frontiers | Research Topics

Single-cell OMICs analyses in cardiovascular diseases

Edited by
Hanjoong Jo, Abhijeet R. Sonawane
and Michel Puceat

Published in
Frontiers in Cardiovascular Medicine
Frontiers in Cell and Developmental Biology

ISSN 1664-8714
ISBN 978-2-8325-4900-1
DOI 10.3389/978-2-8325-4900-1

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Single-cell OMICs analyses in cardiovascular diseases

Topic editors

Hanjoong Jo — Emory University, United States

Abhijeet R. Sonawane — Brigham and Women's Hospital, Harvard Medical School, United States

Michel Puceat — Institut National de la Santé et de la Recherche Médicale (INSERM), France

Citation

Jo, H., Sonawane, A. R., Puceat, M., eds. (2024). *Single-cell OMICs analyses in cardiovascular diseases*. Lausanne: Frontiers Media SA.

doi: 10.3389/978-2-8325-4900-1

Table of contents

05 Editorial: Single-cell OMICs analyses in cardiovascular diseases
Abhijeet Rajendra Sonawane, Michel Pucéat and Hanjoong Jo

08 Next-Generation and Single-Cell Sequencing Approaches to Study Atherosclerosis and Vascular Inflammation Pathophysiology: A Systematic Review
Liam W. McQueen, Shameem S. Ladak, Riccardo Abbasciano, Sarah J. George, M-Saadeh Suleiman, Gianni D. Angelini, Gavin J. Murphy and Mustafa Zakkar

27 Identification of hub biomarkers of myocardial infarction by single-cell sequencing, bioinformatics, and machine learning
Qunhui Zhang, Yang Guo, Benyin Zhang, Hairui Liu, Yanfeng Peng, Di Wang and Dejun Zhang

41 Advances in application of single-cell RNA sequencing in cardiovascular research
Yue Hu, Ying Zhang, Yutong Liu, Yan Gao, Tiantian San, Xiaoying Li, Sensen Song, Binglong Yan and Zhuo Zhao

55 Single-cell transcriptome *in silico* analysis reveals conserved regulatory programs in macrophages/monocytes of abdominal aortic aneurysm from multiple mouse models and human
Shiyong Wu, Shibiao Liu, Baoheng Wang, Meng Li, Chao Cheng, Hairong Zhang, Ningheng Chen and Xueli Guo

68 Immune heterogeneity in cardiovascular diseases from a single-cell perspective
Xin Su, Li Wang, Ning Ma, Xinyu Yang, Can Liu, Fan Yang, Jun Li, Xin Yi and Yanwei Xing

84 Integrative single-cell analysis of cardiac and pulmonary sarcoidosis using publicly available cardiac and bronchoalveolar lavage fluid sequencing datasets
Abdel Daoud, Diego A. Lema, Taejoon Won and Daniela Čiháková

102 Single-Cell RNA sequencing investigation of female-male differences under PAD conditions
Gloriani Sánchez Marrero, Nicolas Villa-Roel, Feifei Li, Christian Park, Dong-Won Kang, Katherine E. Hekman, Hanjoong Jo and Luke P. Brewster

111 Exploring key genes associated with neutrophil function and neutrophil extracellular traps in heart failure: a comprehensive analysis of single-cell and bulk sequencing data
Xudong Li, Changhao Xu, Qiaoqiao Li, Qingxiang Shen and Long Zeng

123 Application and challenges of TCR and BCR sequencing to investigate T- and B-cell clonality in elastase-induced experimental murine abdominal aortic aneurysm
Christin Elster, Miriam Ommer-Bläsius, Alexander Lang, Tanja Vajen, Susanne Pfeiler, Milena Feige, Tin Yau Pang, Marius Böttenberg, Sarah Verheyen, Khang Lê Quý, Maria Chernigovskaya, Malte Kelm, Holger Winkels, Susanne V. Schmidt, Victor Greiff and Norbert Gerdes

138 Single-cell RNA sequencing and ATAC sequencing identify novel biomarkers for bicuspid aortic valve-associated thoracic aortic aneurysm
Xu-Wen Liu, Pei Wang, Li Zhang, Yu Zhu, Jun-Yu Zhai, Chang-Nan Wang, Jun Li and Jian Xiao

frontiers | Frontiers in Cardiovascular Medicine

TYPE Editorial
PUBLISHED 06 May 2024
DOI 10.3389/fcvm.2024.1413184

OPEN ACCESS

EDITED AND REVIEWED BY
Maximillian A. Rogers,
Intellia Therapeutics, Inc., United States

