

PBMCpedia: a harmonized PBMC scRNA-seq database with unified mapping and enhanced celltype annotation

Emma S. Hoffmann¹, Lara Hombrecher¹, Ian F. Diks¹, Matthias Flotho^{1,2}, Pascal Hirsch¹, Andreas Keller^{1,2,3,*}, Friederike Grandke^{1,2,†}

¹Clinical Bioinformatics, Saarland University, 66123 Saarbrücken, Germany

²Helmholtz Institute for Pharmaceutical Research Saarland–Helmholtz Centre for Infection Research, 66123 Saarbrücken, Germany

³PharmaScienceHub, Saarland University, 66123 Saarbrücken, Germany

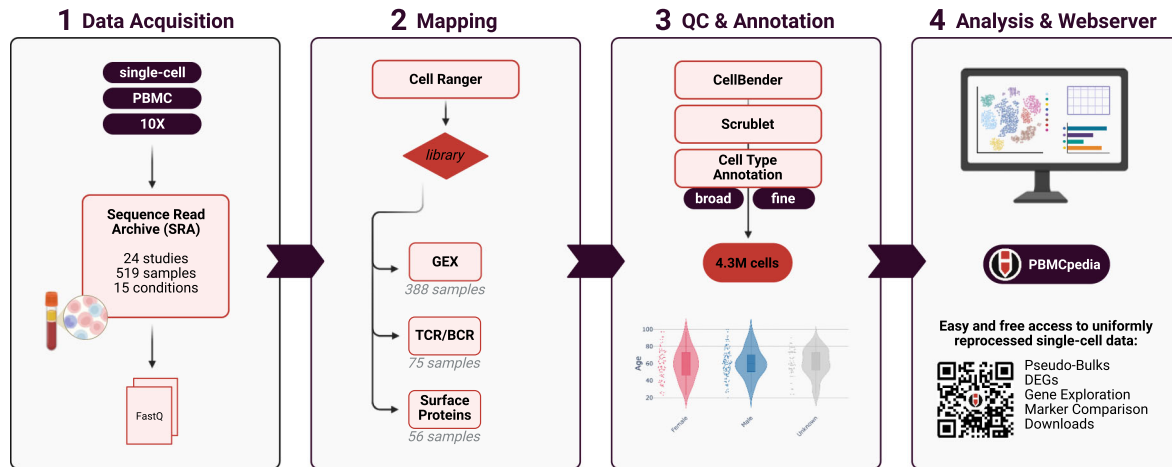
*To whom correspondence should be addressed. Email: andreas.keller@ccb.uni-saarland.de

†These authors jointly supervised the work.

Abstract

Single-cell transcriptomic studies of peripheral blood mononuclear cells (PBMCs) offer valuable insights into immune states across diverse biological conditions, yet cross-study integration remains difficult due to divergent preprocessing and annotations. PBMCpedia addresses this by uniformly reprocessing 519 samples (over 4.3 million cells) from 24 publicly available single-cell RNA sequencing studies using a standardized pipeline with consistent quality control and hierarchical cell type annotation. Spanning 14 disease contexts, including autoimmune, infectious, and neurodegenerative disorders, as well as healthy controls, PBMCpedia supports metadata-aware comparisons across diseases, cell types, sexes, and age groups. It also includes T-cell receptor/B-cell receptor repertoire data for 75 samples and surface protein measurements for 56 samples, enabling integrative immune profiling at both the transcriptomic and proteogenomic levels. To support exploration and accessibility, we provide an interactive web interface (<https://web.ccb.uni-saarland.de/pbmcpedia/>) for querying gene expression, marker genes, and pathway enrichment across cell types, conditions, sexes, and age groups. PBMCpedia fills a critical gap by offering a transparent, harmonized, and disease-diverse PBMC resource designed for cross-study immune profiling and discovery.

Graphical abstract



Introduction

Single-cell RNA sequencing (scRNA-seq) of peripheral blood mononuclear cells (PBMCs) is a powerful approach to profile immune heterogeneity across diverse biological contexts, including infection, autoimmunity, aging, and neurodegeneration. Despite the availability of hundreds of public PBMC datasets, comparative analyses remain difficult due to inconsistent preprocessing, divergent annotation schemes, and batch effects. These chal-

lenges limit reproducibility and hinder the reuse of existing datasets.

