

# TcEVdb: a database for T-cell-derived small extracellular vesicles from single-cell transcriptomes

Tao Luo<sup>1,2</sup>, Wen-Kang Shen<sup>2,3</sup>, Chu-Yu Zhang<sup>2,3</sup>, Dan-Dan Song<sup>2</sup>, Xiu-Qing Zhang<sup>1</sup>, An-Yuan Guo<sup>2,\*</sup>, Qian Lei<sup>2,\*</sup>

<sup>1</sup>College of Life Sciences, University of Chinese Academy of Sciences, Beijing 100049, China

<sup>2</sup>Department of thoracic surgery, West China Biomedical Big Data Center, West China Hospital, Sichuan University, #37 Guoxue Alley, Wuhou District, Chengdu 610041, China

<sup>3</sup>Hubei Bioinformatics & Molecular Imaging Key Laboratory, College of Life Science and Technology, Huazhong University of Science and Technology, 1037 Luoyu Road, Hongshan District, Wuhan 430074, China

\*Corresponding authors. An-Yuan, Guo, West China Biomedical Big Data Center, West China Hospital, Sichuan University, Chengdu 610041, China.  
E-mail: [guoanyuan@wchscu.cn](mailto:guoanyuan@wchscu.cn); Qian Lei, E-mail: [leiqianqian@wchscu.cn](mailto:leiqianqian@wchscu.cn)

Citation details: Luo, T., Shen, W., Zhang, C. *et al.* TcEVdb: a database for T-cell-derived small extracellular vesicles from single-cell transcriptomes. *Database* (2025) Vol. 2025: article ID baaf012; DOI: <https://doi.org/10.1093/database/baaf012>

## Abstract

T-Cell-derived extracellular vesicles (TcEVs) play key roles in immune regulation and tumor microenvironment modulation. However, the heterogeneity of TcEV remains poorly understood due to technical limitations of EV analysis and the lack of comprehensive data. To address this, we constructed TcEVdb, a comprehensive database that explores the expression and cluster of TcEV by the SEVtras method from T-cell single-cell RNA sequencing data. TcEVdb contains 277 265 EV droplets from 51 T-cell types across 221 samples from 21 projects, covering 9 tissue sources and 23 disease conditions. The database provides two main functional modules. The Browse module enables users to investigate EV secretion activity indices across samples, visualize TcEV clusters, analyze differentially expressed genes (DEGs) and pathway enrichment in TcEV subpopulations, and compare TcEV transcriptomes with their cellular origins. The Search module allows users to query specific genes across all datasets and visualize their expression distribution. Furthermore, our analysis of TcEV in diffuse large B-cell lymphoma revealed increased EV secretion in CD4<sup>+</sup> T exhausted cells compared to healthy controls. Subsequent analyses identified distinct droplet clusters with differential expression genes, including clusters enriched for genes associated with cell motility and mitochondrial function. Overall, TcEVdb serves as a comprehensive resource for exploring the transcriptome of TcEV, which will contribute to advancements in EV-based diagnostics and therapeutics across a wide range of diseases.

Database URL: <https://guolab.wchscu.cn/TcEVdb>

## Introduction

Extracellular vesicles (EVs) are lipid bilayer-enclosed vesicles secreted by nearly all cell types, carrying a diverse array of biological molecules, including proteins, lipids, and nucleic acids [1]. They act as intercellular messengers and play crucial roles in intercellular communication and various pathological processes [2, 3]. Specifically, T-cell-derived EV (TcEV) have emerged as critical mediators in immune regulation, modulating immune responses, shaping the tumor microenvironment, and influencing the pathogenesis of autoimmune diseases [4–6]. However, the high heterogeneity of the EV population presents a significant challenge in elucidating their specific functions [3, 7].

In the EV research field, one of the most attractive and challenging questions is to understand the heterogeneity and functional diversity of EVs [8]. Traditional bulk analysis techniques often obscure this complexity, potentially masking important subpopulations with distinct functional characteristics. Recent technological advancements have made remarkable progress in single-vesicle-level analysis, employing techniques such as nanoflow cytometry [9], total internal

reflection fluorescence microscopy [10], microfluidic devices [11], and single-EV sequencing [12]. While these methods have provided valuable insights into EV heterogeneity, they fall short of enabling high-throughput analysis of EVs derived from a wide variety of cell types.

The advent of droplet-based single-cell RNA sequencing (scRNA-seq) has revolutionized the investigation of cellular heterogeneity at single-cell resolution. Recently, He *et al.* demonstrated that scRNA-seq data contain information about EVs, which are captured alongside individual cells in droplets. Based on this insight, they developed SEVtras, an innovative algorithm designed to identify EV-containing droplets in scRNA-seq data and estimate the EV secretion activity of individual cells [13]. SEVtras represents a significantly advancement in EV research by enabling the characterization and analyzing the expression of single or few EVs within scRNA-seq data. This approach facilitates high-throughput analysis of EV heterogeneity and secretion patterns, leveraging extensive existing scRNA-seq data without necessitating specialized EV isolation techniques.

Despite these advancements, a comprehensive resource for exploring TcEV across diverse tissues and pathological conditions remains unavailable. To address this gap, we present TcEVdb (a database for T-cell-derived small extracellular vesicles), a novel database that provides an unprecedented view of TcEV from scRNA-seq data. TcEVdb consists of 277,265 EV droplets from 21 projects containing 211 samples, encompassing 9 tissue sources and 23 different disease states and treatment conditions. By providing this comprehensive resource, TcEVdb will catalyze new discoveries in the field of TcEV and contribute to the advancement of EV-based diagnostics and therapeutics.