*CORRESPONDENCE
Abhijeet Rajendra Sonawane
asonawane@bwh.harvard.edu

RECEIVED 06 April 2024
ACCEPTED 23 April 2024
PUBLISHED 06 May 2024

CITATION
Sonawane AR, Pucéat M and Jo H (2024) Editorial: Single-cell OMICs analyses in cardiovascular diseases.
Front. Cardiovasc. Med. 11:1413184.
doi: 10.3389/fcvm.2024.1413184

Editorial: Single-cell OMICs analyses in cardiovascular diseases

Abhijeet Rajendra Sonawane[1]*, Michel Pucéat[2] and Hanjoong Jo[3]

[1]Center for Interdisciplinary Cardiovascular Sciences and Center for Excellences in Vascular Biology, Division of Cardiovascular Medicine, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, United States, [2]INSERM, Cardiovascular and Nutrition Center (C2VN), Aix-Marseille University, Marseille, France, [3]Wallace H. Coulter Department of Biomedical Engineering, Emory University and Georgia Institute of Technology, Atlanta, GA, United States

KEYWORDS

cardiovascular diseases, single-cell omics, data analysis, atherosclerosis, heart valve disease (HVD)

Editorial on the Research Topic
Single-cell OMICs analyses in cardiovascular diseases

Single-cell technologies have revolutionized the understanding of biological and pathological processes. They are now the driving force of cellular profiling across cardiovascular diseases (CVD) (1–5). They have enabled analysis of cellular composition with tissues and cell cultures, revealing the subtle nuances of heterogeneity among cell types and uncovering the dynamic changes in communication and signaling patterns triggered by disease.

Various omics modalities data can be generated now at single-cell resolution including joint and concomitant assays such as single cell RNA-sequencing (scRNA-seq) and single nucleus RNA-seq (snRNA-seq) for transcriptomics profiling, single cell assay for transposase-accessible chromatin with sequencing (scATAC-seq) for cell type specific profiling of epigenetic, and in turn, regulatory landscape, and cellular indexing of transcriptomes and epitopes-seq (CITE-seq) for profiling the surface proteins on different cell types. Recent advances, including spatial transcriptomics and single cell proteomics (6, 7) have allowed us to spatially resolve cell subtypes in active areas of disease tissues. These multiomics data types allow us to develop novel computational and bioinformatics methods to obtain a multidimensional view of the disease to uncover mechanisms, explain etiologies, and discover biomarkers. Single-cell data analysis tools such as Seurat (8, 9), and SingleCellExperiment (10) have democratized the capability of single cell analysis by beckoning researchers of all backgrounds to embark on a journey of exploration, regardless of their level of bioinformatics expertise. Applications such as SCHNAPPs have also enabled bench scientists to perform basic single cell analyses (11).

Understanding the impact of single cell research on CVD, including vascular and heart valve diseases, is crucial as it allows us to evaluate advancements in the field. Specific cells such as immune cells as well as cell-cell communication play a significant role in disease pathogenesis. Identifying unique cell types (cell subpopulations) and their behavior in the diseased tissue can elucidate their role in altering tissue structure and function.

In this special research topic, we intended to create a forum for current advances in single-cell techniques applied to cardiovascular disease areas. We invited articles on single cell technique and methods development, probing cellular heterogeneity, modern laboratory and in silico techniques, and its application to various vascular and valvular diseases. We received 18 submissions, out of which we selected 10 high quality peer reviewed articles for our collections. Below, we provide a brief overview of articles in this collection.

The systematic review conducted according to PRISMA standards from McQueen et al. provides us with extensive overview of the current state of single-cell research in cardiovascular medicine. Using proper exclusion and inclusion criterion, addressing the risk of bias analyses using tools, they provide a narrative synthesis of 34 articles (out of 791). They provide a comprehensive review of studies which include various single cell omics modalities such as scRNA-seq, scATAC-seq, CITE-seq, or combination thereof. They also include cell-type-specific discussions of the progress done in the research in cardiovascular diseases including, smooth muscle cells, macrophages, endothelial cells, and lymphocytes. Results from tissue and blood-specific analyses which included atherosclerotic lesions, cardiac and adventitial tissues, and blood, are also summarized to provide the reader important processes involved in atherosclerosis development and progression. Hu et al. review recent advances in scRNA-seq technology along with comparison of various gene amplification methods. They also discuss some of the standard workflow associated with single cell studies applied in CVD medicine. Su et al. present a focused review on role of diversity and abundance of immune cells such as macrophages, dendritic cells, and T cells, in cardiac homeostasis and effect of their infiltration on the diseased areas. They focus on immune heterogeneity in atherosclerosis, myocardial ischemia, and heart failure along with suggesting potentially new marker genes.