Several large-scale PBMC atlases have addressed specific aspects of this issue. The Allen Institute's Human Immune Health Atlas [1] includes over 16 million immune cells from healthy donors and provides a standardized annotation framework stratified by age, but excludes disease cohorts. Jiménez-Gracia *et al.* [2] reprocessed raw data from 356 patients across 18 inflammatory conditions, generating

Received: August 15, 2025. Revised: October 14, 2025. Accepted: October 15, 2025

© The Author(s) 2025. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

the “Inflammation Landscape” using interpretable machine learning. However, their harmonized dataset is not accessible via an interactive platform or downloadable in full. Conolly *et al.* [3] created a PBMC aging atlas by integrating 2.8 million cells from 35 studies, but used heterogeneous preprocessing pipelines and do not provide a centralized portal. CancerSCEM 2.0 [4] aggregates over 1400 datasets across 74 cancer types, with limited emphasis on PBMCs.

To address these limitations, we developed PBMCpedia, a harmonized, multidisease PBMC atlas comprising 4 293 193 single-cell transcriptomes from 519 samples across 24 publicly available scRNA-seq studies (Supplementary Table S1). PBMCpedia introduces three core advances. First, all FASTQ files are uniformly reprocessed with standardized quality control, batch correction, and hierarchical annotation, ensuring comparability across studies. Whenever available, all samples include metadata on age, sex, and disease status, enabling stratified, demographic-aware analyses. Second, the resource integrates downstream analysis modules—including differential expression and pathway enrichment—exposed through a reproducible, application programming interface (API)-backed web interface. Third, PBMCpedia combines transcriptomic profiles with paired T-cell receptor (TCR)/B-cell receptor (BCR) repertoires and surface protein measurements, enabling multimodal exploration of immune states at single-cell resolution. Together, these features establish PBMCpedia as both a harmonized reference and a practical diagnostic toolkit for cross-study immune analysis, extending beyond prior PBMC atlases by combining breadth, standardization, and accessible downstream functionality.

Materials and methods

We included 24 publicly available scRNA-seq PBMC datasets covering 14 disease conditions and healthy controls. All datasets were generated with 10x Genomics technology, included raw FASTQ files, and contained usable sample metadata (age, sex, disease status). Datasets lacking raw data or essential metadata were excluded. In total, 519 samples and 4 293 193 high-quality cells were retained following uniform quality control.

Data acquisition

Sequencing data were retrieved from the Sequence Read Archive (SRA) using the SRA Toolkit (v3.0.0). We downloaded .sra files using `prefetch`, verified them with `vdb-validate`, and extracted paired reads using `fasterq-dump` with `-split-files` and `-include-technical` to preserve index reads. Because submission formats varied, files were manually curated to ensure correct pairing and orientation.

Metadata were collected via literature review and the NCBI SRA Run Selector. Age was treated as a continuous variable; sex was categorized as male, female, or unknown to reflect inconsistent reporting.

Uniform preprocessing and quality control

All datasets were aligned with Cell Ranger (v9.0.0) to GRCh38-2024-A. Count matrices were processed using Scanpy (v1.11.1) [5], CellBender (v0.3.2) [6] for ambient RNA correction, and Scrublet (via Scanpy) [7] for doublet de-

tection. Doublet detection was performed on a per-sample basis using a combined feature matrix that included both transcriptomic and surface protein measurements whenever CITE-seq data were available. Cells with fewer than 200 genes or above the 98th percentile for gene/unique molecular identifier (UMI) counts were removed. We did not apply a strict filter based on mitochondrial gene content because, in PBMC datasets, distinguishing truly stressed or low-quality cells from biologically relevant variation in mitochondrial reads is challenging. Immune cell subsets naturally vary in mitochondrial content, and applying a universal cutoff could disproportionately remove specific cell types.

Integration was performed using Harmony (via `harmonypy`, version 0.0.10) [8] on 15 PCs, correcting batch effects by project and sample ID. UMAP visualizations confirmed dataset mixing and clustering by immune identity. UMAP for webserver was downsampled to 10k cells per cell type.

Multiplexed samples were demultiplexed by mapping cells to published donor assignments. This ensured accurate donor attribution.

Annotation strategy

Cell type labels were transferred using the Allen Institute’s `cell_type_mapper` tool at two hierarchical levels: AIFI_L1 (major lineages) and AIFI_L2 (subtypes) using the Human Immune Health Atlas [1] as reference. This annotation supports flexible downstream analysis. The cell type labels were manually verified using known marker genes.

