *Nature Communications*, *8*(1), 14049. https://doi.org/10.1038/ncomms14049

Zheng, Y., Liu, X., Le, W., Xie, L., Li, H., Wen, W., Wang, S., Ma, S., Huang, Z., Ye, J., Shi, W., Ye, Y., Liu, Z., Song, M., Zhang, W., Han, J.-D. J., Belmonte, J. C. I., Xiao, C., Qu, J., … Su, W. (2020). A human circulating immune cell landscape in aging and COVID-19. *Protein & Cell*, *11*(10), 740–770. https://doi.org/10.1007/s13238-020-00762-2

Zhu, H., Chen, J., Liu, K., Gao, L., Wu, H., Ma, L., Zhou, J., Liu, Z., & Han, J.-D. J. (2023). Human PBMC scRNA-seq–based aging clocks reveal ribosome to inflammation balance as a single-cell aging hallmark and super longevity. *Science Advances*, *9*(26), eabq7599. https://doi.org/10.1126/sciadv.abq7599

Zinatizadeh, M. R., Zarandi, P. K., Ghiasi, M., Kooshki, H., Mohammadi, M., Amani, J., & Rezaei, N. (2023). Immunosenescence and inflamm-ageing in COVID-19. *Ageing Research Reviews*, *84*, 101818. https://doi.org/10.1016/j.arr.2022.101818

Zuhair, M., Smit, G. S. A., Wallis, G., Jabbar, F., Smith, C., Devleesschauwer, B., & Griffiths, P. (2019). Estimation of the worldwide seroprevalence of cytomegalovirus: A systematic review and meta-analysis. *Reviews in Medical Virology*, *29*(3), e2034. https://doi.org/10.1002/rmv.2034

ACKNOWLEDGMENTS

I am profoundly grateful to my supervisor, Dr. Pärt Peterson, for his steadfast guidance and support, which have been indispensable throughout every phase of my PhD journey. I also extend my gratitude to my industrial managers, Dr. Leif Schauser and Dr. Vivi Gregersen, for granting me the freedom to independently explore and shape the direction of this research.

My heartfelt thanks go to the ENLIGHT-TEN+ training network for cultivating an environment that facilitated research collaborations across labs and countries. To my fellow ENLIGHT-TEN+ ESRs, thank you for the camaraderie, support, and shared laughter that lightened even the most demanding moments. I am especially grateful to Dr. Lisa Hönicke for her dedication in managing the ENLIGHT-TEN+ network.

I am deeply thankful to the research labs of Dr. Jochen Hühn, Dr. Wilfried Ellmeier, Dr. Luís Graça, and Dr. Marc Veldhoen for hosting my research stays.

I also extend my gratitude to the Molecular Pathology Research Group at Tartu University for their support and experimental work. A special thanks to Dr. Kai Kisand for the opportunity to collaborate and to all my collaborators whose contributions have enriched this research.

I am especially grateful to all the patients who generously donated blood for these studies, without whom this work would not have been possible.

To my dear friends, Ivan and Vladimir, thank you for your unwavering support. Ivan, for being my enduring link to the past, and Vladimir, for lifting my spirits with your humor and encouragement over the years.

Lastly, I owe my deepest gratitude to my family. To my mother, Lidiia, for instilling in me the enduring value of education, and to my wife, Olga, for her unconditional love and support, which have been my foundation throughout this journey.

To all of you, thank you for being part of this chapter of my life.

PUBLICATIONS

CURRICULUM VITAE

Name Ihor (Igor) Filippov
Born August 13th 1991
Address Molecular Pathology Research Group, Institute of Biomedicine and Translational Medicine, University of Tartu, Ravila 19, Tartu, 50411, Estonia
Email ihor.filippov@ut.ee, i.am.filippov@gmail.com

Education

2021–…	Tartu University, doctoral studies
2019–2021	Ghent University, MSc Bioinformatics
2008–2012	Sevastopol State University, BSc Systems Engineering

Employment

01.09.2024–31.08.2025	Junior research fellow, Tartu University, Tartu, Estonia
01.09.2021–31.08.2024	Doctoral student, Qiagen Aarhus, Aarhus, Denmark
01.08.2015–31.07.2019	Engineering manager and software developer, HeyJobs GmbH (formerly Memorado GmbH), Berlin, Germany
2010–2015	Mobile software developer, various companies

Conference presentations

03.2022	Genes 2023 – Single-Cell Genomics Moving Forward, oral presentation
02.2024	Institute of Biomedicine and Translational Medicine Science Seminar, oral presentation

Professional training

07.2022	Course on Bioinformatics for T-Cell immunology, European Bioinformatics Institute, UK
11.2023	Course “Translational Medicine Explained”, Bayer AG, Germany
09.2024	Course “Molecular mechanisms of ageing and regeneration” Spetses, Greece

List of publications

1. **Filippov, I.**, Philip, C. S., Schauser, L., & Peterson, P. (2024). Comparative transcriptomic analyses of thymocytes using 10x Genomics and Parse scRNA-seq technologies. *BMC Genomics*, *25*(1), 1069. https://doi.org/10.1186/s12864-024-10976-x
2. **Filippov, I.**, Schauser, L., & Peterson, P. (2024). An integrated single-cell atlas of blood immune cells in aging. *Npj Aging,* 10(1), 59. https://doi.org/10.1038/s41514-024-00185-x

3. Sjøgren, T., Islam, S., **Filippov, I.**, Jebrzycka, A., Sulen, A., Breivik, L. E., Hellesen, A., Jørgensen, A. P., Lima, K., Tserel, L., Kisand, K., Peterson, P., Ranki, A., Husebye, E. S., Oftedal, B. E., & Wolff, A. S. B. (2024). Single cell characterization of blood and expanded regulatory T cells in autoimmune polyendocrine syndrome type 1. *iScience*, *27*(4), 109610. https://doi.org/10.1016/j.isci.2024.109610
4. Türk, L., **Filippov, I.**, Arnold, C., Zaugg, J., Tserel, L., Kisand, K., & Peterson, P. (2024). Cytotoxic CD8+ Temra cells show loss of chromatin accessibility at genes associated with T cell activation. *Frontiers in Immunology*, *15*, 1285798. https://doi.org/10.3389/fimmu.2024.1285798

ELULOOKIRJELDUS

Nimi	Ihor (Igor) Filippov
Sündinud	13. august 1991
Aadress	Molekulaarpatoloogia uurimisgrupp, Bio-ja siirdemeditsiini instituut, Tartu Ülikool, Ravila 19, Tartu, 50411, Eesti
Email	ihor.filippov@ut.ee, i.am.filippov@gmail.com

Hariduskäik

2021–…	Tartu Ülikool, doktorantuur
2019–2021	Genti Ülikool, magistrikraad bioinformaatikas
2008–2012	Sevastopoli Riiklik Ülikool, bakalaureusekraad süsteemitehnikas

Teenistuskäik

01.09.2024–31.08.2025	Molekulaarimmunoloogia nooremteadur, Tartu Ülikool, Tartu, Eesti
01.09.2021–31.08.2024	Doktorant, Qiagen Aarhus, Aarhus, Taani
01.08.2015–31.07.2019	Insenerijuht ja tarkvaraarendaja, HeyJobs GmbH (endine Memorado GmbH), Berliin, Saksamaa
2010–2015	Mobiilitarkvara arendaja, erinevates ettevõtetes

Konverentsiettekanded

03.2022	Genes 2023 – Single-Cell Genomics Moving Forward, suuline ettekanne
02.2024	Bio-ja siirdemeditsiini instituut teadusseminar, suuline ettekanne

Enesetäiendus

07.2022	Kursus Bioinformatics for T-Cell immunology, European Bioinformatics Institute, UK
11.2023	Kursus “Translational Medicine Explained”, Bayer AG, Germany
09.2024	Kursus “Molecular mechanisms of ageing and regeneration” Spetses, Greece