Li et al. continue the discussion on heart failure using the role of, relatively little explored, neutrophile extracellular traps (NET). Using conventional bioinformatic analysis on bulk and single cell RNA-seq datasets, they identify differentially expressed NET-related genes followed by neutrophile cell heterogeneity in heart failure and normal cardiac tissues followed by cellular differentiation and communication analyses identifying biomarkers associated with NETs in heart failure. Zhang et al. study myocardial infarction (MI) using public datasets and standard bioinformatics pipeline to identify IL1B and TLR2 as differentially expressed genes (DEGs) with most neighbors in protein-protein interaction network (PPI) in MI. While their findings provide valuable insights, there is a need for cautious interpretation due to the limitations inherent in the utilization of public datasets. Additionally, several seminal scRNA-seq studies highlighted the importance of sexual differences at the cell level in CVD. Marrero et al. study on sexual differences in peripheral artery disease (PAD) etiology utilizing scRNA-seq data. Mizrak et al. (12) using spatial transcriptomics uncovered male-specific smooth muscle cells subpopulation playing a key role in human thoracic aortic aneurysm. ScRNA-seq data from Shin et al. (13) allow to delineate the sex-difference in endothelial cells characteristics and function providing new clues about atherosclerotic diseases. These studies show that inclusion of biological sex in a proper experimental design (14) is of utmost importance to consider representativeness of the sample population and potential confounding variables.

Autoimmune conditions like cardiac and pulmonary sarcoidosis (CS/PS) are understudied yet potentially impactful on cardiovascular health. Daoud et al. conducted a systematic review using single-cell RNA-seq datasets to investigate PS and CS. They found increased immune cells and stromal populations in sarcoidosis tissues compared to controls, with sarcoidosis T cells and macrophages showing attenuated activation profiles. Abdominal aortic aneurysm (AAA) involves complex immune cell interactions and may exhibit autoimmune characteristics. Elster et al. studied the clonal expansion of T cells and B cells in AAA tissue using sc-RNA T cell receptor (TCR) and B cell receptor (BCR) sequencing using porcine pancreatic elastase mouse model. Wu et al. investigated macrophage regulation in AAA using scRNA-seq datasets from mouse models and humans. They identified IL-1B and THBS1 as co-upregulated genes across datasets, highlighting the complex immune involvement in AAA. Similarly, thoracic arch aneurysm (TAA), linked to bicuspid aortic valves (BAV), was investigated by Liu et al., revealing potential therapeutic genes. These studies enhance our grasp of aortic aneurysm pathophysiology, urging further investigation for improved cardiovascular disease management.

To summarize, the emergence of single-cell technologies has completely transformed our comprehension of cellular dynamics within cardiovascular disorders, notably in dissecting the intricate cellular compositions, diversities, and signaling modifications associated with diseases such as heart valve diseases and atherosclerosis. Through thorough assessments and leading-edge investigations, researchers have elucidated the presence and functions of diverse vascular and immune cell populations, indicated novel biomarkers, and underlined the molecular pathways influencing disease progression. These studies now complemented with spatial transcriptomics indicate the need of development of sophisticated bioinformatics tools for the integration of multiomics methodologies, which are needed to unravel the intricate complexities of cardiovascular pathophysiology, that will open avenues for targeted treatments and enhanced patient outcomes.

Author contributions

ARS: Writing – original draft, Writing – review & editing. MP: Writing – review & editing. HJ: Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article.

ARS acknowledges funding support from American Heart Association (23CDA1052394/Sonawane AR/2023). HJ's work was supported by funding from National Institutes of Health grants

HL119798, HL139757, and HL151358 and a Wallace H. Coulter Distinguished Faculty Chair Fund.

Acknowledgments

The editors would like to thank all the authors and expert reviewers who have participated in the preparation and evaluation of manuscripts presented in this Research Topic.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

The handling editor MR declared a past co-authorship with the author ARS.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Iqbal F, Schlotter F, Becker-Greene D, Lupieri A, Goettsch C, Hutcheson JD, et al. Sortilin enhances fibrosis and calcification in aortic valve disease by inducing interstitial cell heterogeneity. *Eur Heart J*. (2023) 44(10):885–898. doi: 10.1093/eurheartj/ehac818

2. Sonawane AR, Aikawa E, Aikawa M. Connections for matters of the heart: network medicine in cardiovascular diseases. *Front Cardiovasc Med*. (2022) 9:873582. doi: 10.3389/fcvm.2022.873582

