

Immunopipe: a comprehensive and flexible scRNA-seq and scTCR-seq data analysis pipeline

Panwen Wang¹, Yue Yu², Haidong Dong³, Shuwen Zhang², Zhifu Sun², Hu Zeng²,
Patrizia Mondello⁴, Jean-Pierre A. Kocher², Junwen Wang^{1,5}, Yan W. Asmann⁶, Yi Lin⁴,
Ying Li^{6,*}

¹Department of Quantitative Health Sciences, Mayo Clinic, Scottsdale, AZ 85259, United States

²Department of Quantitative Health Sciences, Mayo Clinic, Rochester, MN 55905, United States

³Department of Urology and Immunology, Mayo Clinic, Rochester, MN 55902, United States

⁴Division of Hematology, Department of Internal Medicine, Mayo Clinic, Rochester, MN 55905, United States

⁵Division of Applied Oral Sciences & Community Dental Care, Faculty of Dentistry, The University of Hong Kong, Hong Kong SAR, China

⁶Department of Quantitative Health Sciences, Mayo Clinic, Jacksonville, FL 32224, United States

*To whom correspondence should be addressed. Email: li.ying@mayo.edu

Abstract

Single-cell sequencing technologies provide us with information at the level of individual cells. Combining single-cell RNA-seq and single-cell TCR-seq profiling enables the exploration of cell heterogeneity and T-cell receptor repertoires simultaneously. Integrating both types of data can play a crucial role in enhancing our understanding of T-cell-mediated immunity and, in turn, facilitate the advancement of immunotherapy. Here, we present immunopipe, a comprehensive and flexible pipeline to perform integrated analysis of scRNA-seq and scTCR-seq data. In addition to the command line tool, we provide a user-friendly web interface for pipeline configuration and execution monitoring, benefiting researchers without extensive programming experience. With its comprehensive functionality and ease of use, immunopipe empowers researchers to uncover valuable insights from scRNA-seq and scTCR-seq data, ultimately advancing the understanding of immune responses and immunotherapy development.

Introduction

T cells play an essential role in the adaptive immune system and are critical for the success of immunotherapy. The immune system safeguards our bodies from various illnesses, including cancer, by generating a diverse list of T-cell receptors (TCRs) through V(D)J recombination. These TCRs enable T cells to recognize and target specific antigens [1]. To study the diversity and clonality of the TCR repertoire at the single-cell level, single-cell TCR-sequencing (scTCR-seq) has emerged as a powerful technique [2]. Additionally, single-cell RNA sequencing (scRNA-seq) allows for the analysis of gene expression in individual cells, providing insights into cellular heterogeneity and functional states [3]. By integrating scRNA-seq and scTCR-seq data, researchers can gain a comprehensive understanding of T-cell-mediated immunity and its potential applications in immunotherapy [4, 5].

Recent advances in single-cell sequencing have enabled the simultaneous capture of transcriptomic profiles and TCR repertoires, which is crucial for understanding immune cell heterogeneity, clonal dynamics, and mechanisms underlying immunotherapy response and resistance. Although a few tools, including Dcirpy [6], Dandelion [7], Playtpus [8], scRepertoire [9], and immunarch (<https://immunarch.com/>), are available to analyze these data, immunopipe incorpo-

rates novel analytical steps—such as the physiochemical attribute analysis of CDR3 sequences and the incorporation of TESSA for phenotype-associated TCR repertoire analysis. The pipeline's comprehensive design supports both quality control and downstream analysis, including differential gene expression (DGE), gene set enrichment analysis (GSEA), and metabolic landscape analysis, thereby addressing a broader spectrum of analytical needs compared with existing implementations.

In this work, we have developed a flexible pipeline named immunopipe that facilitates the analysis of scRNA-seq and scTCR-seq data, empowering researchers to unlock valuable insights and accelerate the development of immunotherapeutic strategies. With its user-friendly web interface and comprehensive functionality, immunopipe is accessible to researchers with varying levels of programming experience.