Differential expression and pathway analysis

Pseudobulk differential gene expression (DGE) analysis was conducted using Scanpy (v1.11.2) [5] and `limma` (v3.58.1 via `rpy2` v3.5.11) [9], comparing disease versus control samples across cell types, annotation levels, sexes, and age groups. We performed preranked GSEA using GSEAPy (v1.1.9) with the “GO_Biological_Process_2021” [10, 11] library. Genes were ranked by \log_2 fold-change (\logFC), retaining the top 800 by absolute \logFC and excluding those with $|\logFC| < 0.15$. Gene sets with 10–2000 members were considered, using 250 permutations and 3 threads. For each comparison, DGE was computed using the Wilcoxon rank-sum test with Benjamini–Hochberg correction. Stratifications included sex (female/male versus same-sex controls) and age group (young: <25 , adult: 25–64, elderly: >64), with an additional pooled comparison to maximize power.

TCR and BCR data extraction

For all datasets containing immune repertoire information, we processed raw FASTQ files using the 10x Genomics Cell Ranger (v9.0.0) `vdj` or `multi` pipeline. The corresponding V(D)J reference (GRCh38, 10x Genomics v9.0.0) was used to align reads and assemble full-length, paired TCR and BCR sequences. Output files (`filtered_contig_annotations.csv`) were parsed to extract chain type, V/D/J gene calls, and pairing information using `scirpy` (v0.22.1) [12]. Paired $\alpha\beta$ (TCR), and light/heavy (BCR) chain assignments were determined per cell barcode.

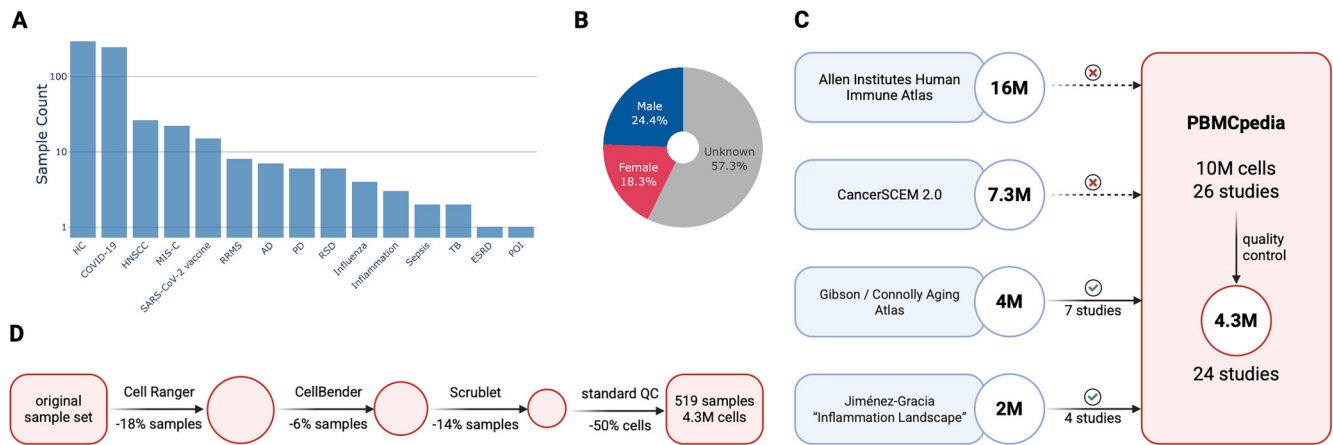


Figure 1. (A) Disease distribution. (B) Overall sex distribution across all 519 samples. (C) Overlap of studies included in PBMCpedia compared to four major PBMC atlases: the Allen Immune Human Health Atlas, Jiménez-Gracia's Inflammation Landscape, Connolly *et al.*'s aging atlas, and CancerSCEM 2.0. PBMCpedia shares a small subset of studies with Connolly and Jiménez-Gracia but is otherwise largely nonoverlapping. (D) Cell and sample filtering across preprocessing steps. Created in BioRender: Hoffmann, E. (2026) <https://BioRender.com/lqe6s4z>.

Web server implementation

PBMCpedia is hosted as a Django-based web application with visualizations rendered via Plotly and DataTables. An API provides easy and fast programmatic access to the data. Integrated knowledge bases, including CellMarker [13], PanglaoDB [14], MSigDB [15, 16], and the Human Protein Atlas [17], enhance interpretability. Contextual links for each gene allow real-time reference to known functions and expression profiles.