3. Kumar S, Andueza A, Kim J, Kang DW, Jo H. Isolation of endothelial cells from the lumen of mouse carotid arteries for single-cell multi-omics experiments. *J Vis Exp*. (2021) (176). doi: 10.3791/63128

4. Andueza A, Kumar S, Kim J, Kang DW, Mumme HL, Perez JI, et al. Endothelial reprogramming by disturbed flow revealed by single-cell RNA and chromatin accessibility study. *Cell Rep*. (2020) 33:108491. doi: 10.1016/j.celrep.2020.108491

5. Kalluri AS, Vellarikkal SK, Edelman ER, Nguyen L, Subramanian A, Ellinor PT, et al. Single-cell analysis of the normal mouse aorta reveals functionally distinct endothelial cell populations. *Circulation*. (2019) 140:147–63. doi: 10.1161/CIRCULATIONAHA.118.038362

6. Tian L, Chen F, Macosko EZ. The expanding vistas of spatial transcriptomics. *Nat Biotechnol*. (2023) 41:773–82. doi: 10.1038/s41587-022-01448-2

7. Gatto L, Aebersold R, Cox J, Demichev V, Derks J, Emmott E, et al. Initial recommendations for performing, benchmarking and reporting single-cell proteomics experiments. *Nat Methods*. (2023) 20:375–86. doi: 10.1038/s41592-023-01785-3

8. Stuart T, Satija R. Integrative single-cell analysis. *Nat Rev Genet*. (2019) 20:257–72. doi: 10.1038/s41576-019-0093-7

9. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol*. (2018) 36:411–20. doi: 10.1038/nbt.4096

10. Amezquita RA, Lun ATL, Becht E, Carey VJ, Carpp LN, Geistlinger L, et al. Orchestrating single-cell analysis with bioconductor. *Nat Methods*. (2020) 17:137–45. doi: 10.1038/s41592-019-0654-x

11. Jagla B, Libri V, Chica C, Rouilly V, Mella S, Puceat M, et al. SCHNAPPs - single cell sHiNy APPlication(s). *J Immunol Methods*. (2021) 499:113176. doi: 10.1016/j.jim.2021.113176

12. Mizrak D, Zhao Y, Feng H, Macaulay J, Tang Y, Sultan Z, et al. Single-molecule spatial transcriptomics of human thoracic aortic aneurysms uncovers calcification-related CARTPT-expressing smooth muscle cells. *Arterioscler Thromb Vasc Biol*. (2023) 43:2285–97. doi: 10.1161/ATVBAHA.123.319329

13. Shin J, Hong J, Edwards-Glenn J, Krukovets I, Tkachenko S, Adelus ML, et al. Unraveling the role of sex in endothelial cell dysfunction: evidence from lineage tracing mice and cultured cells. *Arterioscler Thromb Vasc Biol*. (2024) 44:238–53. doi: 10.1161/ATVBAHA.123.319833

14. Pape M, Miyagi M, Ritz SA, Boulicault M, Richardson SS, Maney DL. Sex contextualism in laboratory research: enhancing rigor and precision in the study of sex-related variables. *Cell*. (2024) 187:1316–26. doi: 10.1016/j.cell.2024.02.008

frontiers in Cardiovascular Medicine

SYSTEMATIC REVIEW
published: 28 March 2022
doi: 10.3389/fcvm.2022.849675

Next-Generation and Single-Cell Sequencing Approaches to Study Atherosclerosis and Vascular Inflammation Pathophysiology: A Systematic Review

Liam W. McQueen[1], Shameem S. Ladak[1], Riccardo Abbasciano[1], Sarah J. George[2], M-Saadeh Suleiman[2], Gianni D. Angelini[2], Gavin J. Murphy[1] and Mustafa Zakkar[1]*

[1] Department of Cardiovascular Sciences, Clinical Science Wing, Glenfield Hospital, University of Leicester, Leicester, United Kingdom, [2] Bristol Heart Institute and Translational Biomedical Research Centre, Bristol Medical School, Bristol Royal Infirmary, University of Bristol, Bristol, United Kingdom

OPEN ACCESS

Edited by:
Emiel Van Der Vorst, Institute for Molecular Cardiovascular Research (IMCAR), Germany

Reviewed by:
Marten A. Hoeksema, Amsterdam UMC, Netherlands
Judith Sluimer, Maastricht University, Netherlands

***Correspondence:**
Liam W. McQueen
lwm3@leicester.ac.uk

Specialty section:
This article was submitted to Atherosclerosis and Vascular Medicine, a section of the journal Frontiers in Cardiovascular Medicine