## Materials and Methods

### Data collection and preprocessing

We systematically collected T-cell scRNA-seq datasets from public repositories, including the National Center for Biotechnology Information Sequence Read Archive (SRA) (<http://www.ncbi.nlm.nih.gov/sra>) and the European Molecular Biology Laboratory - European Bioinformatics Institute BioStudies database [14], utilizing the keywords “T cell” and “scRNA-seq” to identify relevant datasets, and subsequently selecting only those with the organism specified as *Homo sapiens*. For each dataset, we documented relevant metadata, which included disease state, tissue source, sequencing platform information, and corresponding PubMed article references. Raw sequencing data in SRA format were downloaded using the *sra-toolkit* (<https://github.com/ncbi/sra-tools>) *prefetch* command and subsequently converted to FASTQ format with the *fastq-dump* command. Quality control of the sequencing reads was conducted using *fastp* (v0.20.1) with default parameters [15, 16].

The quality-controlled reads were aligned to the built-in human reference genome (GRCh38) using Cell Ranger (v6.0.1). The resulting gene expression matrixes underwent further analysis using the Scanpy package [17] in Python3 (<https://www.python.org/>). Quality control was performed at the cellular level, retaining cells with gene detection counts ranging from 300 to 6500, total counts between 500 and 40,000, and mitochondrial gene expression below 15%. Gene-level filtering retained genes detected in more than three cells. Mitochondrial and ribosomal genes were excluded from further analysis.

To isolate T cells from the scRNA-seq data, we retained cells expressing CD3D or CD3G, while removing cells marked by B-cell (CD79A), epithelial cell (EPCAM), or fibroblast (PDGFR4) markers. Potential doublets were filtered out using Scrublet [18].

### Cell type annotation

Cell type annotation was performed using a two-step approach that combined automated and manual annotation. Firstly, we employed CellTypist [19] for the initial automated annotation, utilizing five built-in models (Immune\_All\_Low, COVID19\_HumanChallenge\_Blood, Healthy\_COVID19\_PB MC, Cells\_Intestinal\_Tract, and Cells\_Lung\_Airway). These models primarily focused on annotating major CD4<sup>+</sup> T-cell subtypes, CD8<sup>+</sup> T-cell subtypes, nonconventional T cells, and other cell types. For each cell, the most frequently predicted cell type across the five models was assigned as the initial

classification, after which cells classified as “others” were excluded.

In the manual annotation step, we employed a marker-based annotation method based on the results of the automated annotation. To achieve a more granular annotation of T-cell subtypes and states, we collected a comprehensive list of marker genes from an extensive literature review (Supplementary Table S2). Leveraging the Scanpy [17] package, we implemented scRNA-seq data analysis steps. First, dimensionality reduction was performed using the Uniform Manifold Approximation and Projection (UMAP) algorithm (*sc.tl.umap*). Subsequently, unsupervised clustering was conducted using the Leiden algorithm (*sc.tl.leiden*) with a resolution set to 2.0. Marker genes for each cluster were identified using the *rank\_genes\_groups* function (*sc.tl.rank\_genes\_groups*) with default parameters. A cell cluster was assigned to a specific T-cell subtype or state if there was a substantial overlap (>50%) between its marker genes and the literature markers for that subtype. In addition to the cell type annotation conducted as part of TcEVdb, the dataset has also contributed to the development of TCellAtlas (<https://guolab.wchscu.cn/TCellAtlas>) and Single T Cell Annotation Tool (STCAT) (<https://github.com/GuoBioinfoLab/STCAT>). TCellAtlas serves as a reference for T-cell transcriptomic profiles, encompassing 68 subtypes and states derived from 1.6 million T cells, while STCAT is a specialized and automated annotation tool for T-cell subtype and state identification.

### SEVtras-based sEV droplet identification

To identify sEV-containing droplets in scRNA-seq datasets, we employed the SEVtras algorithm, a computational approach specifically designed for detecting EV signals at single-droplet resolution [13]. SEVtras utilizes an expectation-maximization framework, starting with a curated gene set enriched for EV markers, to iteratively compute droplet-specific EV signal scores. Through this process, cell-free droplets are classified as either sEV-containing or not based on their transcriptional profiles. The algorithm quantifies the enrichment of sEV marker genes in each droplet while effectively distinguishing these signals from background noise, such as cellular debris. By aggregating the results across samples, SEVtras generates a unified scoring system that facilitates cross-sample comparisons. This computational method allows for high-throughput identification of sEV populations from large-scale scRNA-seq datasets without requiring additional experimental steps. SEVtras has been validated on multiple datasets, demonstrating its robustness in capturing EV heterogeneity and secretion activity at the single-droplet level.

### EV droplet identification and EV secretion activity index calculation

We employed the SEVtras package to identify EV-containing droplets and calculate EV secretion activity across each dataset [13]. The *sEV\_recognizer* method was employed with default parameters to infer gene expression profiles of EV droplets. It is important to note that while SEVtras identifies droplets containing EVs, these droplets likely contain multiple EVs rather than a single vesicle, as EVs from similar cell types tend to form aggregates [13, 20]. In cases where datasets contained only one sample, we adhered to the SEVtras recommendation of duplicating the data to enhance

result reliability. Subsequently, we applied the `ESAI_calculator` method to calculate EV secretion activity for each sample and each cell type within the samples, utilizing the inferred EV droplet expression profiles alongside the annotated T-cell counts. Default parameters were maintained throughout this calculation.