Materials and methods

Immunopipe is implemented on top of pipen (<https://github.com/pwwang/pipen>), a flexible and extensible pipeline framework written in Python (<https://python.org>). Its architecture makes it easier to build complex workflows by breaking them

Received: December 17, 2024. Revised: March 2, 2025. Editorial Decision: May 4, 2025. Accepted: May 7, 2025

© The Author(s) 2025. Published by Oxford University Press on behalf of NAR Genomics and Bioinformatics.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License

(<https://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact reprints@oup.com for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact journals.permissions@oup.com.

down into smaller and reusable processes. The extensibility of pipen enables a variety of plug-ins that add to its already comprehensive capabilities. Immunopipe benefits from two particularly essential plug-ins, namely pipen-report (<https://github.com/pwwang/pipen-report>) which streamlines the process of organizing and generating reports for pipelines, and pipen-board (<https://github.com/pwwang/pipen-board>) which supplies a user-friendly web interface for configuring the pipeline, starting and monitoring the execution.

The implementation of each process is not restricted by programming languages, allowing for the incorporation of existing tools and packages. Immunopipe wraps both existing tools and in-house scripts to perform the analysis. Seurat (<https://satijalab.org/seurat>) is the core tool for the pipeline, not only for the common practices of scRNA-seq data analysis but also for data integration. The TCR clonal information is loaded into the Seurat object as metadata at the cell level and passed down to the downstream analysis. Enrichr [10] is used for enrichment analysis, and fgsea [11] is employed for GSEA. Automated cell type annotation is performed using scCATCH [12], sctype [13], celltypist [14], or hitype (<https://github.com/pwwang/hitype>). The scTCR-seq data are loaded into the pipeline by immunarch (<https://immunarch.com>), which performs basic analysis, including statistics on TCR clones, diversity metrics, gene usage, repertoire overlap, and clonotype tracking. Clustering by TCR sequences is conducted by GLANA [15] or clusTCR [16]. Metabolic landscape analysis is performed by the pipeline provided by Xiao *et al.* [17], which is modified to fit into immunopipe. The physicochemical attribute analysis of the CDR3 amino acid sequences is performed using TiRP [18]. TESSA [19] is incorporated to identify the phenotype-associated TCR repertoire. The subsequent analyses and visualizations are conducted using in-house scripts, which include T-cell selection, visualization of gene expression patterns, and exploration of cell distribution across different cell types.

We provide the documentation for immunopipe online as a manual and a web interface to generate the configuration file, where the description of the options is shown immediately upon focus. To illustrate the utilization of immunopipe in the analysis of scRNA-seq and scTCR-seq data, we provide an illustrative example employing a publicly available dataset [5] (see Data availability). The example includes a minimal configuration file with only the necessary configuration for running the pipeline, ideal for a quick start, and other configuration files with different options enabled, which can be used as a reference for users to adjust the parameters to suit their specific needs. It also contains the results and report generated by the pipeline, which can be viewed online (see Data availability). We have also compiled a gallery of repositories containing the configuration files for the pipeline and the results of the reanalysis of nine publicly available datasets (see Data availability).

Results

Abstraction of the analysis for scRNA-seq and scTCR-seq data

The abstraction of the analysis enhances the robustness and flexibility of the pipeline. Figure 1 illustrates the fundamental analyses implemented in immunopipe by showcasing representative sample figures for key steps, including the common

steps for scRNA-seq and scTCR-seq data (marked as essential processes in [Supplementary Figures S1 and S2](#)), respectively. T-cell selection is performed to avoid bias introduced by non-T cells. The selected T cells are then re-clustered and integrated with the clonal information. Various downstream analyses, including DGE and GSEA, are performed for the integrated data. In addition, the pipeline includes innovative approaches from recent methodological advancements in the literature.

Common practices for scRNA-seq data analysis

The scRNA-seq data analysis starts with quality control to assess the overall data quality and identify potential outliers or technical artifacts. Several metrics commonly used by the community [20] are included, such as the number of cells with expression for a gene and, for individual cells, the number of features and the proportion of mitochondrial, ribosomal, hemoglobin, and platelet genes [20]. Normalization, sample integration, dimension reduction, and clustering are conducted to group cells with similar expression profiles, which aids in identifying distinct cell types or populations. The markers of the clusters and the enriched pathways offer valuable insights into the functional states and potential roles of various cell subsets, thereby aiding in cell population identification. Additionally, the cell type annotation process can be automated or achieved using supervised clustering by mapping the cells to well-annotated reference datasets.