Publikatsioonide nimekiri

1. **Filippov, I.**, Philip, C. S., Schauser, L., & Peterson, P. (2024). Comparative transcriptomic analyses of thymocytes using 10x Genomics and Parse scRNA-seq technologies. *BMC Genomics*, *25*(1), 1069. https://doi.org/10.1186/s12864-024-10976-x
2. **Filippov, I.**, Schauser, L., & Peterson, P. (2024). An integrated single-cell atlas of blood immune cells in aging. *Npj Aging,* 10(1), 59. https://doi.org/10.1038/s41514-024-00185-x
3. Sjøgren, T., Islam, S., **Filippov, I.**, Jebrzycka, A., Sulen, A., Breivik, L. E., Hellesen, A., Jørgensen, A. P., Lima, K., Tserel, L., Kisand, K., Peterson, P., Ranki, A., Husebye, E. S., Oftedal, B. E., & Wolff, A. S. B. (2024). Single cell

characterization of blood and expanded regulatory T cells in autoimmune polyendocrine syndrome type 1. *iScience*, *27*(4), 109610. https://doi.org/10.1016/j.isci.2024.109610

4. Türk, L., **Filippov, I.**, Arnold, C., Zaugg, J., Tserel, L., Kisand, K., & Peterson, P. (2024). Cytotoxic CD8+ Temra cells show loss of chromatin accessibility at genes associated with T cell activation. *Frontiers in Immunology*, *15*, 1285798. https://doi.org/10.3389/fimmu.2024.1285798

DISSERTATIONES MEDICINAE UNIVERSITATIS TARTUENSIS

1. **Heidi-Ingrid Maaroos**. The natural course of gastric ulcer in connection with chronic gastritis and *Helicobacter pylori*. Tartu, 1991.
2. **Mihkel Zilmer**. Na-pump in normal and tumorous brain tissues: Structural, functional and tumorigenesis aspects. Tartu, 1991.
3. **Eero Vasar**. Role of cholecystokinin receptors in the regulation of behaviour and in the action of haloperidol and diazepam. Tartu, 1992.
4. **Tiina Talvik**. Hypoxic-ischaemic brain damage in neonates (clinical, biochemical and brain computed tomographical investigation). Tartu, 1992.
5. **Ants Peetsalu**. Vagotomy in duodenal ulcer disease: A study of gastric acidity, serum pepsinogen I, gastric mucosal histology and *Helicobacter pylori*. Tartu, 1992.
6. **Marika Mikelsaar**. Evaluation of the gastrointestinal microbial ecosystem in health and disease. Tartu, 1992.
7. **Hele Everaus**. Immuno-hormonal interactions in chronic lymphocytic leukaemia and multiple myeloma. Tartu, 1993.
8. **Ruth Mikelsaar**. Etiological factors of diseases in genetically consulted children and newborn screening: dissertation for the commencement of the degree of doctor of medical sciences. Tartu, 1993.
9. **Agu Tamm**. On metabolic action of intestinal microflora: clinical aspects. Tartu, 1993.
10. **Katrin Gross**. Multiple sclerosis in South-Estonia (epidemiological and computed tomographical investigations). Tartu, 1993.
11. **Oivi Uibo**. Childhood coeliac disease in Estonia: occurrence, screening, diagnosis and clinical characterization. Tartu, 1994.
12. **Viiu Tuulik**. The functional disorders of central nervous system of chemistry workers. Tartu, 1994.
13. **Margus Viigimaa**. Primary haemostasis, antiaggregative and anticoagulant treatment of acute myocardial infarction. Tartu, 1994.
14. **Rein Kolk**. Atrial versus ventricular pacing in patients with sick sinus syndrome. Tartu, 1994.
15. **Toomas Podar**. Incidence of childhood onset type 1 diabetes mellitus in Estonia. Tartu, 1994.
16. **Kiira Subi**. The laboratory surveillance of the acute respiratory viral infections in Estonia. Tartu, 1995.
17. **Irja Lutsar**. Infections of the central nervous system in children (epidemiologic, diagnostic and therapeutic aspects, long term outcome). Tartu, 1995.
18. **Aavo Lang**. The role of dopamine, 5-hydroxytryptamine, sigma and NMDA receptors in the action of antipsychotic drugs. Tartu, 1995.
19. **Andrus Arak**. Factors influencing the survival of patients after radical surgery for gastric cancer. Tartu, 1996.

20. **Tõnis Karki**. Quantitative composition of the human lactoflora and method for its examination. Tartu, 1996.
21. **Reet Mändar**. Vaginal microflora during pregnancy and its transmission to newborn. Tartu, 1996.
22. **Triin Remmel**. Primary biliary cirrhosis in Estonia: epidemiology, clinical characterization and prognostication of the course of the disease. Tartu, 1996.
23. **Toomas Kivastik**. Mechanisms of drug addiction: focus on positive reinforcing properties of morphine. Tartu, 1996.
24. **Paavo Pokk**. Stress due to sleep deprivation: focus on $GABA_A$ receptor-chloride ionophore complex. Tartu, 1996.
25. **Kristina Allikmets**. Renin system activity in essential hypertension. Associations with atherothrombogenic cardiovascular risk factors and with the efficacy of calcium antagonist treatment. Tartu, 1996.
26. **Triin Parik**. Oxidative stress in essential hypertension: Associations with metabolic disturbances and the effects of calcium antagonist treatment. Tartu, 1996.
27. **Svetlana Päi**. Factors promoting heterogeneity of the course of rheumatoid arthritis. Tartu, 1997.
28. **Maarike Sallo**. Studies on habitual physical activity and aerobic fitness in 4 to 10 years old children. Tartu, 1997.
29. **Paul Naaber**. *Clostridium difficile* infection and intestinal microbial ecology. Tartu, 1997.
30. **Rein Pähkla**. Studies in pinoline pharmacology. Tartu, 1997.
31. **Andrus Juhan Voitk**. Outpatient laparoscopic cholecystectomy. Tartu, 1997.
32. **Joel Starkopf**. Oxidative stress and ischaemia-reperfusion of the heart. Tartu, 1997.
33. **Janika Kõrv**. Incidence, case-fatality and outcome of stroke. Tartu, 1998.
34. **Ülla Linnamägi**. Changes in local cerebral blood flow and lipid peroxidation following lead exposure in experiment. Tartu, 1998.
35. **Ave Minajeva**. Sarcoplasmic reticulum function: comparison of atrial and ventricular myocardium. Tartu, 1998.
36. **Oleg Milenin**. Reconstruction of cervical part of esophagus by revascularised ileal autografts in dogs. A new complex multistage method. Tartu, 1998.
37. **Sergei Pakriev**. Prevalence of depression, harmful use of alcohol and alcohol dependence among rural population in Udmurtia. Tartu, 1998.
38. **Allen Kaasik**. Thyroid hormone control over β-adrenergic signalling system in rat atria. Tartu, 1998.
39. **Vallo Matto**. Pharmacological studies on anxiogenic and antiaggressive properties of antidepressants. Tartu, 1998.
40. **Maire Vasar**. Allergic diseases and bronchial hyperreactivity in Estonian children in relation to environmental influences. Tartu, 1998.
41. **Kaja Julge**. Humoral immune responses to allergens in early childhood. Tartu, 1998.