## EV droplet data processing

Following the identification of EV-containing droplets using `SEVtrass`, projects that yielded fewer than 100 detected droplets were excluded from further analysis. Utilizing the `Scanpy`, gene filtering was applied to retain only those genes detected in more than three droplets. Data normalization was performed using the function `sc.pp.normalize_total` with a target sum of 100. Highly variable genes were identified using `sc.pp.highly_variable_genes`, selecting the top 1000 genes. To mitigate batch effects between samples, `sc.external.pp.bbknn` was employed. Dimensionality reduction was performed using `sc.tl.umap`, followed by unsupervised clustering with `sc.tl.leiden` at a resolution of 0.5. For each cluster, marker genes were identified using `sc.tl.rank_genes_groups` with the Wilcoxon test. Significant markers were defined as those exhibiting a log fold change  $>0.5$  and an adjusted *P*-value of  $<0.05$ . Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses were performed for the marker genes of each cluster using the `clusterProfiler` package [21], applying the Benjamini–Hochberg adjustment for multiple testing with a *q*-value cutoff of 0.05.

## Cell atlas–level integration

To integrate single-cell data from multiple projects, we employed `scVI` for batch correction [22, 23]. The standard `Scanpy` preprocessing pipeline was also applied, which included normalization, log transformation, and the selection of 2000 highly variable genes using the “`seurat_v3`” flavor. The `scVI` model was configured with the following parameters: `n_layers = 2`, `n_latent = 30`, and `gene_likelihood = “nb”`. After training the model, latent representations were extracted and utilized for neighbor graph construction, UMAP visualization, and Leiden clustering (resolution = 0.8).

## Database construction

The TcEVdb database web interface was developed using `Vue.js` (<https://vuejs.org/>) and `Element Plus` (<https://element-plus.org>) for the frontend, with a `Flask` (<https://flask.palletsprojects.com/en/3.0.x>) backend and `MongoDB` (<https://www.mongodb.com/>) for data storage. Interactive visualizations were implemented using `ECharts` (<https://echarts.apache.org>) and `Plotly.js` (<https://plotly.com/javascript>).

# Results

## Database overview and data summary

TcEVdb was designed as a user-friendly platform for researchers to explore EV populations across T cells in different tissues and diseases, facilitating the identification of novel EV populations and their potential functions in health and disease. By integrating extensive EV data with detailed metadata and intuitive visualization plots, TcEVdb aimed to accelerate research on TcEV (Fig. 1).

To comprehensively investigate TcEV, we performed EV droplet data processing and obtained 21 scRNA-seq datasets from 221 samples, covering 9 tissue sources and 23 disease states or therapeutic conditions. After analysis, we identified 277 265 EV droplets from 51 T-cell subtypes of these datasets.

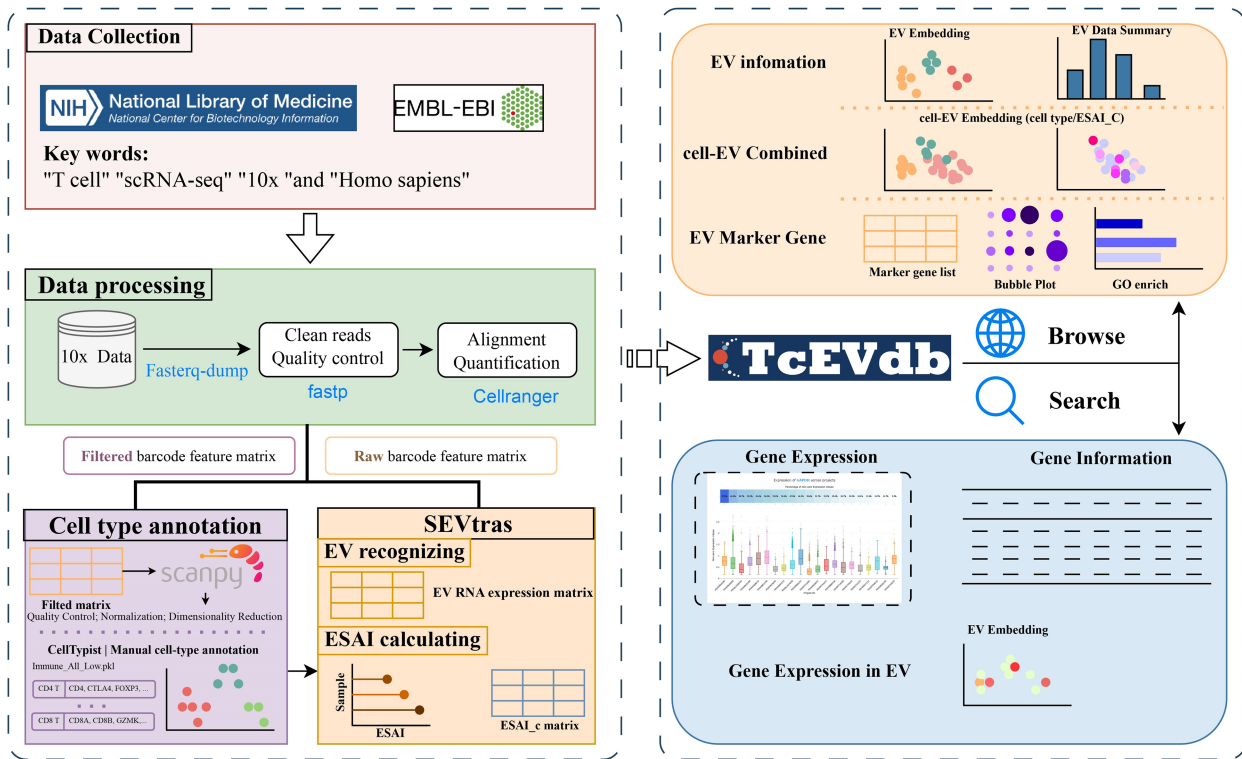
TcEVdb encompassed samples from nine tissue sources, with a predominance of blood samples. The analysis of TcEVdb samples showed that blood samples constituted the largest proportion at 52.5% ( $n = 116$ ), followed by colonic biopsies at 14.5% ( $n = 32$ ) and bone marrow at 12.7% ( $n = 28$ ) (Fig. 2a). The disease landscape presented in TcEVdb exhibited considerable diversity, encompassing samples from both healthy individuals and various pathological conditions (Fig. 2b). Notably, healthy samples constituted the largest subgroup (28.5%,  $n = 63$ ). Among the pathological conditions, Coronavirus Disease 2019 (12.2%,  $n = 27$ ), T-cell large granular lymphocytic leukemia (T-LGL leukemia) (10.0%,  $n = 22$ ), and melanoma (9.5%,  $n = 21$ ) were the most prevalent. To illustrate the data scale in TcEVdb, we presented the cell and EV counts for each project (Fig. 2c). The sizes of these projects varied widely, with the largest project containing 321 221 cells and 24 723 EV droplets (PRJNA714380 [24]), while the smallest project contained 9744 cells and 4369 EV droplets (PRJNA650256 [25]).