Results

Overview

PBMCpedia fills a critical need for a harmonized, multidisease PBMC reference by integrating 24 publicly available scRNA-seq datasets into a harmonized reference atlas of human PBMCs. After uniform preprocessing and stringent quality control, the dataset comprises 4 293 193 high-quality transcriptomes from 519 samples, covering 14 disease conditions and healthy controls (Fig. 1A).

The resource spans a broad range of biological contexts, including infectious diseases (e.g. COVID-19, tuberculosis), autoimmune disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis), and neurodegeneration (e.g. Alzheimer's and Parkinson's disease). Each sample is (where available) annotated with standardized metadata including disease status, donor sex, and age, enabling stratified analysis across demographic variables (Fig. 1B).

All datasets were processed through a consistent pipeline and integrated using Harmony to correct for batch effects while preserving biological structure. Cells were annotated using a two-level hierarchical framework aligned with the Allen Human Immune Health Atlas, capturing both broad immune lineages and fine-grained subtypes.

Where available, PBMCpedia also includes multimodal profiles: 75 samples contain paired TCR/BCR repertoire sequencing and 56 include CITE-seq-based surface protein quantification. Our final PBMCpedia object retains all three modalities as separate layers. We did not integrate them into a single embedding, leaving users the flexibility to combine or

jointly analyze modalities according to their specific research needs.

In total, PBMCpedia provides a reproducible, richly annotated PBMC resource suitable for comparative analysis, biomarker discovery, and immunological exploration across diverse diseases and patient subgroups.

Comparison to existing resources

Several large-scale PBMC resources exist, but each addresses only part of the challenge. The Allen Human Immune Health Atlas [1] profiles over 16 million immune cells from healthy donors with a focus on age-related changes, but it excludes disease cohorts. Jiménez-Gracia *et al.* [2] reprocessed 2 million PBMCs from 356 patients across 18 inflammatory conditions in the "Inflammation Landscape," yet the harmonized dataset is not available as a full download or interactive browser. Connolly *et al.* [3] integrated 2.8 million cells from 35 public studies into an aging-focused PBMC atlas, supplemented by a 1.2-million-cell validation cohort. However, they used heterogeneous pipelines and do not provide a unified, interactive portal. CancerSCEM 2.0 [4] aggregates over 1400 datasets across 74 cancer types, including PBMCs. However, its primary emphasis is on tumor microenvironments rather than systemic immune profiling.

PBMCpedia overcomes these limitations by uniformly reprocessing raw FASTQ files using a consistent pipeline with stringent quality control and Harmony-based integration, eliminating batch effects from study-specific workflows. The atlas covers 14 disease contexts and healthy controls, supports hierarchical annotations aligned with the Allen framework, and provides metadata-stratified differential expression and pathway enrichment analyses. Beyond harmonization, PBMCpedia uniquely incorporates multimodal data: 75 samples with paired TCR/BCR repertoires and 56 with CITE-seq surface proteins, enabling joint exploration of transcriptomic, clonotypic, and proteomic variation (Fig. 1C).

To ensure robust and reproducible insights, we applied strict preprocessing and quality control filters across all studies. These include thresholds for gene and UMI counts, and doublet exclusion based on both transcriptomic and pro-