42. **Heli Grünberg**. The cardiovascular risk of Estonian schoolchildren. A cross-sectional study of 9-, 12- and 15-year-old children. Tartu, 1998.
43. **Epp Sepp**. Formation of intestinal microbial ecosystem in children. Tartu, 1998.
44. **Mai Ots**. Characteristics of the progression of human and experimental glomerulopathies. Tartu, 1998.
45. **Tiina Ristimäe**. Heart rate variability in patients with coronary artery disease. Tartu, 1998.
46. **Leho Kõiv**. Reaction of the sympatho-adrenal and hypothalamo-pituitary-adrenocortical system in the acute stage of head injury. Tartu, 1998.
47. **Bela Adojaan**. Immune and genetic factors of childhood onset IDDM in Estonia. An epidemiological study. Tartu, 1999.
48. **Jakov Shlik**. Psychophysiological effects of cholecystokinin in humans. Tartu, 1999.
49. **Kai Kisand**. Autoantibodies against dehydrogenases of α-ketoacids. Tartu, 1999.
50. **Toomas Marandi**. Drug treatment of depression in Estonia. Tartu, 1999.
51. **Ants Kask**. Behavioural studies on neuropeptide Y. Tartu, 1999.
52. **Ello-Rahel Karelson**. Modulation of adenylate cyclase activity in the rat hippocampus by neuropeptide galanin and its chimeric analogs. Tartu, 1999.
53. **Tanel Laisaar**. Treatment of pleural empyema — special reference to intrapleural therapy with streptokinase and surgical treatment modalities. Tartu, 1999.
54. **Eve Pihl**. Cardiovascular risk factors in middle-aged former athletes. Tartu, 1999.
55. **Katrin Õunap**. Phenylketonuria in Estonia: incidence, newborn screening, diagnosis, clinical characterization and genotype/phenotype correlation. Tartu, 1999.
56. **Siiri Kõljalg**. *Acinetobacter* – an important nosocomial pathogen. Tartu, 1999.
57. **Helle Karro**. Reproductive health and pregnancy outcome in Estonia: association with different factors. Tartu, 1999.
58. **Heili Varendi**. Behavioral effects observed in human newborns during exposure to naturally occurring odors. Tartu, 1999.
59. **Anneli Beilmann**. Epidemiology of epilepsy in children and adolescents in Estonia. Prevalence, incidence, and clinical characteristics. Tartu, 1999.
60. **Vallo Volke**. Pharmacological and biochemical studies on nitric oxide in the regulation of behaviour. Tartu, 1999.
61. **Pilvi Ilves**. Hypoxic-ischaemic encephalopathy in asphyxiated term infants. A prospective clinical, biochemical, ultrasonographical study. Tartu, 1999.
62. **Anti Kalda**. Oxygen-glucose deprivation-induced neuronal death and its pharmacological prevention in cerebellar granule cells. Tartu, 1999.
63. **Eve-Irene Lepist**. Oral peptide prodrugs – studies on stability and absorption. Tartu, 2000.

64. **Jana Kivastik**. Lung function in Estonian schoolchildren: relationship with anthropometric indices and respiratory symptomas, reference values for dynamic spirometry. Tartu, 2000.
65. **Karin Kull**. Inflammatory bowel disease: an immunogenetic study. Tartu, 2000.
66. **Kaire Innos**. Epidemiological resources in Estonia: data sources, their quality and feasibility of cohort studies. Tartu, 2000.
67. **Tamara Vorobjova**. Immune response to *Helicobacter pylori* and its association with dynamics of chronic gastritis and epithelial cell turnover in antrum and corpus. Tartu, 2001.
68. **Ruth Kalda**. Structure and outcome of family practice quality in the changing health care system of Estonia. Tartu, 2001.
69. **Annika Krüüner**. *Mycobacterium tuberculosis* – spread and drug resistance in Estonia. Tartu, 2001.
70. **Marlit Veldi**. Obstructive Sleep Apnoea: Computerized Endopharyngeal Myotonometry of the Soft Palate and Lingual Musculature. Tartu, 2001.
71. **Anneli Uusküla**. Epidemiology of sexually transmitted diseases in Estonia in 1990–2000. Tartu, 2001.
72. **Ade Kallas**. Characterization of antibodies to coagulation factor VIII. Tartu, 2002.
73. **Heidi Annuk**. Selection of medicinal plants and intestinal lactobacilli as antimicrobil components for functional foods. Tartu, 2002.
74. **Aet Lukmann**. Early rehabilitation of patients with ischaemic heart disease after surgical revascularization of the myocardium: assessment of health-related quality of life, cardiopulmonary reserve and oxidative stress. A clinical study. Tartu, 2002.
75. **Maigi Eisen**. Pathogenesis of Contact Dermatitis: participation of Oxidative Stress. A clinical – biochemical study. Tartu, 2002.
76. **Piret Hussar**. Histology of the post-traumatic bone repair in rats. Elaboration and use of a new standardized experimental model – bicortical perforation of tibia compared to internal fracture and resection osteotomy. Tartu, 2002.
77. **Tõnu Rätsep**. Aneurysmal subarachnoid haemorrhage: Noninvasive monitoring of cerebral haemodynamics. Tartu, 2002.
78. **Marju Herodes**. Quality of life of people with epilepsy in Estonia. Tartu, 2003.
79. **Katre Maasalu**. Changes in bone quality due to age and genetic disorders and their clinical expressions in Estonia. Tartu, 2003.
80. **Toomas Sillakivi**. Perforated peptic ulcer in Estonia: epidemiology, risk factors and relations with *Helicobacter pylori*. Tartu, 2003.
81. **Leena Puksa**. Late responses in motor nerve conduction studies. F and A waves in normal subjects and patients with neuropathies. Tartu, 2003.
82. **Krista Lõivukene**. *Helicobacter pylori* in gastric microbial ecology and its antimicrobial susceptibility pattern. Tartu, 2003.

83. **Helgi Kolk**. Dyspepsia and *Helicobacter pylori* infection: the diagnostic value of symptoms, treatment and follow-up of patients referred for upper gastrointestinal endoscopy by family physicians. Tartu, 2003.
84. **Helena Soomer**. Validation of identification and age estimation methods in forensic odontology. Tartu, 2003.
85. **Kersti Oselin**. Studies on the human MDR1, MRP1, and MRP2 ABC transporters: functional relevance of the genetic polymorphisms in the *MDR1* and *MRP1* gene. Tartu, 2003.
86. **Jaan Soplepmann**. Peptic ulcer haemorrhage in Estonia: epidemiology, prognostic factors, treatment and outcome. Tartu, 2003.
87. **Margot Peetsalu**. Long-term follow-up after vagotomy in duodenal ulcer disease: recurrent ulcer, changes in the function, morphology and *Helicobacter pylori* colonisation of the gastric mucosa. Tartu, 2003.
88. **Kersti Klaamas**. Humoral immune response to *Helicobacter pylori* a study of host-dependent and microbial factors. Tartu, 2003.
89. **Pille Taba**. Epidemiology of Parkinson's disease in Tartu, Estonia. Prevalence, incidence, clinical characteristics, and pharmacoepidemiology. Tartu, 2003.
90. **Alar Veraksitš**. Characterization of behavioural and biochemical phenotype of cholecystokinin-2 receptor deficient mice: changes in the function of the dopamine and endopioidergic system. Tartu, 2003.
91. **Ingrid Kalev**. CC-chemokine receptor 5 (CCR5) gene polymorphism in Estonians and in patients with Type I and Type II diabetes mellitus. Tartu, 2003.
92. **Lumme Kadaja**. Molecular approach to the regulation of mitochondrial function in oxidative muscle cells. Tartu, 2003.
93. **Aive Liigant**. Epidemiology of primary central nervous system tumours in Estonia from 1986 to 1996. Clinical characteristics, incidence, survival and prognostic factors. Tartu, 2004.
94. **Andres, Kulla**. Molecular characteristics of mesenchymal stroma in human astrocytic gliomas. Tartu, 2004.
95. **Mari Järvelaid**. Health damaging risk behaviours in adolescence. Tartu, 2004.
96. **Ülle Pechter**. Progression prevention strategies in chronic renal failure and hypertension. An experimental and clinical study. Tartu, 2004.
97. **Gunnar Tasa**. Polymorphic glutathione S-transferases – biology and role in modifying genetic susceptibility to senile cataract and primary open angle glaucoma. Tartu, 2004.
98. **Tuuli Käämbre**. Intracellular energetic unit: structural and functional aspects. Tartu, 2004.
99. **Vitali Vassiljev**. Influence of nitric oxide syntase inhibitors on the effects of ethanol after acute and chronic ethanol administration and withdrawal. Tartu, 2004.