Common practices for scTCR-seq data analysis

The scTCR-seq data analysis starts with basic analysis and statistics of TCR clones, such as the number of unique TCRs, clonality, and diversity metrics. Advanced analysis, such as gene usage, clonotype change tracking, expanded or enriched TCR clone identification, and repertoire overlap detection between different samples or conditions, can be performed subsequently. The clone residency analysis is useful to examine the presence and persistence of T-cell clones across different samples or time points [5]. V-J usage plots investigate the V-J gene conjunction patterns within the TCR sequences, shedding light on the repertoire diversity and potential antigen specificity [21]. Moreover, clustering based on the TCR sequences allows identification of T-cell clusters or refinement of clonotypes, aiding in characterizing specific T-cell populations [15, 16].

T-cell selection

To integrate scTCR-seq data with scRNA-seq data and avoid the bias introduced by non-T cells, a T-cell selection process allows for the seamless integration of T-cell populations with scRNA-seq data. Cells are first clustered and T cells are selected based on the clonotype percentage in each cluster from scTCR-seq data and the expression of marker genes, including positive markers CD3E, CD3D, and CD3G [22], and negative markers, or exclusive markers for other cell types, such as CD14, CD19, and CD68 [23]. The selected T cells are then re-clustered and analyzed following the common practices for scRNA-seq data analysis.

Integration of scRNA-seq and scTCR-seq data

Immunopipe integrates scRNA-seq and scTCR-seq data through cell barcodes so that the TCR repertoire can be explored with different cell types annotated by scRNA-seq data. For example, we can compare the diversity of TCR