Received: *07 January 2022*
Accepted: *07 March 2022*
Published: *28 March 2022*

Citation:
McQueen LW, Ladak SS, Abbasciano R, George SJ, Suleiman M-S, Angelini GD, Murphy GJ and Zakkar M (2022) Next-Generation and Single-Cell Sequencing Approaches to Study Atherosclerosis and Vascular Inflammation Pathophysiology: A Systematic Review. Front. Cardiovasc. Med. 9:849675. doi: 10.3389/fcvm.2022.849675

Background and Aims: Atherosclerosis is a chronic inflammatory disease that remains the leading cause of morbidity and mortality worldwide. Despite decades of research into the development and progression of this disease, current management and treatment approaches remain unsatisfactory and further studies are required to understand the exact pathophysiology. This review aims to provide a comprehensive assessment of currently published data utilizing single-cell and next-generation sequencing techniques to identify key cellular and molecular contributions to atherosclerosis and vascular inflammation.

Methods: Electronic searches of Cochrane Central Register of Controlled Trials, MEDLINE, and EMBASE databases were undertaken from inception until February 2022. A narrative synthesis of all included studies was performed for all included studies. Quality assessment and risk of bias analysis was evaluated using the ARRIVE and SYRCLE checklist tools.

Results: Thirty-four studies were eligible for narrative synthesis, with 16 articles utilizing single-cell exclusively, 10 utilizing next-generation sequencing and 8 using a combination of these approaches. Studies investigated numerous targets, ranging from exploratory tissue and plaque analysis, cell phenotype investigation and physiological/hemodynamic contributions to disease progression at both the single-cell and whole genome level. A significant area of focus was placed on smooth muscle cell, macrophage, and stem/progenitor contributions to disease, with little focus placed on contributions of other cell types including lymphocytes and endothelial cells. A significant level of heterogeneity exists in the outcomes from single-cell sequencing of similar samples, leading to inter-sample and inter-study variation.

Conclusions: Single-cell and next-generation sequencing methodologies offer novel means of elucidating atherosclerosis with significantly higher resolution than previous methodologies. These approaches also show significant potential for translatability

into other vascular disease states, by facilitating cell-specific gene expression profiles between disease states. Implementation of these technologies may offer novel approaches to understanding the disease pathophysiology and improving disease prevention, management, and treatment.

Systematic Review Registration: https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42021229960, identifier: CRD42021229960.

Keywords: vascular inflammation, atherosclerosis, single-cell sequencing, next-generation sequencing, systematic review

INTRODUCTION

Atherosclerosis of the arteries remains the leading cause of morbidity and mortality worldwide (1, 2). Several risk factors such as hyperlipidemia, smoking, hypertension and diabetes have been implicated in facilitating atherogenesis through activation of the inflammasome (3). LDL accumulated in sub-endothelial areas of arterial bends and branches can be oxidized (oxLDL), activating endothelial cells (ECs), and initiating the recruitment of monocytes. These monocytes subsequently differentiate into macrophages and endocytose oxLDL, which accumulates in these cells leading to foam cell formation. Foam cells cannot process oxLDL and can rupture if large amounts accumulate inside them. This can lead to the deposition of more oxLDL into the artery wall triggering more inflammatory reactions and thus completing a vicious cycle (4, 5). EC activation by modified lipids can induce the expression of adhesion molecules, chemokines, and cytokines leading to monocytes recruitment from the blood stream which is a key event in the development of atherosclerosis (6). Several genetic deletion studies have demonstrated that inflammation is required for lesion formation. Moreover, vascular inflammation can also contribute to the development of thrombotic complications of atheroma as activated macrophages can produce proteolytic enzymes that degrade collagen and thereby alter the structure of the fibrous cap (7–9).

Over the years, many methods have been established to study the development and progression of vascular inflammation and atherosclerosis. The recent development and utilization of next-generation and single-cell sequencing techniques are additional tools that will improve our understanding of complex diseases including atherosclerosis progression (10, 11).

This review aims to provide a narrative synthesis of published literature that utilize single-cell and next-generation sequencing methodologies to explore key molecular and cellular targets related to atherosclerosis and vascular inflammation onset and progression, as well as determining the broader utility and translatability of these methodologies.

METHODS

This systematic review was performed following guidance from the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement standard (12). A study protocol was designed which conformed to the PRISMA protocol standard (13) and was registered at the International Prospective Register of Systematic Reviews (14).

Study Eligibility

The inclusion criteria were: (1) Any studies utilizing single-cell and/or next-generation sequencing methodologies to study vascular inflammation, atherosclerosis, or both; (2) Human or animal subjects demonstrating atherosclerosis and/or vascular inflammation (all species, all sexes); and (3) All study models (*in vivo*, *in vitro*, and *ex vivo*).