100. **Aune Rehema**. Assessment of nonhaem ferrous iron and glutathione redox ratio as markers of pathogeneticity of oxidative stress in different clinical groups. Tartu, 2004.
101. **Evelin Seppet**. Interaction of mitochondria and ATPases in oxidative muscle cells in normal and pathological conditions. Tartu, 2004.
102. **Eduard Maron**. Serotonin function in panic disorder: from clinical experiments to brain imaging and genetics. Tartu, 2004.
103. **Marje Oona**. *Helicobacter pylori* infection in children: epidemiological and therapeutic aspects. Tartu, 2004.
104. **Kersti Kokk**. Regulation of active and passive molecular transport in the testis. Tartu, 2005.
105. **Vladimir Järv**. Cross-sectional imaging for pretreatment evaluation and follow-up of pelvic malignant tumours. Tartu, 2005.
106. **Andre Õun**. Epidemiology of adult epilepsy in Tartu, Estonia. Incidence, prevalence and medical treatment. Tartu, 2005.
107. **Piibe Muda**. Homocysteine and hypertension: associations between homocysteine and essential hypertension in treated and untreated hypertensive patients with and without coronary artery disease. Tartu, 2005.
108. **Külli Kingo**. The interleukin-10 family cytokines gene polymorphisms in plaque psoriasis. Tartu, 2005.
109. **Mati Merila**. Anatomy and clinical relevance of the glenohumeral joint capsule and ligaments. Tartu, 2005.
110. **Epp Songisepp**. Evaluation of technological and functional properties of the new probiotic *Lactobacillus fermentum* ME-3. Tartu, 2005.
111. **Tiia Ainla**. Acute myocardial infarction in Estonia: clinical characteristics, management and outcome. Tartu, 2005.
112. **Andres Sell**. Determining the minimum local anaesthetic requirements for hip replacement surgery under spinal anaesthesia – a study employing a spinal catheter. Tartu, 2005.
113. **Tiia Tamme**. Epidemiology of odontogenic tumours in Estonia. Pathogenesis and clinical behaviour of ameloblastoma. Tartu, 2005.
114. **Triine Annus**. Allergy in Estonian schoolchildren: time trends and characteristics. Tartu, 2005.
115. **Tiia Voor**. Microorganisms in infancy and development of allergy: comparison of Estonian and Swedish children. Tartu, 2005.
116. **Priit Kasenõmm**. Indicators for tonsillectomy in adults with recurrent tonsillitis – clinical, microbiological and pathomorphological investigations. Tartu, 2005.
117. **Eva Zusinaite**. Hepatitis C virus: genotype identification and interactions between viral proteases. Tartu, 2005.
118. **Piret Kõll**. Oral lactoflora in chronic periodontitis and periodontal health. Tartu, 2006.
119. **Tiina Stelmach**. Epidemiology of cerebral palsy and unfavourable neurodevelopmental outcome in child population of Tartu city and county, Estonia Prevalence, clinical features and risk factors. Tartu, 2006.

120. **Katrin Pudersell**. Tropane alkaloid production and riboflavine excretion in the field and tissue cultures of henbane (*Hyoscyamus niger* L.). Tartu, 2006.
121. **Külli Jaako**. Studies on the role of neurogenesis in brain plasticity. Tartu, 2006.
122. **Aare Märtson**. Lower limb lengthening: experimental studies of bone regeneration and long-term clinical results. Tartu, 2006.
123. **Heli Tähepõld**. Patient consultation in family medicine. Tartu, 2006.
124. **Stanislav Liskmann**. Peri-implant disease: pathogenesis, diagnosis and treatment in view of both inflammation and oxidative stress profiling. Tartu, 2006.
125. **Ruth Rudissaar**. Neuropharmacology of atypical antipsychotics and an animal model of psychosis. Tartu, 2006.
126. **Helena Andreson**. Diversity of *Helicobacter pylori* genotypes in Estonian patients with chronic inflammatory gastric diseases. Tartu, 2006.
127. **Katrin Pruus**. Mechanism of action of antidepressants: aspects of serotoninergic system and its interaction with glutamate. Tartu, 2006.
128. **Priit Põder**. Clinical and experimental investigation: relationship of ischaemia/reperfusion injury with oxidative stress in abdominal aortic aneurysm repair and in extracranial brain artery endarterectomy and possibilities of protection against ischaemia using a glutathione analogue in a rat model of global brain ischaemia. Tartu, 2006.
129. **Marika Tammaru**. Patient-reported outcome measurement in rheumatoid arthritis. Tartu, 2006.
130. **Tiia Reimand**. Down syndrome in Estonia. Tartu, 2006.
131. **Diva Eensoo**. Risk-taking in traffic and Markers of Risk-Taking Behaviour in Schoolchildren and Car Drivers. Tartu, 2007.
132. **Riina Vibo**. The third stroke registry in Tartu, Estonia from 2001 to 2003: incidence, case-fatality, risk factors and long-term outcome. Tartu, 2007.
133. **Chris Pruunsild**. Juvenile idiopathic arthritis in children in Estonia. Tartu, 2007.
134. **Eve Õiglane-Šlik**. Angelman and Prader-Willi syndromes in Estonia. Tartu, 2007.
135. **Kadri Haller**. Antibodies to follicle stimulating hormone. Significance in female infertility. Tartu, 2007.
136. **Pille Ööpik**. Management of depression in family medicine. Tartu, 2007.
137. **Jaak Kals**. Endothelial function and arterial stiffness in patients with atherosclerosis and in healthy subjects. Tartu, 2007.
138. **Priit Kampus**. Impact of inflammation, oxidative stress and age on arterial stiffness and carotid artery intima-media thickness. Tartu, 2007.
139. **Margus Punab**. Male fertility and its risk factors in Estonia. Tartu, 2007.
140. **Alar Toom**. Heterotopic ossification after total hip arthroplasty: clinical and pathogenetic investigation. Tartu, 2007.

141. **Lea Pehme**. Epidemiology of tuberculosis in Estonia 1991–2003 with special regard to extrapulmonary tuberculosis and delay in diagnosis of pulmonary tuberculosis. Tartu, 2007.
142. **Juri Karjagin**. The pharmacokinetics of metronidazole and meropenem in septic shock. Tartu, 2007.
143. **Inga Talvik**. Inflicted traumatic brain injury shaken baby syndrome in Estonia – epidemiology and outcome. Tartu, 2007.
144. **Tarvo Rajasalu**. Autoimmune diabetes: an immunological study of type 1 diabetes in humans and in a model of experimental diabetes (in RIP-B7.1 mice). Tartu, 2007.
145. **Inga Karu**. Ischaemia-reperfusion injury of the heart during coronary surgery: a clinical study investigating the effect of hyperoxia. Tartu, 2007.
146. **Peeter Padrik**. Renal cell carcinoma: Changes in natural history and treatment of metastatic disease. Tartu, 2007.
147. **Neve Vendt**. Iron deficiency and iron deficiency anaemia in infants aged 9 to 12 months in Estonia. Tartu, 2008.
148. **Lenne-Triin Heidmets**. The effects of neurotoxins on brain plasticity: focus on neural Cell Adhesion Molecule. Tartu, 2008.
149. **Paul Korrovits**. Asymptomatic inflammatory prostatitis: prevalence, etiological factors, diagnostic tools. Tartu, 2008.
150. **Annika Reintam**. Gastrointestinal failure in intensive care patients. Tartu, 2008.
151. **Kristiina Roots**. Cationic regulation of Na-pump in the normal, Alzheimer's and CCK_2 receptor-deficient brain. Tartu, 2008.
152. **Helen Puusepp**. The genetic causes of mental retardation in Estonia: fragile X syndrome and creatine transporter defect. Tartu, 2009.
153. **Kristiina Rull**. Human chorionic gonadotropin beta genes and recurrent miscarriage: expression and variation study. Tartu, 2009.
154. **Margus Eimre**. Organization of energy transfer and feedback regulation in oxidative muscle cells. Tartu, 2009.
155. **Maire Link**. Transcription factors FoxP3 and AIRE: autoantibody associations. Tartu, 2009.
156. **Kai Haldre**. Sexual health and behaviour of young women in Estonia. Tartu, 2009.
157. **Kaur Liivak**. Classical form of congenital adrenal hyperplasia due to 21-hydroxylase deficiency in Estonia: incidence, genotype and phenotype with special attention to short-term growth and 24-hour blood pressure. Tartu, 2009.
158. **Kersti Ehrlich**. Antioxidative glutathione analogues (UPF peptides) – molecular design, structure-activity relationships and testing the protective properties. Tartu, 2009.
159. **Anneli Rätsep**. Type 2 diabetes care in family medicine. Tartu, 2009.
160. **Silver Türk**. Etiopathogenetic aspects of chronic prostatitis: role of mycoplasmas, coryneform bacteria and oxidative stress. Tartu, 2009.