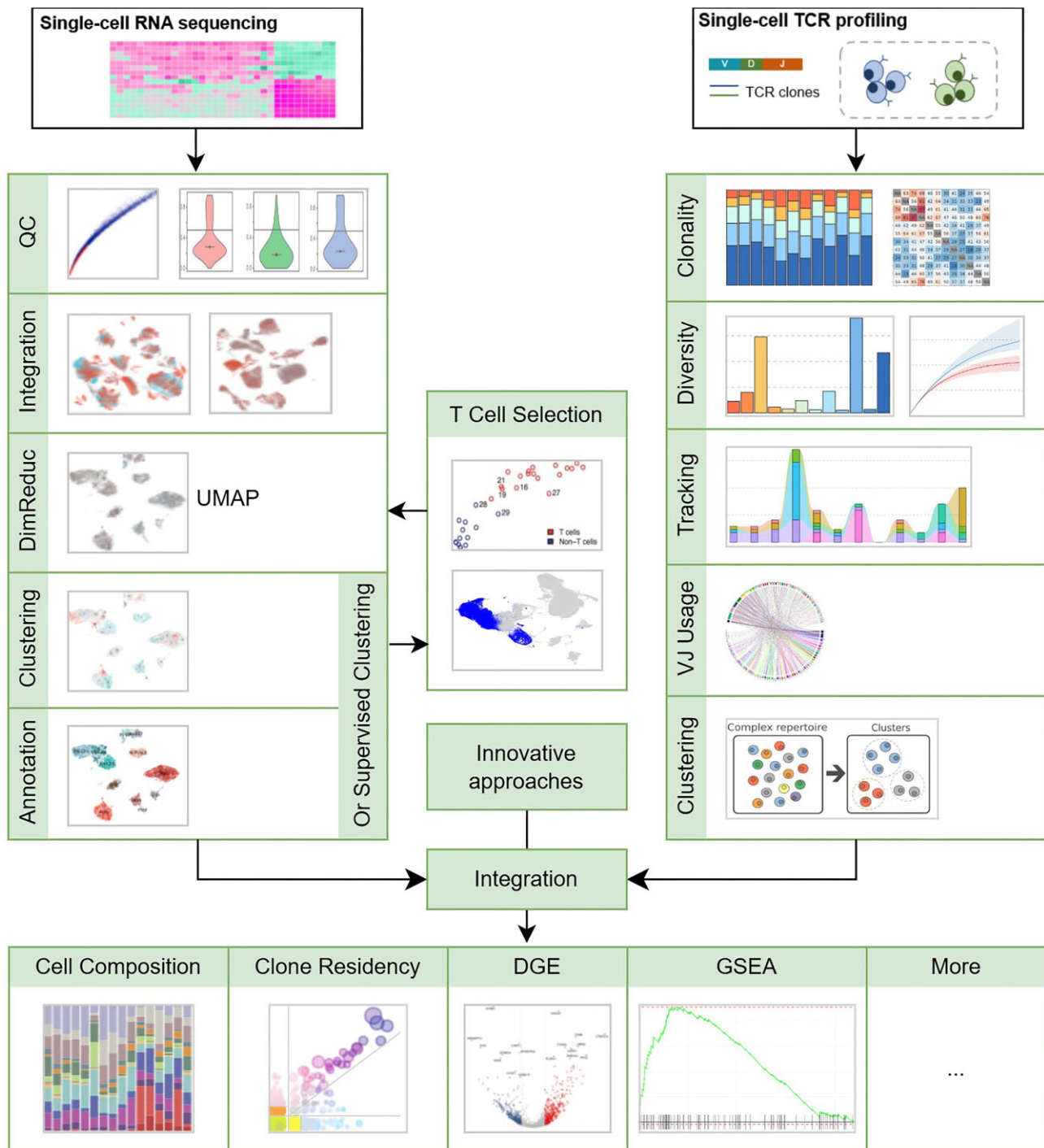


Figure 1. Illustration of analysis by immunopipe. The analysis initiates with quality control (QC) of scRNA-seq data, followed by comprehensive data integration. Dimensionality reduction techniques are subsequently applied, facilitating clustering and annotation or supervised clustering of the data. Concurrently, single-cell TCR profiling encompasses clonality analysis and diversity assessment. The tracking of TCR clones over time, VJ gene usage, and clustering of TCR repertoires further refine the analysis. The integration of T-cell selection data from both scRNA-seq and TCR profiling is centrally depicted, highlighting an innovative approach for comprehensive analysis. The outputs, illustrated in the bottom panel, include cell composition, clone residency, differential gene expression (DGE), gene set enrichment analysis (GSEA), and other analyses. This integrated methodology delivers an in-depth understanding of immune cell heterogeneity and dynamics.

Table 1. Publicly available datasets reanalyzed by immunopipe

GEO	Condition	No. of individuals	No. of samples	No. of cells	No. of matched TCR sequences	Reference
GSE144469	Colitis and therapy	22	22	75 569	68 760	[24]
GSE176201	COVID-19 lung tissue	5	6	34 781	23 081	[21]
GSE180268	HPV and head and neck cancer	6	19	53 303	26 844	[20]
GSE114724	Breast cancer	3	5	28 341	24 039	[25]
GSE161192	Lewy body dementia	2	4	6438	5642	[19]
GSE179994	Anti-PD-1 therapy in lung cancer	38	47	150 849	77 030	[4]
GSE148190	Skin cancer	1	2	8794	4904	[23]
GSE139555	Anti-PD1 therapy	14	32	194 519	67 700	[5]
GSE145370	Esophageal cancer	7	14	108 226	35 449	[22]

clones between different cell types or identify the TCR clones specific to a given cell type. In addition, it is possible to perform further analysis that involves clinical phenotypes, such as immunotherapy response. This analysis helps identify clones linked to specific clinical phenotypes, aiding the development of therapeutic strategies.

Differential gene expression analysis and gene set enrichment analysis

DGE analysis is a common practice in scRNA-seq data analysis. It allows for the identification of genes that are differentially expressed between two groups of cells, providing insights into the functional differences between these groups. GSEA is widely used to gain deeper insights into the functional states of cells and the underlying biological mechanisms. Using DGE analysis and GSEA, immunopipe enables researchers to understand the functional states of cells comprehensively.

Incorporation of innovative approaches

Additionally, immunopipe incorporates innovative methods that were introduced in recent publications. Metabolic landscape analysis at single-cell resolution [17] is designed to study metabolic programs in single cells. Physiochemical attribute analysis of the CDR3 amino acid sequences can be performed by comparing two groups of cells, typically regulatory T cells and resting conventional T cells. It quantifies hydrophobicity, isoelectric point, and volume of the amino acids in CDR3, which are associated with the self-reactivity of the TCR [18]. TESSA [19] characterizes the TCR repertoire by embedding TCR sequences and integrating them with scRNA-seq data, which has shown superiority over methods using TCR sequences only, in associating TCR clones with clinical phenotypes.