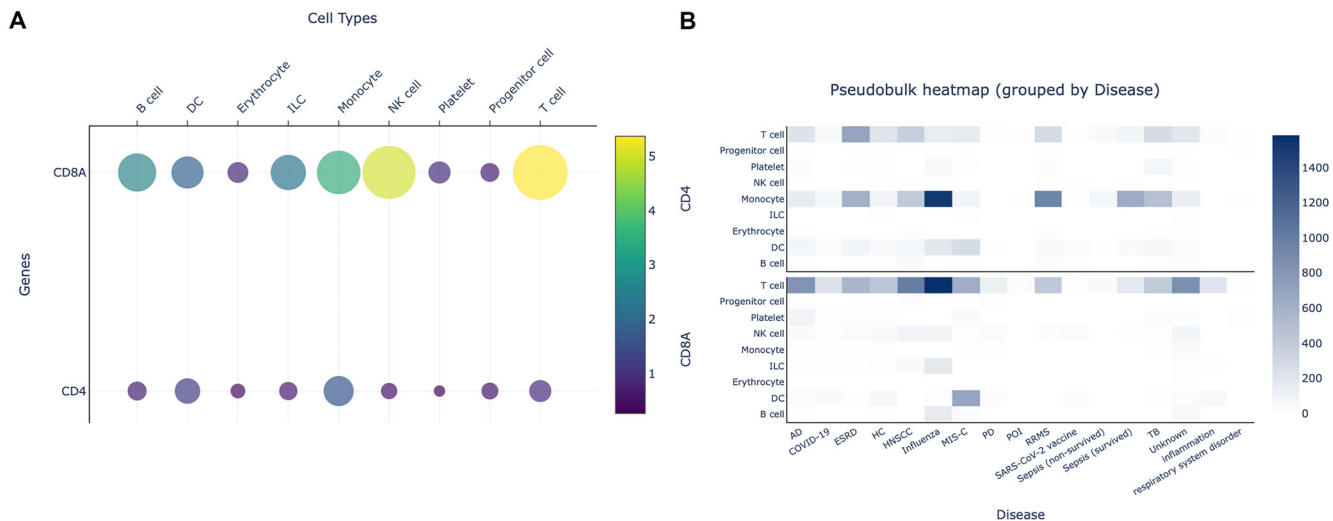


Figure 2. Examples from the PBMCpedia webserver. **(A)** Dot plot of *CD4* and *CD8A*, visualizing their relative expression across annotated cell types at different resolution levels. The dot size represents the proportion of expressing cells, while color intensity reflects average expression. **(B)** Heatmap displaying *CD4* and *CD8A* across broad immune lineages. Created in BioRender: Hoffmann, E. (2026) <https://BioRender.com/lqe6s4z>.

teomic profiles using Scrublet. While these steps are essential to minimize technical artifacts and ensure biological relevance, they come at a cost: over 50% (5 625 609) of initially captured cells were discarded, resulting in a substantial shrinkage of the dataset. We annotated with a binary flag indicating whether cells exceed a 10% mitochondrial threshold, and the counts per project are reported in [Supplementary Table S2](#), allowing users to assess and optionally filter cells according to their own analytical needs. This rigorous curation underscores PBMCpedia's focus on data integrity over raw cell quantity, yielding a high-confidence reference atlas suitable for downstream machine learning and systems immunology applications (Fig. 1 D).

All results are freely accessible through a user-friendly web interface, with downloadable data and code to ensure transparency and reuse. PBMCpedia thus fills a critical gap in the single-cell landscape by offering a reproducible, disease-aware resource for comparative immunology.

Cross-disease immune signatures

Because PBMCpedia spans diverse biological contexts, all processed through a unified pipeline, it enables direct, systematic comparisons across conditions (Fig. 2). Users can explore differences and commonalities in immune responses across diseases (e.g. influenza versus COVID-19), age groups (e.g. young versus adult), and sexes (e.g. male versus female) without the confounding effects of batch differences or inconsistent annotations. This level of comparability is difficult to achieve when integrating data from separate studies, where technical variation and divergent labeling schemes often obscure true biological signals.

With that, PBMCpedia provides a framework not only for hypothesis-driven analysis but also for diagnostic exploration, identifying patterns of immune dysregulation that recur across conditions or stratify patient subgroups. This supports demographic filtering, allowing users to restrict analyses to subsets, such as female patients over 60. Even within such filters, sample sizes remain large enough to detect meaningful differences, facilitating age- and sex-aware disease comparisons.

Discussion

PBMCpedia enables robust, comparative immune profiling across diseases by offering a harmonized and openly accessible dataset with standardized preprocessing. By uniformly reprocessing raw data, applying stringent quality control, and integrating hierarchical cell type annotations, the atlas ensures reproducibility and comparability across studies. Its two-level annotation system serves both novice users and expert immunologists, providing flexibility for various analytical depths. Stratified analyses by sex and age, together with multimodal data where available (TCR/BCR, CITE-seq), further enhance its utility for exploring immune heterogeneity at multiple molecular levels.

PBMCpedia's technical innovations lie in its multistudy harmonization, transparent QC metrics, and integrated multi-omic profiles, which collectively establish a platform for both discovery and diagnostic evaluation. The combination of consistent preprocessing, batch correction with Harmony, and hierarchical annotation addresses a key limitation of prior atlases. Moreover, the availability of standardized metadata and multimodal modalities supports integrative analyses of transcriptomic, clonotypic, and proteomic variation, enhancing both biological insight and methodological rigor.