161. **Kaire Heilman**. Risk markers for cardiovascular disease and low bone mineral density in children with type 1 diabetes. Tartu, 2009.
162. **Kristi Rüütel**. HIV-epidemic in Estonia: injecting drug use and quality of life of people living with HIV. Tartu, 2009.
163. **Triin Eller**. Immune markers in major depression and in antidepressive treatment. Tartu, 2009.
164. **Siim Suutre**. The role of TGF-β isoforms and osteoprogenitor cells in the pathogenesis of heterotopic ossification. An experimental and clinical study of hip arthroplasty. Tartu, 2010.
165. **Kai Kliiman**. Highly drug-resistant tuberculosis in Estonia: Risk factors and predictors of poor treatment outcome. Tartu, 2010.
166. **Inga Villa**. Cardiovascular health-related nutrition, physical activity and fitness in Estonia. Tartu, 2010.
167. **Tõnis Org**. Molecular function of the first PHD finger domain of Autoimmune Regulator protein. Tartu, 2010.
168. **Tuuli Metsvaht**. Optimal antibacterial therapy of neonates at risk of early onset sepsis. Tartu, 2010.
169. **Jaanus Kahu**. Kidney transplantation: Studies on donor risk factors and mycophenolate mofetil. Tartu, 2010.
170. **Koit Reimand**. Autoimmunity in reproductive failure: A study on associated autoantibodies and autoantigens. Tartu, 2010.
171. **Mart Kull**. Impact of vitamin D and hypolactasia on bone mineral density: a population based study in Estonia. Tartu, 2010.
172. **Rael Laugesaar**. Stroke in children – epidemiology and risk factors. Tartu, 2010.
173. **Mark Braschinsky**. Epidemiology and quality of life issues of hereditary spastic paraplegia in Estonia and implemention of genetic analysis in everyday neurologic practice. Tartu, 2010.
174. **Kadri Suija**. Major depression in family medicine: associated factors, recurrence and possible intervention. Tartu, 2010.
175. **Jarno Habicht**. Health care utilisation in Estonia: socioeconomic determinants and financial burden of out-of-pocket payments. Tartu, 2010.
176. **Kristi Abram**. The prevalence and risk factors of rosacea. Subjective disease perception of rosacea patients. Tartu, 2010.
177. **Malle Kuum**. Mitochondrial and endoplasmic reticulum cation fluxes: Novel roles in cellular physiology. Tartu, 2010.
178. **Rita Teek**. The genetic causes of early onset hearing loss in Estonian children. Tartu, 2010.
179. **Daisy Volmer**. The development of community pharmacy services in Estonia – public and professional perceptions 1993–2006. Tartu, 2010.
180. **Jelena Lissitsina**. Cytogenetic causes in male infertility. Tartu, 2011.
181. **Delia Lepik**. Comparison of gunshot injuries caused from Tokarev, Makarov and Glock 19 pistols at different firing distances. Tartu, 2011.
182. **Ene-Renate Pähkla**. Factors related to the efficiency of treatment of advanced periodontitis. Tartu, 2011.

183. **Maarja Krass**. L-Arginine pathways and antidepressant action. Tartu, 2011.
184. **Taavi Lai**. Population health measures to support evidence-based health policy in Estonia. Tartu, 2011.
185. **Tiit Salum**. Similarity and difference of temperature-dependence of the brain sodium pump in normal, different neuropathological, and aberrant conditions and its possible reasons. Tartu, 2011.
186. **Tõnu Vooder**. Molecular differences and similarities between histological subtypes of non-small cell lung cancer. Tartu, 2011.
187. **Jelena Štšepetova**. The characterisation of intestinal lactic acid bacteria using bacteriological, biochemical and molecular approaches. Tartu, 2011.
188. **Radko Avi**. Natural polymorphisms and transmitted drug resistance in Estonian HIV-1 CRF06_cpx and its recombinant viruses. Tartu, 2011, 116 p.
189. **Edward Laane**. Multiparameter flow cytometry in haematological malignancies. Tartu, 2011, 152 p.
190. **Triin Jagomägi**. A study of the genetic etiology of nonsyndromic cleft lip and palate. Tartu, 2011, 158 p.
191. **Ivo Laidmäe**. Fibrin glue of fish (*Salmo salar*) origin: immunological study and development of new pharmaceutical preparation. Tartu, 2012, 150 p.
192. **Ülle Parm**. Early mucosal colonisation and its role in prediction of invasive infection in neonates at risk of early onset sepsis. Tartu, 2012, 168 p.
193. **Kaupo Teesalu**. Autoantibodies against desmin and transglutaminase 2 in celiac disease: diagnostic and functional significance. Tartu, 2012, 142 p.
194. **Maksim Zagura**. Biochemical, functional and structural profiling of arterial damage in atherosclerosis. Tartu, 2012, 162 p.
195. **Vivian Kont**. Autoimmune regulator: characterization of thymic gene regulation and promoter methylation. Tartu, 2012, 134 p.
196. **Pirje Hütt**. Functional properties, persistence, safety and efficacy of potential probiotic lactobacilli. Tartu, 2012, 246 p.
197. **Innar Tõru**. Serotonergic modulation of CCK-4- induced panic. Tartu, 2012, 132 p.
198. **Sigrid Vorobjov**. Drug use, related risk behaviour and harm reduction interventions utilization among injecting drug users in Estonia: implications for drug policy. Tartu, 2012, 120 p.
199. **Martin Serg**. Therapeutic aspects of central haemodynamics, arterial stiffness and oxidative stress in hypertension. Tartu, 2012, 156 p.
200. **Jaanika Kumm**. Molecular markers of articular tissues in early knee osteoarthritis: a population-based longitudinal study in middle-aged subjects. Tartu, 2012, 159 p.
201. **Kertu Rünkorg**. Functional changes of dopamine, endopioid and endocannabinoid systems in CCK2 receptor deficient mice. Tartu, 2012, 125 p.
202. **Mai Blöndal**. Changes in the baseline characteristics, management and outcomes of acute myocardial infarction in Estonia. Tartu, 2012, 127 p.

203. **Jana Lass**. Epidemiological and clinical aspects of medicines use in children in Estonia. Tartu, 2012, 170 p.
204. **Kai Truusalu**. Probiotic lactobacilli in experimental persistent *Salmonella* infection. Tartu, 2013, 139 p.
205. **Oksana Jagur**. Temporomandibular joint diagnostic imaging in relation to pain and bone characteristics. Long-term results of arthroscopic treatment. Tartu, 2013, 126 p.
206. **Katrin Sikk**. Manganese-ephedrone intoxication – pathogenesis of neurological damage and clinical symptomatology. Tartu, 2013, 125 p.
207. **Kai Blöndal**. Tuberculosis in Estonia with special emphasis on drug-resistant tuberculosis: Notification rate, disease recurrence and mortality. Tartu, 2013, 151 p.
208. **Marju Puurand**. Oxidative phosphorylation in different diseases of gastric mucosa. Tartu, 2013, 123 p.
209. **Aili Tagoma**. Immune activation in female infertility: Significance of autoantibodies and inflammatory mediators. Tartu, 2013, 135 p.
210. **Liis Sabre**. Epidemiology of traumatic spinal cord injury in Estonia. Brain activation in the acute phase of traumatic spinal cord injury. Tartu, 2013, 135 p.
211. **Merit Lamp**. Genetic susceptibility factors in endometriosis. Tartu, 2013, 125 p.
212. **Erik Salum**. Beneficial effects of vitamin D and angiotensin II receptor blocker on arterial damage. Tartu, 2013, 167 p.
213. **Maire Karelson**. Vitiligo: clinical aspects, quality of life and the role of melanocortin system in pathogenesis. Tartu, 2013, 153 p.
214. **Kuldar Kaljurand**. Prevalence of exfoliation syndrome in Estonia and its clinical significance. Tartu, 2013, 113 p.
215. **Raido Paasma**. Clinical study of methanol poisoning: handling large outbreaks, treatment with antidotes, and long-term outcomes. Tartu, 2013, 96 p.
216. **Anne Kleinberg**. Major depression in Estonia: prevalence, associated factors, and use of health services. Tartu, 2013, 129 p.
217. **Triin Eglit**. Obesity, impaired glucose regulation, metabolic syndrome and their associations with high-molecular-weight adiponectin levels. Tartu, 2014, 115 p.
218. **Kristo Ausmees**. Reproductive function in middle-aged males: Associations with prostate, lifestyle and couple infertility status. Tartu, 2014, 125 p.
219. **Kristi Huik**. The influence of host genetic factors on the susceptibility to HIV and HCV infections among intravenous drug users. Tartu, 2014, 144 p.
220. **Liina Tserel**. Epigenetic profiles of monocytes, monocyte-derived macrophages and dendritic cells. Tartu, 2014, 143 p.
221. **Irina Kerna**. The contribution of *ADAM12* and *CILP* genes to the development of knee osteoarthritis. Tartu, 2014, 152 p.