Flexible configuration and execution options

Immunopipe offers significant flexibility at various levels, from the overall workflow based on input data to the detailed configuration of specific analyses. Although it is primarily designed to analyze scRNA-seq and scTCR-seq data (as demonstrated in [Supplementary Figure S1](#)), immunopipe can be used for scRNA-seq data analysis alone (as shown in [Supplementary Figure S2](#)). The essential analyses that adhere to common practices are mandatory, while most other analyses are optional. Additionally, immunopipe offers granular control of each analysis, enabling users to adjust the parameters to align with their research design. Many parameters are pre-configured with carefully selected values, making the analysis process flexible and straightforward.

The analysis performed by immunopipe heavily relies on the cell grouping information within the data that could be retrieved from scTCR-seq and clinical data. Users can directly reference the relevant variable from the metadata in the configuration, aligning with their specific research requirements. It is also possible to derive new variables in the configuration from existing variables by modifying the metadata, which can subsequently serve as grouping information for future analysis, adding another layer of flexibility. Moreover, data filtration is allowed through configuration, which enables focusing on a subset of cells of interest.

In addition to running on a local machine, immunopipe is equipped with built-in executors, including Sun Grid Engine, Slurm Workload Manager, and SSH. Docker images were compiled (see Code availability) to enable seamless deployment across diverse computing environments. It also simplifies the installation procedure and mitigates the potential for dependency conflicts which can be a big hurdle for less experienced users.

The pipeline generates reports in HTML format, which can be easily delivered or hosted on a server for easy access and interpretation. Immunopipe is also accompanied by a web interface, which facilitates effortless configuration file generation, and initiation and monitoring of the execution. This interface streamlines the analysis process and provides convenient access to log files, intermediate results, and final reports. The web interface offers detailed descriptions of each configuration option, allowing users to make informed decisions and adjust the parameters to suit their needs.

Reanalysis of publicly available datasets

To demonstrate the capabilities of immunopipe, we have reanalyzed nine publicly available datasets [4, 5, 24–30] (see Data availability). The datasets are provided with both scRNA-seq and scTCR-seq data. However, it is noteworthy that the study by Luoma *et al.* [29] contains an additional dataset with scRNA-seq data only (CD45+ cells), which can also be analyzed by immunopipe. The datasets cover not only cancer but COVID-19 cases. The number of samples in each dataset ranges from 2 to 47, and of individuals from 1 to 38; the number of cells varies from 6438 to 194 519 and the number of matched TCR sequences from 5642 to 77 030 (Table 1). To demonstrate the scalability and efficiency of immunopipe, we benchmarked computational performance metrics, including runtime and memory usage, for each analytical step (see [Supplementary Table S1](#)).

The primary aim of the reanalysis is to demonstrate the capabilities of immunopipe, and provide a reference for researchers to adjust the parameters, visualize the correspond-

ing results, and facilitate the interpretation of the results. The reanalysis is not a reproduction of the analysis in the original publications, and the results may not be directly comparable. This is due to the lack of transparency in some publications regarding the pre-processing steps and the use of different parameters in the analysis. Moreover, the tools used in the original analyses may have been updated since the publication, which can also lead to differences in the results. The configuration files, results, and reports for the reanalysis are available in each repository (see Data availability).

Acknowledgements

We thank the Mayo Clinic Comprehensive Cancer Center Novel Biotherapeutics Data Science Workstream team and the Mayo Clinic Bioinformatics Core staff for their valuable feedback on the pipeline.

Author contributions: Panwen Wang and Ying Li: conceptualization, methodology, formal analysis, writing—original draft, and writing—review & editing; Haidong Dong, Hu Zeng, and Patrizia Mondello: investigation, data curation, methodology, visualization, and writing—review & editing; Yue Yu and Shuwen Zhang: software, validation, resources, and writing—review & editing; Zhifu Sun, Jean-Pierre A. Kocher, Junwen Wang, Yan W. Asmann, Yi Lin, and Ying Li: supervision, project administration, funding acquisition, and writing—review & editing.