To illustrate the diagnostic potential of the database, we applied a simple multilayer perceptron as a proof of concept for incremental training across studies. This experiment does not introduce a new annotation model; rather, it demonstrates how even basic classifiers can detect studies that systematically degrade performance, guiding preprocessing or reannotation and highlighting dataset consistency.

As a community-curated resource, PBMCpedia reflects the availability of public data. Some diseases, including cancers and rare conditions, remain underrepresented. Future versions will incorporate new studies as they become available, leveraging our reproducible pipeline to expand coverage. This includes especially the raw data of the Allen Immune, facilitated by the already shared annotation, and the "Inflammation Landscape" [2].

While transcriptomics is the core modality, PBMCpedia already includes some multimodal datasets. Future updates will

expand this further, including spatial transcriptomics and cell-cell interaction networks. Despite robust batch correction, residual cohort differences may persist. We encourage users to leverage the provided metadata, especially age, sex, and control samples from multiple studies, for careful interpretation.

PBMCpedia addresses the growing need for reproducible, cross-disease immune profiling by enabling scalable, metadata-aware analysis and therefore supporting both discovery and hypothesis-driven research. All data, code, and results are freely available through an interactive web platform.

By prioritizing accessibility, consistency, and scalability, PBMCpedia serves as a valuable tool for immunologists, computational biologists, and clinical researchers. We invite community feedback and contributions to expand and refine PBMCpedia as an open resource for immune research.

Conflict of interest

None declared.

Acknowledgements

The authors used ChatGPT (OpenAI, GPT-4o) to assist with phrasing and language refinement during manuscript preparation. All scientific content, data analysis, figure creation, and the initial manuscript draft were carried out by the authors, who also thoroughly reviewed and revised all text.

The Graphical Abstract was created in BioRender. Hoffmann, E. (2025) <https://BioRender.com/j9pa3p7>

Author contribution: Emma S. Hoffmann (Data curation, Methodology, Writing—original draft), Lara Hombrecher (Methodology), Ian F. Diks (Data curation, Methodology), Matthias Flotho (Conceptualization, Methodology), Pascal Hirsch (Methodology), Andreas Keller (Funding acquisition, Supervision), Friederike Grandke (Conceptualization, Data curation, Methodology, Supervision, Writing—original draft).

Supplementary data

Supplementary data is available at NAR online.

Funding

The project was supported by the Michael J. Fox Foundation for Parkinson's Research (MJFF-021418 to A.K. and T.W.-C.). Computational resources used within this study were financed through the DFG (469073465 to A.K.). Funding to pay the Open Access publication charges for this article was provided by the Saarland University.

Data availability

PBMCpedia is freely available at <https://web.ccb.uni-saarland.de/pbmcpedia/>.