222. **Ingrid Liiv**. Autoimmune regulator protein interaction with DNA-dependent protein kinase and its role in apoptosis. Tartu, 2014, 143 p.
223. **Liivi Maddison**. Tissue perfusion and metabolism during intra-abdominal hypertension. Tartu, 2014, 103 p.
224. **Krista Ress**. Childhood coeliac disease in Estonia, prevalence in atopic dermatitis and immunological characterisation of coexistence. Tartu, 2014, 124 p.
225. **Kai Muru**. Prenatal screening strategies, long-term outcome of children with marked changes in maternal screening tests and the most common syndromic heart anomalies in Estonia. Tartu, 2014, 189 p.
226. **Kaja Rahu**. Morbidity and mortality among Baltic Chernobyl cleanup workers: a register-based cohort study. Tartu, 2014, 155 p.
227. **Klari Noormets**. The development of diabetes mellitus, fertility and energy metabolism disturbances in a Wfs1-deficient mouse model of Wolfram syndrome. Tartu, 2014, 132 p.
228. **Liis Toome**. Very low gestational age infants in Estonia. Tartu, 2014, 183 p.
229. **Ceith Nikkolo**. Impact of different mesh parameters on chronic pain and foreign body feeling after open inguinal hernia repair. Tartu, 2014, 132 p.
230. **Vadim Brjalin**. Chronic hepatitis C: predictors of treatment response in Estonian patients. Tartu, 2014, 122 p.
231. **Vahur Metsna**. Anterior knee pain in patients following total knee arthroplasty: the prevalence, correlation with patellar cartilage impairment and aspects of patellofemoral congruence. Tartu, 2014, 130 p.
232. **Marju Kase**. Glioblastoma multiforme: possibilities to improve treatment efficacy. Tartu, 2015, 137 p.
233. **Riina Runnel**. Oral health among elementary school children and the effects of polyol candies on the prevention of dental caries. Tartu, 2015, 112 p.
234. **Made Laanpere**. Factors influencing women's sexual health and reproductive choices in Estonia. Tartu, 2015, 176 p.
235. **Andres Lust**. Water mediated solid state transformations of a polymorphic drug – effect on pharmaceutical product performance. Tartu, 2015, 134 p.
236. **Anna Klugman**. Functionality related characterization of pretreated wood lignin, cellulose and polyvinylpyrrolidone for pharmaceutical applications. Tartu, 2015, 156 p.
237. **Triin Laisk-Podar**. Genetic variation as a modulator of susceptibility to female infertility and a source for potential biomarkers. Tartu, 2015, 155 p.
238. **Mailis Tõnisson**. Clinical picture and biochemical changes in blood in children with acute alcohol intoxication. Tartu, 2015, 100 p.
239. **Kadri Tamme**. High volume haemodiafiltration in treatment of severe sepsis – impact on pharmacokinetics of antibiotics and inflammatory response. Tartu, 2015, 133 p.

240. **Kai Part**. Sexual health of young people in Estonia in a social context: the role of school-based sexuality education and youth-friendly counseling services. Tartu, 2015, 203 p.
241. **Urve Paaver**. New perspectives for the amorphization and physical stabilization of poorly water-soluble drugs and understanding their dissolution behavior. Tartu, 2015, 139 p.
242. **Aleksandr Peet**. Intrauterine and postnatal growth in children with HLA-conferred susceptibility to type 1 diabetes. Tartu. 2015, 146 p.
243. **Piret Mitt**. Healthcare-associated infections in Estonia – epidemiology and surveillance of bloodstream and surgical site infections. Tartu, 2015, 145 p.
244. **Merli Saare**. Molecular Profiling of Endometriotic Lesions and Endometria of Endometriosis Patients. Tartu, 2016, 129 p.
245. **Kaja-Triin Laisaar**. People living with HIV in Estonia: Engagement in medical care and methods of increasing adherence to antiretroviral therapy and safe sexual behavior. Tartu, 2016, 132 p.
246. **Eero Merilind**. Primary health care performance: impact of payment and practice-based characteristics. Tartu, 2016, 120 p.
247. **Jaanika Kärner**. Cytokine-specific autoantibodies in AIRE deficiency. Tartu, 2016, 182 p.
248. **Kaido Paapstel**. Metabolomic profile of arterial stiffness and early biomarkers of renal damage in atherosclerosis. Tartu, 2016, 173 p.
249. **Liidia Kiisk**. Long-term nutritional study: anthropometrical and clinico-laboratory assessments in renal replacement therapy patients after intensive nutritional counselling. Tartu, 2016, 207 p.
250. **Georgi Nellis**. The use of excipients in medicines administered to neonates in Europe. Tartu, 2017, 159 p.
251. **Aleksei Rakitin**. Metabolic effects of acute and chronic treatment with valproic acid in people with epilepsy. Tartu, 2017, 125 p.
252. **Eveli Kallas**. The influence of immunological markers to susceptibility to HIV, HBV, and HCV infections among persons who inject drugs. Tartu, 2017, 138 p.
253. **Tiina Freimann**. Musculoskeletal pain among nurses: prevalence, risk factors, and intervention. Tartu, 2017, 125 p.
254. **Evelyn Aaviksoo**. Sickness absence in Estonia: determinants and influence of the sick-pay cut reform. Tartu, 2017, 121 p.
255. **Kalev Nõupuu**. Autosomal-recessive Stargardt disease: phenotypic heterogeneity and genotype-phenotype associations. Tartu, 2017, 131 p.
256. **Ho Duy Binh**. Osteogenesis imperfecta in Vietnam. Tartu, 2017, 125 p.
257. **Uku Haljasorg**. Transcriptional mechanisms in thymic central tolerance. Tartu, 2017, 147 p.
258. **Živile Riispere**. IgA Nephropathy study according to the Oxford Classification: IgA Nephropathy clinical-morphological correlations, disease progression and the effect of renoprotective therapy. Tartu, 2017, 129 p.

259. **Hiie Soeorg**. Coagulase-negative staphylococci in gut of preterm neonates and in breast milk of their mothers. Tartu, 2017, 216 p.
260. **Anne-Mari Anton Willmore**. Silver nanoparticles for cancer research. Tartu, 2017, 132 p.
261. **Ott Laius**. Utilization of osteoporosis medicines, medication adherence and the trend in osteoporosis related hip fractures in Estonia. Tartu, 2017, 134 p.
262. **Alar Aab**. Insights into molecular mechanisms of asthma and atopic dermatitis. Tartu, 2017, 164 p.
263. **Sander Pajusalu**. Genome-wide diagnostics of Mendelian disorders: from chromosomal microarrays to next-generation sequencing. Tartu, 2017, 146 p.
264. **Mikk Jürisson**. Health and economic impact of hip fracture in Estonia. Tartu, 2017, 164 p.
265. **Kaspar Tootsi**. Cardiovascular and metabolomic profiling of osteoarthritis. Tartu, 2017, 150 p.
266. **Mario Saare**. The influence of AIRE on gene expression – studies of transcriptional regulatory mechanisms in cell culture systems. Tartu, 2017, 172 p.
267. **Piia Jõgi**. Epidemiological and clinical characteristics of pertussis in Estonia. Tartu, 2018, 168 p.
268. **Elle Põldoja**. Structure and blood supply of the superior part of the shoulder joint capsule. Tartu, 2018, 116 p.
269. **Minh Son Nguyen**. Oral health status and prevalence of temporomandibular disorders in 65–74-year-olds in Vietnam. Tartu, 2018, 182 p.
270. **Kristian Semjonov**. Development of pharmaceutical quench-cooled molten and melt-electrospun solid dispersions for poorly water-soluble indomethacin. Tartu, 2018, 125 p.
271. **Janne Tiigimäe-Saar**. Botulinum neurotoxin type A treatment for sialorrhea in central nervous system diseases. Tartu, 2018, 109 p.
272. **Veiko Vengerfeldt**. Apical periodontitis: prevalence and etiopathogenetic aspects. Tartu, 2018, 150 p.
273. **Rudolf Bichele**. TNF superfamily and AIRE at the crossroads of thymic differentiation and host protection against *Candida albicans* infection. Tartu, 2018, 153 p.
274. **Olga Tšuiko**. Unravelling Chromosomal Instability in Mammalian Preimplantation Embryos Using Single-Cell Genomics. Tartu, 2018, 169 p.
275. **Kärt Kriisa**. Profile of acylcarnitines, inflammation and oxidative stress in first-episode psychosis before and after antipsychotic treatment. Tartu, 2018, 145 p.
276. **Xuan Dung Ho**. Characterization of the genomic profile of osteosarcoma. Tartu, 2018, 144 p.
277. **Karit Reinson**. New Diagnostic Methods for Early Detection of Inborn Errors of Metabolism in Estonia. Tartu, 2018, 201 p.