Supplementary data

Supplementary data is available at NAR Genomics & Bioinformatics online.

Conflict of interest

None declared.

Funding

Mayo Clinic Comprehensive Cancer Center [P30CA15083]; Casey DeSantis Cancer Research Act Florida [Statute 381.922]; Mayo Clinic Center for Individualized Medicine, Bernard E. and Edith B. Waterman, and Henry J. Predolin Foundation.

Data availability

The details about the example can be found at <https://github.com/pwwang/immunopipe-example>. The data used in the example are hosted at GEO (GSE139555). The report for the example generated by the minimal configuration file can be viewed at <http://imp.pwwang.com/minimal/reports/> and the report for the configuration file with all processes enabled is available at <https://imp.pwwang.com/output/reports/>. The gallery of repositories for the reanalysis of the publicly available datasets can be found at <https://pwwang.github.io/immunopipe/latest/gallery/>.

Code availability

The most recent source code for immunopipe is publicly available on GitHub (<https://github.com/pwwang/immunopipe>) and on Figshare (10.6084/m9.figshare.28049426), and the

documentation can be found at <https://pwwang.github.io/immunopipe/>. The docker image for immunopipe is also available on Docker Hub (<https://hub.docker.com/r/justold/immunopipe>).

References

- Davis MM, Bjorkman PJ. T-cell antigen receptor genes and T-cell recognition. *Nature* 1988;334:395–402. <https://doi.org/10.1038/334395a0>
- Pai JA, Satpathy AT. High-throughput and single-cell T cell receptor sequencing technologies. *Nat Methods* 2021;18:881–92. <https://doi.org/10.1038/s41592-021-01201-8>
- Stubbington MJT, Lonnberg T, Proserpio V *et al.* T cell fate and clonality inference from single-cell transcriptomes. *Nat Methods* 2016;13:329–32. <https://doi.org/10.1038/nmeth.3800>
- Liu B, Hu X, Feng K *et al.* Temporal single-cell tracing reveals clonal revival and expansion of precursor exhausted T cells during anti-PD-1 therapy in lung cancer. *Nat Cancer* 2022;3:108–21. <https://doi.org/10.1038/s43018-021-00292-8>
- Wu TD, Madireddi S, de Almeida PE *et al.* Peripheral T cell expansion predicts tumour infiltration and clinical response. *Nature* 2020;579:274–8. <https://doi.org/10.1038/s41586-020-2056-8>
- 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/10.1093/bioinformatics/btaa611>
- Suo C, Polanski K, Dann E *et al.* Dandelion uses the single-cell adaptive immune receptor repertoire to explore lymphocyte developmental origins. *Nat Biotechnol* 2024;42:40–51. <https://doi.org/10.1038/s41587-023-01734-7>
- Yermanos A, Agraftotis A, Kuhn R *et al.* Platypus: an open-access software for integrating lymphocyte single-cell immune repertoires with transcriptomes. *NAR Genom Bioinform* 2021;3:lqab023. <https://doi.org/10.1093/nargab/lqab023>
- Borcherting N, Bormann NL, Kraus G. scRepertoire: an R-based toolkit for single-cell immune receptor analysis. *F1000Res* 2020;9:47. <https://doi.org/10.12688/f1000research.22139.1>
- Xie Z, Bailey A, Kuleshov MV *et al.* Gene set knowledge discovery with Enrichr. *Curr Protoc* 2021;1:e90. <https://doi.org/10.1002/cpz1.90>
- Korotkevich G, Sukhov V, Budin N *et al.* Fast gene set enrichment analysis. bioRxiv. <https://doi.org/10.1101/060012>, 1 February 2021, preprint: not peer reviewed.
- Shao X, Liao J, Lu X *et al.* scCATCH: automatic annotation on cell types of clusters from single-cell RNA sequencing data. *iScience* 2020;23:100882. <https://doi.org/10.1016/j.isci.2020.100882>
- Ianevski A, Giri AK, Aittokallio T. Fully-automated and ultra-fast cell-type identification using specific marker combinations from single-cell transcriptomic data. *Nat Commun* 2022;13:1246. <https://doi.org/10.1038/s41467-022-28803-w>
- Dominguez Conde C, Xu C, Jarvis LB *et al.* Cross-tissue immune cell analysis reveals tissue-specific features in humans. *Science* 2022;376:eabl5197. <https://doi.org/10.1126/science.abl5197>
- Zhang H, Zhan X, Li B. GIANA allows computationally-efficient TCR clustering and multi-disease repertoire classification by isometric transformation. *Nat Commun* 2021;12:4699. <https://doi.org/10.1038/s41467-021-25006-7>
- Valkiers S, Van Houcke M, Laukens K *et al.* ClusTCR: a Python interface for rapid clustering of large sets of CDR3 sequences with unknown antigen specificity. *Bioinformatics* 2021;37:4865–7. <https://doi.org/10.1093/bioinformatics/btab446>
- Xiao Z, Dai Z, Locasale JW. Metabolic landscape of the tumor microenvironment at single cell resolution. *Nat Commun* 2019;10:4865–7. <https://doi.org/10.1038/s41467-019-11738-0>
- Lagattuta KA, Kang JB, Nathan A *et al.* Repertoire analyses reveal T cell antigen receptor sequence features that influence T cell fate.