References

- Gong Q, Sharma M, Kuan EL *et al.* Longitudinal multi-omic immune profiling reveals age-related immune cell dynamics in healthy adults. *Nature* 2025. <https://doi.org/10.1038/s41586-025-09686-5>
- Jiménez-Gracia L, Maspero D, Aguilar-Fernández S *et al.* Interpretable Inflammation Landscape of Circulating Immune cells. *bioRxiv*, <https://doi.org/10.1101/2023.11.28.568839> 1 April 2025, preprint: not peer reviewed.
- Connolly E, Pan T, Aluru M *et al.* Loss of immune cell identity with age inferred from large atlases of single cell transcriptomes. *Aging Cell* 2024;23:e14306. <https://doi.org/10.1111/acer.14306>
- Zeng J, Nie Z, Shang Y *et al.* CancerSCEM 2.0: an updated data resource of single-cell expression map across various human cancers. *Nucleic Acids Res* 2024;53:D1278–86. <https://doi.org/10.1093/nar/gkac954>
- Wolf FA, Angerer P, Theis FJ. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol* 2018;19:15. <https://doi.org/10.1186/s13059-017-1382-0>
- Fleming SJ, Chaffin MD, Arduini A *et al.* Unsupervised removal of systematic background noise from droplet-based single-cell experiments using CellBender. *Nat Methods* 2019;20:1323–35. <https://doi.org/10.1038/s41592-023-01943-7>
- Wolock SL, Lopez R, Klein AM. Scrublet: computational identification of cell doublets in single-cell transcriptomic data. *Cell Syst* 2019;8:281–91. <https://doi.org/10.1016/j.cels.2018.11.005>
- Korsunsky I, Fan J, Slowikowski K *et al.* Fast, sensitive, and accurate integration of single cell data with Harmony. *Nat Methods* 2019;16:1289–96. <https://doi.org/10.1038/s41592-019-0619-0>
- Ritchie ME, Phipson B, Wu DL *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 2015;43:e47. <https://doi.org/10.1093/nar/gkv007>
- Aleksander SA, Ballhoff JP, Carbon S *et al.* The Gene Ontology Knowledgebase in 2023. *Genetics* 2023;224:iyad031. <https://doi.org/10.1093/genetics/iyad031>
- Ashburner M, Ball CA, Blake JA *et al.* Gene Ontology: tool for the unification of biology. *Nat Genet* 2000;25:25–9. <https://doi.org/10.1038/75556>
- Sturm G, Szabo T, Fotakis G *et al.* Scirpy: a Scanpy extension for analyzing single-cell T-cell receptor-sequencing data. *Bioinformatics* 2020;36:4817–8. <https://doi.org/10.1101/2020.04.10.035865>
- Hu C, Li T, Xu Y *et al.* CellMarker 2.0: an updated database of manually curated cell markers in human/mouse and web tools based on scRNA-seq data. *Nucleic Acids Res* 2022;51:D870–6. <https://doi.org/10.1093/nar/gkac947>
- Franzén O, Gan LM, Björkegren JLM. PanglaoDB: a web server for exploration of mouse and human single-cell RNA sequencing data. *Database* 2019;2019:baz046. <https://doi.org/10.1093/database/baz046>
- Liberzon A, Birger C, Thorvaldsdóttir H *et al.* The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst* 2015;1:417–25.
- Liberzon A, Subramanian A, Pinchback R *et al.* Molecular signatures database (MSigDB) 3.0. *Bioinformatics* 2011;27:12:1739–40. <https://doi.org/10.1093/bioinformatics/btr260>
- Uhlén M, Fagerberg L, Hallström BM *et al.* Tissue-based map of the human proteome. *Science* 2015;347:1260419. <https://doi.org/10.1126/science.1260419>
- Varoquaux G, Buitinck L, Louppe G *et al.* Scikit-learn: machine learning without learning the machinery. *GetMobile Mob Comput Commun* 2015;19:29–33. <https://doi.org/10.1145/2786984.2786995>
- Luo OJ, Lei W, Zhu G *et al.* Multidimensional single-cell analysis of human peripheral blood reveals characteristic features of the immune system landscape in aging and frailty. *Nat Aging* 2022;2:348–64. <https://doi.org/10.1038/s43587-022-00198-9>
- Smith SL, Kennedy PR, Stacey KB *et al.* Diversity of peripheral blood human NK cells identified by single-cell RNA sequencing. *Blood Adv* 2020;47:1388–406. <https://doi.org/10.1182/bloodadvances.2019000699>
- Yang C, Siebert JR, Burns R *et al.* Heterogeneity of human bone marrow and blood natural killer cells defined by single-cell