278. **Mari-Anne Vals**. Congenital N-glycosylation Disorders in Estonia. Tartu, 2019, 148 p.
279. **Liis Kadastik-Eerme**. Parkinson's disease in Estonia: epidemiology, quality of life, clinical characteristics and pharmacotherapy. Tartu, 2019, 202 p.
280. **Hedi Hunt**. Precision targeting of intraperitoneal tumors with peptide-guided nanocarriers. Tartu, 2019, 179 p.
281. **Rando Porosk**. The role of oxidative stress in Wolfram syndrome 1 and hypothermia. Tartu, 2019, 123 p.
282. **Ene-Ly Jõgeda**. The influence of coinfections and host genetic factor on the susceptibility to HIV infection among people who inject drugs. Tartu, 2019, 126 p.
283. **Kristel Ehala-Aleksejev**. The associations between body composition, obesity and obesity-related health and lifestyle conditions with male reproductive function. Tartu, 2019, 138 p.
284. **Aigar Ottas**. The metabolomic profiling of psoriasis, atopic dermatitis and atherosclerosis. Tartu, 2019, 136 p.
285. **Elmira Gurbanova**. Specific characteristics of tuberculosis in low default, but high multidrug–resistance prison setting. Tartu, 2019, 129 p.
286. **Van Thai Nguyeni**. The first study of the treatment outcomes of patients with cleft lip and palate in Central Vietnam. Tartu, 2019, 144 p.
287. **Maria Yakoreva**. Imprinting Disorders in Estonia. Tartu, 2019, 187 p.
288. **Kadri Rekker**. The putative role of microRNAs in endometriosis pathogenesis and potential in diagnostics. Tartu, 2019, 140 p.
289. **Ülle Võhma**. Association between personality traits, clinical characteristics and pharmacological treatment response in panic disorder. Tartu, 2019, 121 p.
290. **Aet Saar**. Acute myocardial infarction in Estonia 2001–2014: towards risk-based prevention and management. Tartu, 2019, 124 p.
291. **Toomas Toomsoo**. Transcranial brain sonography in the Estonian cohort of Parkinson's disease. Tartu, 2019, 114 p.
292. **Lidiia Zhytnik**. Inter- and intrafamilial diversity based on genotype and phenotype correlations of Osteogenesis Imperfecta. Tartu, 2019, 224 p.
293. **Pilleriin Soodla**. Newly HIV-infected people in Estonia: estimation of incidence and transmitted drug resistance. Tartu, 2019, 194 p.
294. **Kristiina Ojamaa**. Epidemiology of gynecological cancer in Estonia. Tartu, 2020, 133 p.
295. **Marianne Saard**. Modern Cognitive and Social Intervention Techniques in Paediatric Neurorehabilitation for Children with Acquired Brain Injury. Tartu, 2020, 168 p.
296. **Julia Maslovskaja**. The importance of DNA binding and DNA breaks for AIRE-mediated transcriptional activation. Tartu, 2020, 162 p.
297. **Natalia Lobanovskaya**. The role of PSA-NCAM in the survival of retinal ganglion cells. Tartu, 2020, 105 p.

298. **Madis Rahu**. Structure and blood supply of the postero-superior part of the shoulder joint capsule with implementation of surgical treatment after anterior traumatic dislocation. Tartu, 2020, 104 p.
299. **Helen Zirnask**. Luteinizing hormone (LH) receptor expression in the penis and its possible role in pathogenesis of erectile disturbances. Tartu, 2020, 87 p.
300. **Kadri Toome**. Homing peptides for targeting of brain diseases. Tartu, 2020, 152 p.
301. **Maarja Hallik**. Pharmacokinetics and pharmacodynamics of inotropic drugs in neonates. Tartu, 2020, 172 p.
302. **Raili Müller**. Cardiometabolic risk profile and body composition in early rheumatoid arthritis. Tartu, 2020, 133 p.
303. **Sergo Kasvandik**. The role of proteomic changes in endometrial cells – from the perspective of fertility and endometriosis. Tartu, 2020, 191 p.
304. **Epp Kaleviste**. Genetic variants revealing the role of STAT1/STAT3 signaling cytokines in immune protection and pathology. Tartu, 2020, 189 p.
305. **Sten Saar**. Epidemiology of severe injuries in Estonia. Tartu, 2020, 104 p.
306. **Kati Braschinsky**. Epidemiology of primary headaches in Estonia and applicability of web-based solutions in headache epidemiology research. Tartu, 2020, 129 p.
307. **Helen Vaher**. MicroRNAs in the regulation of keratinocyte responses in *psoriasis vulgaris* and atopic dermatitis. Tartu, 2020, 242 p.
308. **Liisi Raam**. Molecular Alterations in the Pathogenesis of Two Chronic Dermatoses – Vitiligo and Psoriasis. Tartu, 2020, 164 p.
309. **Artur Vetkas**. Long-term quality of life, emotional health, and associated factors in patients after aneurysmal subarachnoid haemorrhage. Tartu, 2020, 127 p.
310. **Teele Kasepalu**. Effects of remote ischaemic preconditioning on organ damage and acylcarnitines' metabolism in vascular surgery. Tartu, 2020, 130 p.
311. **Prakash Lingasamy**. Development of multitargeted tumor penetrating peptides. Tartu, 2020, 246 p.
312. **Lille Kurvits.** Parkinson's disease as a multisystem disorder: whole transcriptome study in Parkinson's disease patients' skin and blood. Tartu, 2021, 142 p.
313. **Mariliis Põld.** Smoking, attitudes towards smoking behaviour, and nicotine dependence among physicians in Estonia: cross-sectional surveys 1982–2014. Tartu, 2021, 172 p.
314. **Triin Kikas**. Single nucleotide variants affecting placental gene expression and pregnancy outcome. Tartu, 2021, 160 p.
315. **Hedda Lippus-Metsaots**. Interpersonal violence in Estonia: prevalence, impact on health and health behaviour. Tartu, 2021, 172 p.

316. **Georgi Dzaparidze.** Quantification and evaluation of the diagnostic significance of adenocarcinoma-associated microenvironmental changes in the prostate using modern digital pathology solutions. Tartu, 2021, 132 p.
317. **Tuuli Sedman.** New avenues for GLP1 receptor agonists in the treatment of diabetes. Tartu, 2021, 118 p.
318. **Martin Padar.** Enteral nutrition, gastrointestinal dysfunction and intestinal biomarkers in critically ill patients. Tartu, 2021, 189 p.
319. **Siim Schneider.** Risk factors, etiology and long-term outcome in young ischemic stroke patients in Estonia. Tartu, 2021, 131 p.
320. **Konstantin Ridnõi.** Implementation and effectiveness of new prenatal diagnostic strategies in Estonia. Tartu, 2021, 191 p.
321. **Risto Vaikjärv.** Etiopathogenetic and clinical aspects of peritonsillar abscess. Tartu, 2021, 115 p.
322. **Liis Preem.** Design and characterization of antibacterial electrospun drug delivery systems for wound infections. Tartu, 2022, 220 p.
323. **Keerthie Dissanayake.** Preimplantation embryo-derived extracellular vesicles: potential as an embryo quality marker and their role during the embryo-maternal communication. Tartu, 2022, 203 p.
324. **Laura Viidik.** 3D printing in pharmaceutics: a new avenue for fabricating therapeutic drug delivery systems. Tartu, 2022, 139 p.
325. **Kasun Godakumara.** Extracellular vesicle mediated embryo-maternal communication – A tool for evaluating functional competency of preimplantation embryos. Tartu, 2022, 176 p.
326. **Hindrek Teder.** Developing computational methods and workflows for targeted and whole-genome sequencing based non-invasive prenatal testing. Tartu, 2022, 138 p.
327. **Jana Tuusov.** Deaths caused by alcohol, psychotropic and other substances in Estonia: evidence based on forensic autopsies. Tartu, 2022, 157 p.
328. **Heigo Reima.** Colorectal cancer care and outcomes – evaluation and possibilities for improvement in Estonia. Tartu, 2022, 146 p.
329. **Liisa Kuhi.** A contribution of biomarker collagen type II neoepitope C2C in urine to the diagnosis and prognosis of knee osteoarthritis. Tartu, 2022, 157 p.
330. **Reeli Tamme.** Associations between pubertal hormones and physical activity levels, and subsequent bone mineral characteristics: a longitudinal study of boys aged 12–18. Tartu, 2022, 118 p.
331. **Deniss Sõritsa.** The impact of endometriosis and physical activity on female reproduction. Tartu, 2022, 152 p.
332. **Mohammad Mehedi Hasan.** Characterization of follicular fluid-derived extracellular vesicles and their contribution to periconception environment. Tartu, 2022, 194 p.
333. **Priya Kulkarni.** Osteoarthritis pathogenesis: an immunological passage through synovium-synovial fluid axis. Tartu, 2022, 268 p.