- Nat Immunol* 2022;23:446–57. <https://doi.org/10.1038/s41590-022-01129-x>
19. Zhang Z, Xiong D, Wang X *et al.* Mapping the functional landscape of T cell receptor repertoires by single-T cell transcriptomics. *Nat Methods* 2021;18:92–9. <https://doi.org/10.1038/s41592-020-01020-3>
 20. Ilicic T, Kim JK, Kolodziejczyk AA *et al.* Classification of low quality cells from single-cell RNA-seq data. *Genome Biol* 2016;17:29. <https://doi.org/10.1186/s13059-016-0888-1>
 21. Shugay M, Bagaev DV, Turchaninova MA *et al.* VDJtools: unifying post-analysis of T cell receptor repertoires. *PLoS Comput Biol* 2015;11:e1004503. <https://doi.org/10.1371/journal.pcbi.1004503>
 22. Dong D, Zheng L, Lin J *et al.* Structural basis of assembly of the human T cell receptor–CD3 complex. *Nature* 2019;573:546–52. <https://doi.org/10.1038/s41586-019-1537-0>
 23. Strobl H, Scheinecker C, Csmarits B *et al.* Flow cytometric analysis of intracellular CD68 molecule expression in normal and malignant haemopoiesis. *Br J Haematol* 1995;90:774–82. <https://doi.org/10.1111/j.1365-2141.1995.tb05195.x>
 24. Gate D, Tapp E, Leventhal O *et al.* CD4⁺ T cells contribute to neurodegeneration in Lewy body dementia. *Science* 2021;374:868–74. <https://doi.org/10.1126/science.abf7266>
 25. Eberhardt CS, Kissick HT, Patel MR *et al.* Functional HPV-specific PD-1⁺ stem-like CD8 T cells in head and neck cancer. *Nature* 2021;597:279–84. <https://doi.org/10.1038/s41586-021-03862-z>
 26. Cheon IS, Li C, Son YM *et al.* Immune signatures underlying post-acute COVID-19 lung sequelae. *Sci Immunol* 2021;6:eabk1741. <https://doi.org/10.1126/sciimmunol.abk1741>
 27. Zheng Y, Chen Z, Han Y *et al.* Immune suppressive landscape in the human esophageal squamous cell carcinoma microenvironment. *Nat Commun* 2020;11:6268. <https://doi.org/10.1038/s41467-020-20019-0>
 28. Mahuron KM, Moreau JM, Glasgow JE *et al.* Layilin augments integrin activation to promote antitumor immunity. *J Exp Med* 2020;217:e20192080. <https://doi.org/10.1084/jem.20192080>
 29. Luoma AM, Suo S, Williams HL *et al.* Molecular pathways of colon inflammation induced by cancer immunotherapy. *Cell* 2020;182:655–71. <https://doi.org/10.1016/j.cell.2020.06.001>
 30. Azizi E, Carr AJ, Plitas G *et al.* Single-cell map of diverse immune phenotypes in the breast tumor microenvironment. *Cell* 2018;174:1293–308. <https://doi.org/10.1016/j.cell.2018.05.060>