- transcriptome. *Nat Commun* 2019;10:3931. <https://doi.org/10.1038/s41467-019-11947-7>
22. Yazar S, Alquicira-Hernández J, Wing K *et al*. Single-cell eQTL mapping identifies cell type-specific genetic control of autoimmune disease. *Science* 2022;376:eabf3041. <https://doi.org/10.1126/science.abf3041>
 23. Volden R, Vollmers C. Single-cell isoform analysis in human immune cells. *Genome Biol* 2022;23:47. <https://doi.org/10.1186/s13059-022-02615-z>
 24. Xu H, Jia J. Single-Cell RNA sequencing of peripheral blood reveals immune cell signatures in Alzheimer's disease. *Front Immunol* 2021;12:645666. <https://doi.org/10.3389/fimmu.2021.645666>
 25. Xiong L, Xue LL, Du RL *et al*. Single-cell RNA sequencing reveals B cell-related molecular biomarkers for Alzheimer's disease. *Exp Mol Med* 2021;53:1888–901. <https://doi.org/10.1038/s12276-021-00714-8>
 26. Bai H, Ma J, Mao W *et al*. Identification of TCR repertoires in asymptomatic COVID-19 patients by single-cell T-cell receptor sequencing. *Blood Cells Mol Dis* 2022;97:102678. <https://doi.org/10.1016/j.bcmd.2022.102678>
 27. Ivanova EN, Shwetar J, Devlin JC *et al*. mRNA COVID-19 vaccine elicits potent adaptive immune response without the acute inflammation of SARS-CoV-2 infection. *iScience* 2023;26:108572. <https://doi.org/10.1016/j.isci.2023.108572>
 28. van der Wijst MG, Vazquez SE, Hartoularos GC *et al*. Type I interferon autoantibodies are associated with systemic immune alterations in patients with COVID-19. *Sci Trans Med* 2021;13:eabh2624. <https://doi.org/10.1126/scitranslmed.abh2624>
 29. Liu C, Martins AJ, Lau WW *et al*. Time-resolved systems immunology reveals a late juncture linked to fatal COVID-19. *Cell* 2021;184:1836–57. <https://doi.org/10.1016/j.cell.2021.02.018>
 30. Ramaswamy A, Brodsky NN, Sumida TS *et al*. Immune dysregulation and autoreactivity correlate with disease severity in SARS-CoV-2-associated multisystem inflammatory syndrome in children. *Immunity* 2021;54:1083–95. <https://doi.org/10.1016/j.immuni.2021.04.003>
 31. Unterman A, Sumida TS, Nouri N *et al*. Single-cell multi-omics reveals dyssynchrony of the innate and adaptive immune system in progressive COVID-19. *Nat Commun* 2022;13:440. <https://doi.org/10.1038/s41467-021-27716-4>
 32. Lee JS, Park S, Jeong HW *et al*. Immunophenotyping of COVID-19 and influenza highlights the role of type I interferons in development of severe COVID-19. *Sci Immunol* 2020;5:eabd1554. <https://doi.org/10.1126/sciimmunol.abd1554>
 33. Hwang N, Huh YJ, Bu SM *et al*. Single-cell sequencing of PBMC characterizes the altered transcriptomic landscape of classical monocytes in BNT162b2-induced myocarditis. *Front Immunol* 2022;13:979188. <https://doi.org/10.3389/fimmu.2022.979188>
 34. Waickman AT, Friberg H, Gromowski GD *et al*. Temporally integrated single cell RNA sequencing analysis of PBMC from experimental and natural primary human DENV-1 infections. *PLoS Pathogens* 2021;17 1:e1009240. <https://doi.org/10.1371/journal.ppat.1009240>
 35. Zhang R, Liu X, Ma Y *et al*. Identification of Cell-Cell Communications by Single-Cell RNA Sequencing in End Stage Renal Disease Provides New Insights into Immune Cell Heterogeneity. *J Inflamm Res* 2023;16:4977–5000. <https://doi.org/10.2147/JIR.S424911>
 36. Cillo AR, Kürten CHL, Tabib T *et al*. Immune landscape of viral- and carcinogen-driven head and neck cancer. *Immunity* 2019;52:183–99. <https://doi.org/10.1016/j.immuni.2019.11.014>
 37. Zhou L, Adrianto I, Wang J *et al*. Single-Cell RNA-Seq Analysis Uncovers Distinct Functional Human NKT Cell Sub-Populations in Peripheral Blood. *Front Cell Dev Biol* 2020;8:384. <https://doi.org/10.3389/fcell.2020.00384>
 38. Seyedadr M, Wang Y, Elzoheiry M *et al*. IL-11 induces NLRP3 inflammasome activation in monocytes and inflammatory cell migration to the central nervous system. *Proc Natl Acad Sci USA* 2023;120:e2221007120. <https://doi.org/10.1073/pnas.2221007120>
 39. Wang P, Luo M, Zhou W *et al*. Global characterization of peripheral B cells in Parkinson's disease by single-cell RNA and BCR sequencing. *Front Immunol* 2022;13:814239. <https://doi.org/10.3389/fimmu.2022.814239>
 40. Zhang C, Yu D, Mei Y *et al*. Single-cell RNA sequencing of peripheral blood reveals immune cell dysfunction in premature ovarian insufficiency. *Front Endocrinol* 2023;14:1129657. <https://doi.org/10.3389/fendo.2023.1129657>
 41. Qiu X, Li J, Bonenfant J *et al*. Dynamic changes in human single-cell transcriptional signatures during fatal sepsis. *J Leukocyte Biol* 2021;110:1253–68. <https://doi.org/10.1002/JLB.5MA0721-825R>
 42. Cai Y, Dai Y, Wang Y *et al*. Single-cell transcriptomics of blood reveals a natural killer cell subset depletion in tuberculosis. *EBioMedicine* 2020;53:102686. <https://doi.org/10.1016/j.ebiom.2020.102686>