334. **Nigul Ilves.** Brain plasticity and network reorganization in children with perinatal stroke: a functional magnetic resonance imaging study. Tartu, 2022, 169 p.
335. **Marko Murruste.** Short- and long-term outcomes of surgical management of chronic pancreatitis. Tartu, 2022, 180 p.
336. **Marilin Ivask.** Transcriptomic and metabolic changes in the WFS1-deficient mouse model. Tartu, 2022, 158 p.
337. **Jüri Lieberg.** Results of surgical treatment and role of biomarkers in pathogenesis and risk prediction in patients with abdominal aortic aneurysm and peripheral artery disease. Tartu, 2022, 160 p.
338. **Sanna Puusepp.** Comparison of molecular genetics and morphological findings of childhood-onset neuromuscular disorders. Tartu, 2022, 216 p.
339. **Khan Nguyen Viet.** Chemical composition and bioactivity of extracts and constituents isolated from the medicinal plants in Vietnam and their nanotechnology-based delivery systems. Tartu, 2023, 172 p.
340. **Getnet Balcha Midekessa.** Towards understanding the colloidal stability and detection of Extracellular Vesicles. Tartu, 2023, 172 p.
341. **Kristiina Sepp.** Competency-based and person-centred community pharmacy practice – development and implementation in Estonia. Tartu, 2023, 242 p.
342. **Linda Sõber.** Impact of thyroid disease and surgery on patient's quality of voice and swallowing. Tartu, 2023, 114 p.
343. **Anni Lepland.** Precision targeting of tumour-associated macrophages in triple negative breast cancer. Tartu, 2023, 160 p.
344. **Sirje Sammul.** Prevalence and risk factors of arterial hypertension and cardiovascular mortality: 13-year longitudinal study among 35- and 55-year-old adults in Estonia and Sweden. Tartu, 2023, 158 p.
345. **Maarjaliis Paavo**. Short-Wavelength and Near-Infrared Autofluorescence Imaging in Recessive Stargardt Disease, Choroideremia, *PROM1*-Macular Dystrophy and Ocular Albinism. Tartu, 2023, 202 p.
346. **Kaspar Ratnik.** Development of predictive multimarker test for preeclampsia in early and late pregnancy. Tartu, 2023, 134 p.
347. **Kärt Simre.** Development of coeliac disease in two populations with different environmental backgrounds. Tartu, 2023, 161 p.
348. **Qurat Ul Ain Reshi.** Characterization of the maternal reproductive tract and spermatozoa communication during periconception period via extracellular vesicles. Tartu, 2023, 182 p.
349. **Stanislav Tjagur.** *Mycoplasma genitalium* and other sexually transmitted infections causing urethritis – their prevalence, impact on male fertility parameters and prostate health. Tartu, 2023, 225 p.
350. **Lagle Lehes.** The first study of voice and resonance related treatment outcomes of Estonian cleft palate children. Tartu, 2023, 126 p.
351. **Liis Ilves.** Metabolomic profiling of chronic inflammatory skin diseases. Tartu, 2023, 146 p.

352. **Marina Šunina**. Flow cytometric analysis of T and B cell properties in healthy donors and subjects with vitiligo. Tartu, 2023, 164 p.
353. **Jaanus Suumann.** Gastric biomarkers and their dynamics as a less invasive method to evaluate stomach health in bariatric surgery patients. Tartu, 2023, 122 p.
354. **Ele Hanson.** Clinical and biochemical markers for the prediction and early diagnosis of pregnancy related complications. Tartu, 2023, 145 p.
355. **Priit Pauklin.** Hemodynamic and biochemical characteristics of patients with atrial fibrillation and anticoagulation of ≥65-year-old patients with atrial fibrillation in Estonia. Tartu, 2023, 144 p.
356. **Triinu Keskpaik.** Quality Indicators and Non-Ischemic Myocardial Injury in Emergency Medicine. Tartu, 2023, 121 p.
357. **Laura Roht.** Hereditary colorectal cancer syndromes in Estonia. Tartu, 2023, 178 p.
358. **Norman Ilves.** Risk factors and onset time of periventricular hemorrhagic infarction in preterm born children and periventricular venous infarction in term born children. Tartu, 2024, 177 p.
359. **Edgar Lipping.** Postoperative antibacterial therapy in complicated appendicitis and appendectomy in pregnancy. Tartu, 2024, 121 p.
360. **Celia Teresa Pozo Ramos.** Preparation and assessment of antimicrobial electrospun matrices for prospective applications in wound healing. Tartu, 2024, 203 p.
361. **Karl Kuusik.** Effects of remote ischaemic preconditioning on arterial stiffness, organ damage and metabolomic profile in patients with lower extremity artery disease. Tartu, 2024, 173 p.
362. **Kelli Somelar-Duracz.** The molecular and cellular mechanisms of brain plasticity impairing factors. Tartu, 2024, 245 p.
363. **Aleksei Baburin.** Breast cancer incidence, mortality and survival in Estonia in the context of health care system changes and screening. Tartu, 2024, 130 p.
364. **Marina Loid.** Molecular and cellular determinants of healthy receptive and aged endometrium. Tartu, 2024, 159 p.
365. **Ulvi Vaher.** Epilepsy after ischemic perinatal stroke in term born children: neuroimaging predictors, clinical course and cognitive outcome. Tartu, 2024, 160 p.
366. **Allan Tobi.** Development of Smart Nanoparticles for Experimental Treatment of Cancer. Tartu, 2024, 160 p.
367. **Leho Rips.** The influence of vitamin D on the physical performance of conscripts in the Estonian Defence Forces. Tartu, 2024, 147 p.
368. **Kati Kärberg.** Factors and markers predicting subclinical atherosclerosis in type 2 diabetes. Tartu, 2024, 161 p.
369. **Valeria Sidorenko.** Novel anthracycline-loaded nanoparticles for precision cancer therapy. Tartu, 2024, 183 p.

370. **Kadri Kõivumägi.** Acute gastroenteritis hospitalizations in Estonia after implementation of universal mass vaccination against rotavirus. Tartu, 2024, 150 p.
371. **Ingrid Oit-Wiscombe.** Genetic markers of enzymatics in the pathogenesis of chronic obstructive pulmonary disease as a systemic disease and the effects of antioxidant peptides. Tartu, 2025, 170 p.
372. **Gerli Mõts.** Ethical issues in nursing before and during the COVID-19 pandemic: a multi-method study. Tartu, 2025, 150 p.
373. **Annika Valner.** Changes in structure and function of extremities in early rheumatoid arthritis. Tartu, 2025, 129 p.
374. **Meruert Sarsenova.** Molecular and cellular landscape of endometriosis. Tartu, 2025, 157 p.
375. **Mailis Liiv.** Role of mitochondrial dynamics in Wolfram syndrome. Tartu, 2025, 180 p.
376. **Marta Velgan.** Addressing the family physician shortage: Career and migration intentions in Estonia and Europe. Tartu, 2025, 164 p.