Exclusion criteria included: (1) No implementation of single-cell and/or next-generation sequencing methodologies; (2) Aortic, transplant, or neurological inflammation and/or atherosclerosis; (3) Vasculitis or other vascular conditions; (4) Genome-wide association studies or meta-analysis studies; and (5) Subjects with co-morbidities or non-healthy control conditions (6); Conference and meeting abstracts, case reports and literature reviews; (7) Studies not published in English.

Data Sources and Search Strategy

Electronic searches were conducted using the Cochrane Central Register of Controlled Trials, MEDLINE, and EMBASE without date or language restriction from inception until February 2022. The search strategy employed to determine studies of relevance utilized combinations of keywords such as "scRNA-seq," "single-cell sequencing," "next-generation sequencing," "vascular inflammation," "vasculitis," "atherosclerosis," and "arteriosclerosis." A full description of the search terms is listed in the **Supplementary Materials**. In addition, the reference lists of all retrieved articles were hand searched for further relevant studies not previously identified. Only papers that were published in English were considered for subsequent analysis. References from selected papers were additionally scanned for relevant articles to ensure the literature search was thorough. Reviewers L.W.M and R.A. performed the database searches.

Search results were imported into the Rayyan QCRI web app (15), and duplicates were identified and removed. To select relevant papers identified by the electronic search, papers were assessed initially by their title, then by analysis of their abstracts. Reviewers L.W.M., S.L. and M.Z performed this stage independently. Studies not excluded after this stage were then examined in full to assess their relevance. All authors then validated the final selected papers, and any discrepancies will be resolved by discussion.

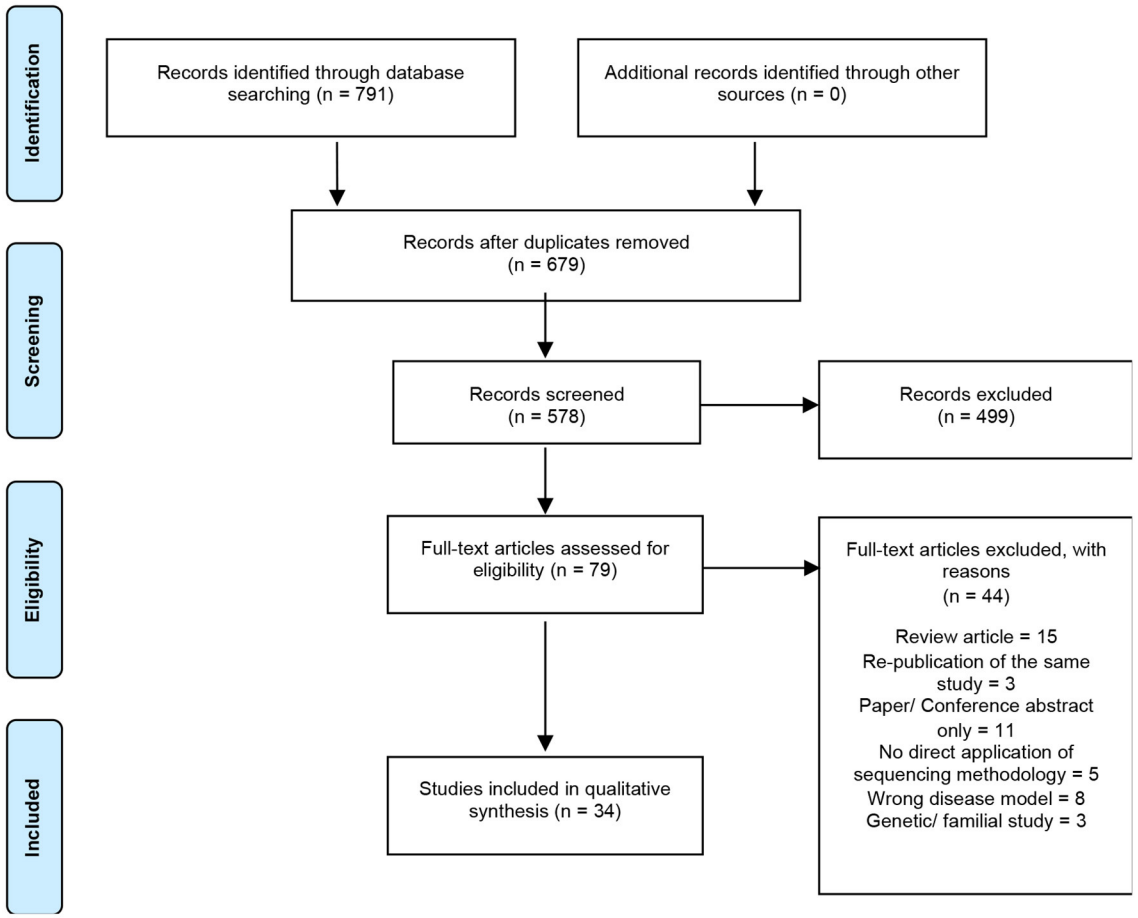


FIGURE 1 | Summary of electronic literature search protocol. PRISMA flowchart detailing process of systematic literature searching, screening and selection (*n* = number of studies).