One of the key features of TcEVdb was the inclusion of EV secretory activity index data. Figure 2d illustrates the distribution of EV secretion activity index (ESAI) values across various tissue sources, demonstrating significant differences in EV secretory activity among these tissues. Notably, the highest median ESAI value was observed in lung-derived samples, followed closely by cerebrospinal fluid and lymph node samples. This finding suggested a potential tissue-specific regulation of EV secretion by T cells, which may be crucial for understanding the role of EVs in local immune responses and tissue homeostasis.

## Database functions and features

TcEVdb provided a user-friendly interface with two main functions: data browsing and gene querying, providing researchers to investigate the complex landscape of TcEV across biological conditions. The project overview interface (Fig. 3a) offered users comprehensive information about each project, displaying essential metadata such as project name, conditions, tissue sources, and sequencing platforms. The project information panel summarized key metrics, including cell and EV numbers, while the sample information table provided a detailed view of individual samples.

A key feature of TcEVdb was its visualization of ESAI (Fig. 3b). TcEVdb employed interactive heatmaps and bar charts to illustrate ESAI patterns at both the sample and cell type levels, providing valuable insights into the heterogeneity of EV production within a sample. The visualization capabilities of TcEVdb extended to EV droplet analyses through interactive UMAP plots, allowing users to explore EV distribution based on various metadata categories (Fig. 3c). Complementary statistics, such as cluster composition, sample distribution, and cell type origin, provided a multifaceted view of EV subpopulation heterogeneity. Furthermore, gene expression patterns could be visualized via violin plots, facilitating the identification of EV subpopulations with distinct molecular profiles. TcEVdb also provided detailed analyses for each cluster of EVs, including information on DEGs and



**Figure 1.** Overview of TcEVdb data collection, analysis workflow, and functionality.

their functional enrichment for each cluster (Fig. 3d). Users could query DEGs, GO terms, and KEGG pathway lists and images associated with each cluster. These functionalities facilitated the identification of potential functions of specific EV subpopulations.

Another notable attribute of TcEVdb was the integrated visualization of cells and EV droplets (Fig. 3e), facilitating a direct comparison between cellular transcriptomes and those of their secreted EV. The ESAI value for each cell was visually represented in the cell type-level ESAI (ESAI\_c) Embedding section. Additionally, the TcEVdb query function (Fig. 3f) enabled users to search for specific genes of interest across all datasets. Users could visualize the expression distribution of a given gene across different projects, identify datasets where that gene served as a cluster marker, and examine its expression pattern across datasets.

### Analysis of T-cell-derived EV from blood samples

TcEVdb contains EV from T cells in blood samples across a diverse range of disease states, demonstrating its potential to reveal disease-specific EV profiles. A comparative analysis of ESAI in blood samples from different cancer conditions showed significant differences (Fig. 4a). Notably, ESAI was significantly elevated in diffuse large B-cell lymphoma (DLBCL) compared to healthy controls ( $P \leq .05$ ). Further investigation into cell type-specific EV secretion patterns in DLBCL and healthy samples revealed different characteristics of various T-cell subsets (Fig. 4b). Among these, CD4<sup>+</sup> T exhausted cells (CD4 Tex) in DLBCL samples exhibited significantly high ESAI. To validate this, we performed a comparison of CD4 Tex-derived EV secretion between DLBCL and healthy samples (Fig. 4c). The analysis confirmed that

ESAI was significantly increased in DLBCL CD4 Tex cells ( $P \leq .0001$ ), suggesting a potential role of these cells in the pathogenesis or progression of DLBCL.

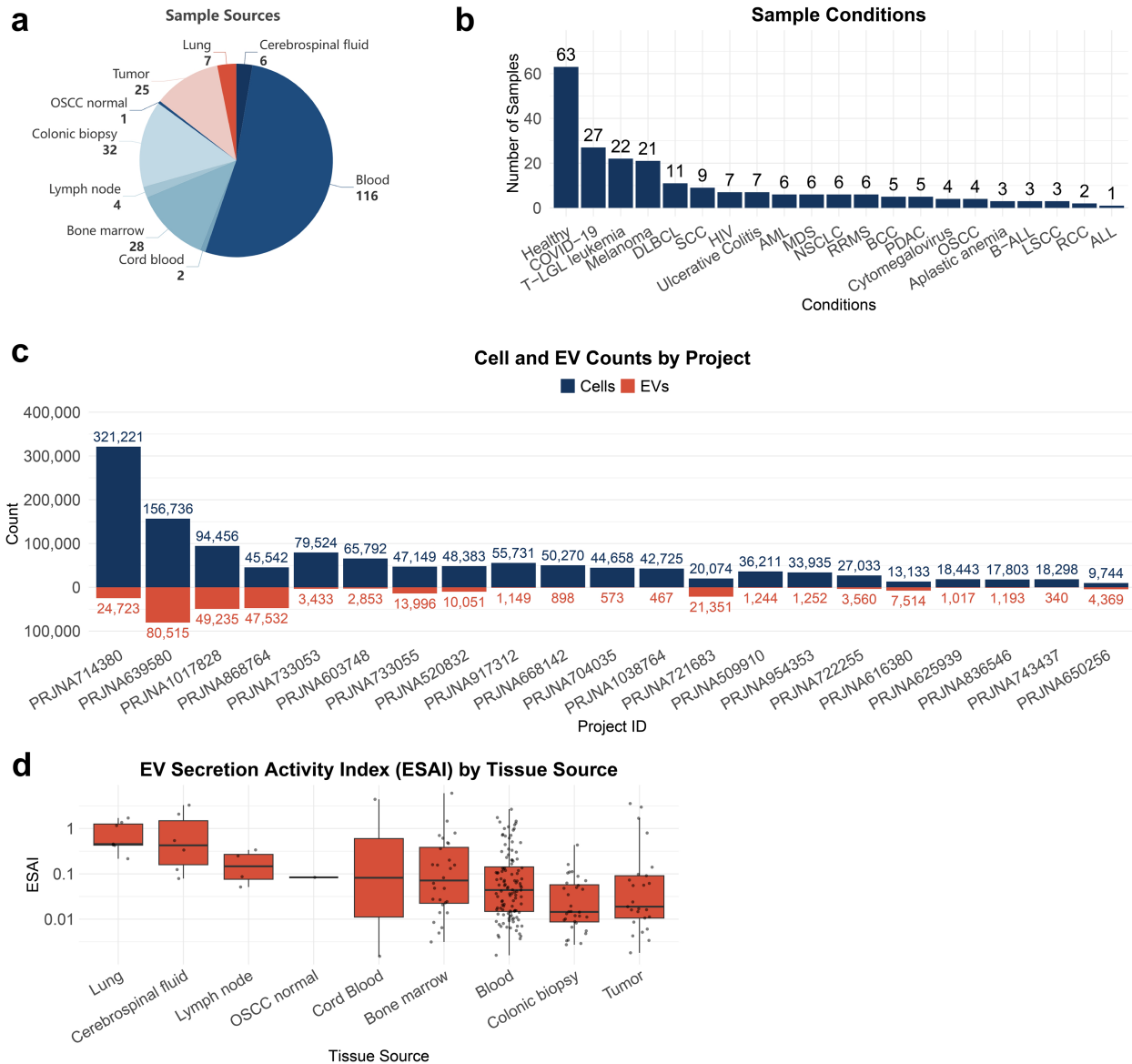
To elucidate the underlying cellular heterogeneity, we performed cell atlas-level integration of single-cell transcriptomic data of CD4 Tex cells from DLBCL and healthy samples using scVI (Methods: Cell atlas-level integration) (Fig. 4d and e). The integrated analysis encompassed 14,243 cells, with 12,719 cells from healthy samples and 1,524 cells from DLBCL samples. Although distinct disease-associated clusters were not immediately discernible, DLBCL-derived cells exhibited a tendency to cluster within Cluster 2 (Fig. 4d). This observation indicates potential transcriptional differences in CD4 Tex cells between DLBCL and healthy samples.

To further characterize DLBCL CD4 Tex cell-derived EVs, we conducted a downscaling and clustering analysis of the EV transcriptome data (Fig. 4f). The results revealed four distinct EV clusters, each characterized by unique gene expression features (Fig. 4g, Supplementary Table S3). Cluster 1 exhibited high expression levels of TMSB10 and MALAT1, genes associated with cell motility and metastasis [26, 27]. Cluster 2 demonstrated elevated expression of EEF1A1, B2M, and RPLP1. Cluster 3 was primarily defined by mitochondrial genes (MT-CO1, MT-CO2, MT-CO3, MT-ATP6, and MT-CYB), suggesting the presence of a distinct EV subpopulation enriched in mitochondrial cargos.

### Discussion

Recent studies have increasingly recognized the importance of T-cell-derived EVs (TcEV) in modulating immune responses and influencing disease progression [28, 29]. Despite the increasing acknowledgment of their significance,





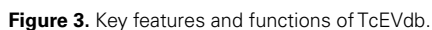
**Figure 2.** Sample distribution and ESAI across tissue sources of TcEVdb.

the investigation of TcEV has been hindered by the lack of comprehensive resources for analyzing their heterogeneity. To address this gap, we developed TcEVdb, a database that integrated large-scale single-cell RNA sequencing data processed using the SEVtrass algorithm, thereby providing unprecedented insights into TcEV. TcEVdb offered a comprehensive platform for the exploration of TcEV across various biological conditions.