Data Extraction

A standardized form was developed to extract data from the included studies for assessment of study quality and evidence synthesis. This form was tabulated using Microsoft Excel 2016 (Microsoft, Redmond, Washington). Data extraction first considered data from figures, tables and graphs (using digital ruler software where appropriate), followed by data extraction from the main text. Data extracted to the standardized form were categorized under the following headings: author, title, year, journal, research aims and objectives, subjects examined, cell type examined, sequencing method, research methodology, and key findings. Author L.W.M. performed data extraction, all authors validated the findings, and any discrepancies were resolved by discussion.

Study Outcomes

The primary outcome measure was to determine the contributions of single-cell and/or next-generation sequencing techniques in the field of atherosclerosis and/or vascular inflammation development and progression. The secondary outcome measure was to determine the utility and translatability of these findings and sequencing methodologies to other related vascular disease research.

Bias and Quality Assessment

Quality of included studies was assessed using the ARRIVE guidelines checklist (16). For assessment of the risk of bias, SYRCLE, a modified version of the Cochrane Risk of Bias tool specifically for animal intervention studies, was utilized (17). Following data extraction, author L.W.M. performed quality and risk of bias assessment. Any discrepancies were resolved by discussion between all authors.

Data Synthesis

A narrative synthesis of all included studies was performed, with all relevant data tabulated where appropriate. For all outcomes, data was extracted in text format, or as mean ± standard deviation for numerical values. Where applicable, continuous variables were summarized with standardized mean difference, and dichotomous data was summarized with risk and odd ratios. Where relevant data was missing, authors were contacted (where possible); otherwise, data was presented as described in the article

or **Supplementary Material**. Given the anticipated diversity of outcome measures, limited scope for statistical analysis was expected and as such, meta-analysis was not undertaken.

RESULTS

A total of 791 articles were identified by the search strategy, 679 following the removal of duplicates, which were then screened against the inclusion/exclusion criteria. Of these papers, only 34 were eligible for final synthesis (**Figure 1**). A summary of the characteristics of all included studies are reported in **Table 1**.

Studies using single-cell and/or next-generation sequencing to explore atherosclerosis and/or vascular inflammation appeared as early as 2012, however, there was an upsurge in the prevalence of these studies starting in 2017 (**Figure 2A**). The most employed sequencing methodology in these studies is single-cell sequencing ($n = 23$), which was first utilized in 2017 and features more frequently in subsequent years. Comparatively, next-generation sequencing was utilized less frequently ($n = 12$), with no significant trend in its use over time. A total of 10 articles utilized a combination of sequencing methods, with some articles using chromatin immunoprecipitation sequencing (ChIP-Seq) ($n = 4$), microRNA sequencing ($n = 4$), assay for transposase-accessible chromatin sequencing (ATAC-Seq) ($n = 2$) and cellular indexing of transcriptomes and epitopes sequencing (CITE-Seq) ($n = 1$) (**Figure 2B**). The most commonly studies cell types were smooth muscle cells (SMCs) ($n = 10$), followed by macrophages ($n = 6$) and stem/progenitor cells ($n = 5$), although these techniques are also applied more broadly to investigate the cellular and molecular characteristics of tissue, atherosclerotic plaque, and blood samples (**Figure 2C**). Interestingly, there appears to be a trend in the types of cells studied over time, with focus on smooth muscle cells almost tripling in prevalence from 2019 ($n = 2$) to 2021 ($n = 5$), whilst focus on macrophages has been consistent for a longer period (2017–2022), but at a lower prevalence (**Figure 2D**). All included studies were published in journals with impact factor $\geq$4, with the vast majority appearing in the publications Circulation and Circulation Research ($n = 10$) (**Figure 2E**).