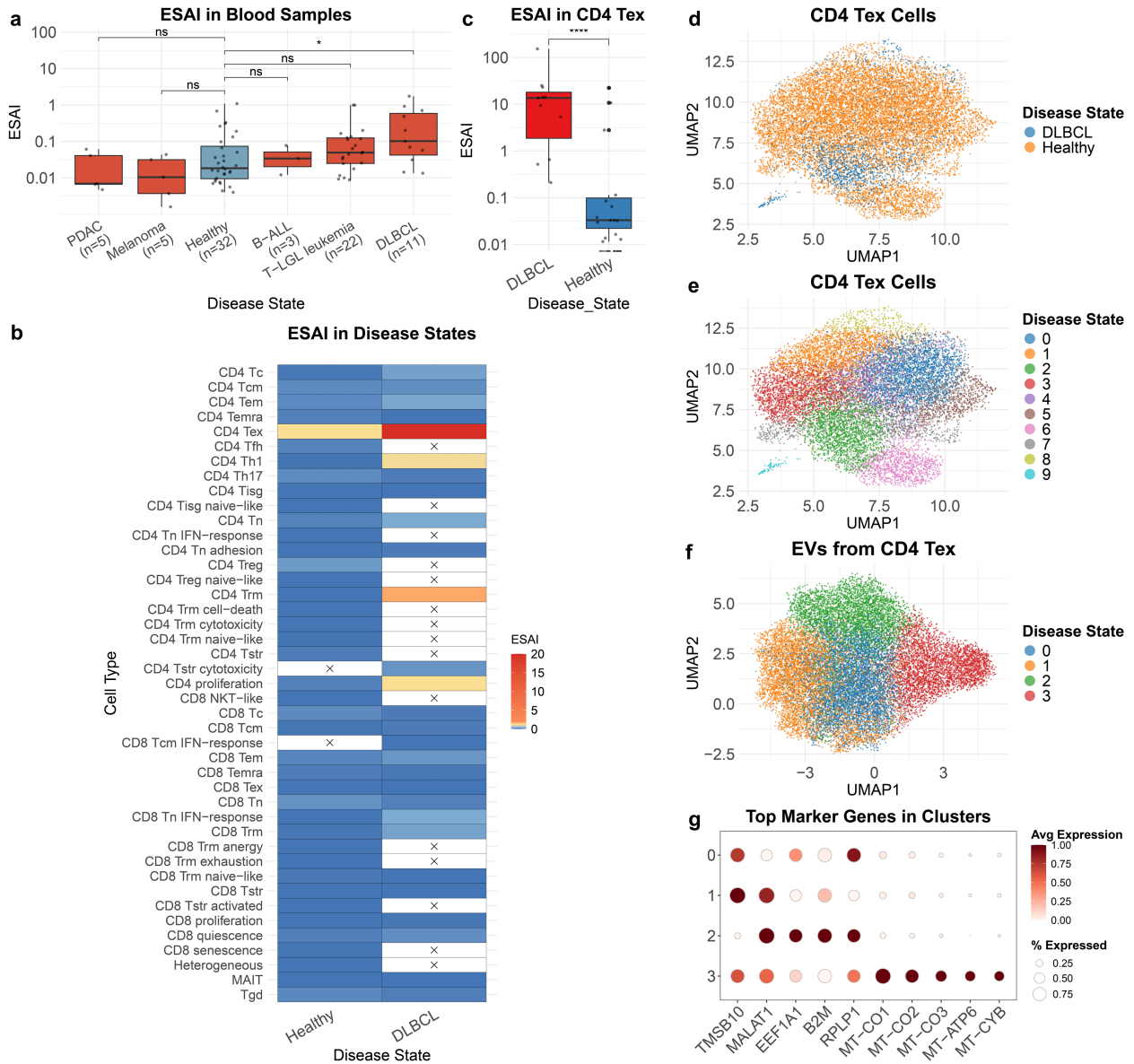
Several databases have been established to enhance research on EVs, but TcEVdb introduces a novel approach utilizing scRNA-seq data, which enables droplet resolution identification of sEV. TcEVdb focuses exclusively on T-cell-derived EVs, providing detailed characterization across 51 T-cell subtypes. Unlike Vesiclepedia 2024 and exoRBase 2.0 [30, 31], which aggregate data across diverse cell types without resolving cell-specific dynamics, TcEVdb's specificity facilitates the exploration of the heterogeneity and secretion activity of T-cell EVs at single-cell resolution. Additionally,

ExoBCD focuses primarily on biomarkers for breast cancer [32], and EVAtlas and EVmiRNA enrich EV research by concentrating on microRNA expression profiles and their functional implications [33, 34]. However, these resources are limited to bulk-level datasets. By employing scRNA-seq datasets, TcEVdb not only uncovers EV secretion patterns but also provides cell-specific biological insights, establishing it as a powerful resource for advancing our understanding of T-cell-derived EVs.

The development of TcEVdb involved a comprehensive process of data collection, processing, and analysis. T-cell scRNA-seq datasets were collected from public databases and underwent rigorous quality control and preprocessing. Detailed cell type annotation was accomplished through a combination of automated and manual methods. Subsequently, the SEVtrass algorithm was applied to identify EV-containing droplets and calculate EV secretion activity, yielding insights into EV heterogeneity and secretion patterns.



of detected genes per droplet compared to the recent single-EV study [12]. Additionally, the current version of TcEVdb focuses exclusively on TcEV, which restricts our understanding of the broader EV landscape. The biological significance of EVs extends beyond TcEV and immune regulation, with EVs from numerous cell types contributing to diverse physiological processes [35, 36]. Future iterations of the database could benefit from including EV-containing droplets derived from



**Figure 4.** Analysis of TcEV from blood samples.

other immune and nonimmune cell types, providing a more comprehensive view of intercellular communication networks in various physiological and pathological contexts.

The case study of TcEV from blood samples exemplified the potential of TcEVdb in generating novel biological insights. Our results revealed a disease-specific pattern of EV secretion, particularly in DLBCL. The observed increase in EV secretion from CD4 Tex cells in DLBCL, along with the characterization of various EV subpopulations, may provide a foundation for future studies investigating the role of TcEV in the pathogenesis of cancers [37].

In summary, TcEVdb was the first dedicated database for EVs derived from T cells, offering a platform for exploring the EV transcriptome across various biological conditions. TcEVdb included interactive visualization, gene expression profiling, and multigene search capabilities, enabling researchers to conduct comprehensive functional analyses of EV subpopulations. As the field of EV research continues to

evolve, TcEVdb is anticipated to advance EV biology research and contribute to the development of innovative EV-based diagnostics and therapeutics for a diverse array of diseases.

## Acknowledgements

We express our sincere gratitude to all researchers who have made their data publicly available, as this has formed the foundation of TcEVdb. We also thank the developers of SEV-tras for their innovative approach to analyzing EV-containing droplets. The computations in this paper were supported by the High Performance Computing platform at the West China Biomedical Big Data Center, West China Hospital, Sichuan University.

## Supplementary data

Supplementary data is available at *Database* online.

Conflict of interest: None declared.

## Funding

Financial support was provided by the National Natural Science Foundation of China (82271599 and 32370717) and the 1.3.5 project for disciplines of excellence from West China Hospital of Sichuan University (ZYYC23007).

## Data availability

All raw data used in this study were obtained from publicly available databases. The accession numbers used in this study are provided in [Supplementary Table S1](#). Additionally, all processed data, metadata, are freely accessible through TcEVdb (<https://guolab.wchscu.cn/TcEVdb>).

## References

- Zhang X, Zhang H, Gu J *et al*. Engineered extracellular vesicles for cancer therapy. *Adv Mater* 2021;33:e2005709. <https://doi.org/10.1002/adma.202005709>
- van Niel G, D'Angelo G, Raposo G. Shedding light on the cell biology of extracellular vesicles. *Nat Rev Mol Cell Biol* 2018;19:213–28. <https://doi.org/10.1038/nrm.2017.125>
- Kalluri R, LeBleu VS. The biology, function, and biomedical applications of exosomes. *Science* 2020;367:eaau6977.
- Buzas EI. The roles of extracellular vesicles in the immune system. *Nat Rev Immunol* 2023;23:236–50. <https://doi.org/10.1038/s41577-022-00763-8>
- Kalluri R. The biology and function of extracellular vesicles in immune response and immunity. *Immunity* 2024;57:1752–68. <https://doi.org/10.1016/j.immuni.2024.07.009>
- Yang P, Peng Y, Feng Y *et al*. Immune cell-derived extracellular vesicles – new strategies in cancer immunotherapy. *Front Immunol* 2021;12:771551. <https://doi.org/10.3389/fimmu.2021.771551>
- Lee Y-J, Chae S, Choi D. Monitoring of single extracellular vesicle heterogeneity in cancer progression and therapy. *Front Oncol* 2023;13:1256585. <https://doi.org/10.3389/fonc.2023.1256585>
- Shao H, Im H, Castro CM *et al*. New technologies for analysis of extracellular vesicles. *Chem Rev* 2018;118:1917–50. <https://doi.org/10.1021/acs.chemrev.7b00534>
- Liu H, Tian Y, Xue C *et al*. Analysis of extracellular vesicle DNA at the single-vesicle level by nano-flow cytometry. *J Extracell Vesicles* 2022;11:e12206. <https://doi.org/10.1002/jev2.12206>
- He D, Wang H, Ho S-L *et al*. Total internal reflection-based single-vesicle in situ quantitative and stoichiometric analysis of tumor-derived exosomal microRNAs for diagnosis and treatment monitoring. *Theranostics* 2019;9:4494–507. <https://doi.org/10.7150/thno.33683>
- Zhou J, Wu Z, Hu J *et al*. High-throughput single-EV liquid biopsy: rapid, simultaneous, and multiplexed detection of nucleic acids, proteins, and their combinations. *Sci Adv* 2020;6:eabc1204. <https://doi.org/10.1126/sciadv.abc1204>
- Luo T, Chen S-Y, Qiu Z-X *et al*. Transcriptomic features in a single extracellular vesicle via single-cell RNA sequencing. *Small Methods* 2022;6:e2200881. <https://doi.org/10.1002/smtd.202200881>
- He R, Zhu J, Ji P *et al*. SEVtrac delineates small extracellular vesicles at droplet resolution from single-cell transcriptomes. *Nat Methods* 2024;21:259–66. <https://doi.org/10.1038/s41592-023-02117-1>
- Sarkans U, Gostev M, Athar A *et al*. The BioStudies database-one stop shop for all data supporting a life sciences study. *Nucleic Acids Res* 2018;46:D1266–70. <https://doi.org/10.1093/nar/gkx965>
- Chen S, Zhou Y, Chen Y *et al*. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 2018;34:i884–90. <https://doi.org/10.1093/bioinformatics/bty560>
- Chen S. Ultrafast one-pass FASTQ data preprocessing, quality control, and deduplication using fastp. *Imeta* 2023;2:e107. <https://doi.org/10.1002/imt2.107>
- 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>
- Wolock SL, Lopez R, Klein AM. Scrublet: computational identification of cell doublets in single-cell transcriptomic data. *Cell Syst* 2019;8:281–91.e9. <https://doi.org/10.1016/j.cels.2018.11.005>
- Domínguez 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>
- Zomer A, Maynard C, Verweij FJ *et al*. In vivo imaging reveals extracellular vesicle-mediated phenocopying of metastatic behavior. *Cell* 2015;161:1046–57. <https://doi.org/10.1016/j.cell.2015.04.042>
- Xu S, Hu E, Cai Y *et al*. Using clusterProfiler to characterize multi-omics data. *Nat Protoc* 2024;19:3292–3320. <https://doi.org/10.1038/s41596-024-01020-z>
- Virshup I, Bredikhin D, Heumos L *et al*. The scverse project provides a computational ecosystem for single-cell omics data analysis. *Nat Biotechnol* 2023;41:604–06. <https://doi.org/10.1038/s41587-023-01733-8>
- Gayoso A, Lopez R, Xing G *et al*. A Python library for probabilistic analysis of single-cell omics data. *Nat Biotechnol* 2022;40:163–66. <https://doi.org/10.1038/s41587-021-01206-w>
- Gao S, Wu Z, Arnold B *et al*. Single-cell RNA sequencing coupled to TCR profiling of large granular lymphocyte leukemia T cells. *Nat Commun* 2022;13:1982. <https://doi.org/10.1038/s41467-022-29175-x>
- Chen J, Yang J, Li H *et al*. Single-cell transcriptomics reveal the intratumoral landscape of infiltrated T-cell subpopulations in oral squamous cell carcinoma. *Mol Oncol* 2021;15:866–86. <https://doi.org/10.1002/1878-0261.12910>
- Sribenja S, Li M, Wongkham S *et al*. Advances in thymosin beta10 research: differential expression, molecular mechanisms, and clinical implications in cancer and other conditions. *Cancer Invest* 2009;27:1016–22. <https://doi.org/10.3109/07357900902849640>
- Amodio N, Raimondi L, Juli G *et al*. MALAT1: a druggable long non-coding RNA for targeted anti-cancer approaches. *J Hematol Oncol* 2018;11:63. <https://doi.org/10.1186/s13045-018-0606-4>
- Hansen AS, Jensen LS, Gammelgaard KR *et al*. T-Cell derived extracellular vesicles prime macrophages for improved STING based cancer immunotherapy. *J Extracell Vesicles* 2023;12:e12350. <https://doi.org/10.1002/jev2.12350>
- Dang G, Li T, Yang D *et al*. T lymphocyte-derived extracellular vesicles aggravate abdominal aortic aneurysm by promoting macrophage lipid peroxidation and migration via pyruvate kinase muscle isozyme 2. *Redox Biol* 2022;50:102257. <https://doi.org/10.1016/j.redox.2022.102257>
- Chitti SV, Gummadi S, Kang T *et al*. Vesiclepedia 2024: an extracellular vesicles and extracellular particles repository. *Nucleic Acids Res* 2024;52:D1694–98. <https://doi.org/10.1093/nar/gkad1007>
- Lai H, Li Y, Zhang H *et al*. exoRBase 2.0: an atlas of mRNA, lncRNA and circRNA in extracellular vesicles from human biofluids. *Nucleic Acids Res* 2022;50:D118–28. <https://doi.org/10.1093/nar/gkab1085>
- Wang X, Chai Z, Pan G *et al*. ExoBCD: a comprehensive database for exosomal biomarker discovery in breast cancer. *Brief Bioinform* 2021;22:bbaa088. <https://doi.org/10.1093/bib/bbaa088>
- Liu C-J, Xie G-Y, Miao Y-R *et al*. EVAtlas: a comprehensive database for ncRNA expression in human extracellular vesicles. *Nucleic Acids Res* 2022;50:D111–17. <https://doi.org/10.1093/nar/gkab668>



34. Liu T, Zhang Q, Zhang J *et al.* EVmiRNA: a database of miRNA profiling in extracellular vesicles. *Nucleic Acids Res* 2019;**47**:D89–93. <https://doi.org/10.1093/nar/gky985>
35. Lei Q, Gao F, Liu T *et al.* Extracellular vesicles deposit PCNA to rejuvenate aged bone marrow-derived mesenchymal stem cells and slow age-related degeneration. *Sci Transl Med* 2021;**13**:eaz8697. <https://doi.org/10.1126/scitranslmed.aaz8697>
36. Hu M, Kenific CM, Boudreau N *et al.* Tumor-derived nanoseeds condition the soil for metastatic organotropism. *Semin Cancer Biol* 2023;**93**:70–82. <https://doi.org/10.1016/j.semcancer.2023.05.003>
37. Zhou Q, Wei S, Wang H *et al.* T cell-derived exosomes in tumor immune modulation and immunotherapy. *Front Immunol* 2023;**14**:1130033. <https://doi.org/10.3389/fimmu.2023.1130033>