Cell-Specific Sequencing

Smooth Muscle Cells

Ten articles (20, 27, 28, 31, 37, 41, 45–47, 50) addressed the role of SMCs in atherosclerosis and vascular inflammation. Four of these articles (20, 45, 46, 50) utilized single-cell sequencing exclusively, one (31) utilized next-generation RNA sequencing exclusively, one (47) utilized next-generation micro-RNA sequencing, and the remaining four (27, 28, 37, 41) utilized a combination of these methodologies, including the use of ChIP-Seq (27, 28, 37), ATAC-Seq (28) and CITE-Seq (27).

Pan et al. (20) identified a novel transient cell state in mouse models related to SMC phenotypic switching using single-cell sequencing, which exhibited upregulation of markers *Ly6A*, *VCAM1*, and *Ly6C1*, attributed to stem, endothelial and mesenchymal cells respectively, accompanied by loss of SMC-specific markers. These cells, termed SEM cells, were the result of transdifferentiation of SMCs. lineage analysis revealed potential for SEMs to differentiate back to an SMC phenotype, or into a "fibrochondrocyte" or macrophage-like cell. Additionally, these cells did not exhibit traditional mesenchymal markers, suggesting a unique SMC-derived cell state, and this unique state was shown to exist in human atherosclerotic lesions.

Wirka et al. (27) undertook single-cell sequencing of SMC-specific lineage-traced mice to determine the fate of SMCs at baseline and after high-fat diet. Sequencing of aortic root atherosclerotic plaques revealed two SMC clusters with distinct gene expression profiles which become less defined as the disease progressed. This second, modulated cluster expressed markers of SMC differentiation such as *CNN1*, *FN1*, and *COL1A1*, and appeared to be more closely related to a fibroblast-like phenotype with expression of markers decorin and biglycan. These cells, termed fibromyocytes (FMCs), were localized to vessel medias, suggesting phenotypic switching occurs prior to migration into the plaque, and FMCs appeared to be absent of any markers associated with macrophages suggesting FMCs exist as a discrete species. The *TCF21* gene was shown to exert control over the fate of SMC phenotypes, with upregulation in progressive atherosclerotic burden, and association with GWAS data suggesting *TCF21* action on SMC phenotype modulation toward an FMC phenotype has an atheroprotective role.

Kim et al. (28), utilizing an established SMC-specific lineage-traced mice models (27), performed single-cell sequencing of atherosclerotic aortic root tissue to identify the role of aryl hydrocarbon receptors (AHRs) on affecting SMC phenotype. AHR expression was found to colocalise with the previously identified FMC cell cluster, as well as with *TCF21*, to the lesion intima and fibrotic cap, further validated using a combination of RNAScope and immunofluorescence. Sequencing analysis of AHR-knockout models revealed an increased population of FMCs in the lesion intima vs. wild-type (WT), as well as a distinct cell cluster with expression profiles related to ossification, collagen organization and *TGFβ* signaling. Additionally, upregulation of factors specifically strongly related to chondrocytes were identified (e.g., *Sox9* and *Runx2*) in this cluster which the researchers termed "chondromyocytes" (CMC).

Mendez-Barbero et al. (31) undertook RNA sequencing of cultured murine vascular SMCs to identify the *TWEAK/Fn14* axis as a central regulator of SMC proliferation and migration in response to vascular injury. Comparison of gene expression between cells in the presence and absence of recombinant *TWEAK* (*rTWEAK*) treatment revealed upregulated expression of markers related to cell proliferation, migration, and motility in response to injury. The *TWEAK/Fn14* interaction was shown to downregulate cell cycle regulator $p15^{INK4B}$ and upregulated cyclin-dependent kinase 4/6 and cyclin D1 *via* induced phosphorylation of *MAPK* extracellular signal-regulated kinases 1 and 2 (*ERK1/2*) and *AKT*, as well as *NFκβ* subunit *p65*. Conversely, the absence of *TWEAK* using mouse knockout models was shown to inhibit SMC proliferation and markedly reduce the neointimal area of mice with wire injury-induced vessel damage, suggesting a central role for *TWEAK/Fn14* in facilitating vascular remodeling.

Alencar et al. (37) utilized a combination of single-cell and next-generation RNA sequencing, combined with ChIP-Seq, to

TABLE 1 | Data characterization for included studies.

References	Aims and objectives	Target	Subject examined	Sequencing methodology	Cell type examined
Tang et al. (18)	Aim to determine the existence of Sca1+ vascular stem cells *in vivo*, and their role in vascular repair	Vessel repair	Mouse	Single-cell RNA-sequencing	Stem cells
Sharma et al. (19)	Understand whether Tregs are essential for the regression of atherosclerotic plaques, and if so, to identify key mechanisms by which Tregs contribute to plaque repair and contraction	Atherosclerotic regression	Mouse and cell culture	Single-cell RNA-sequencing	Lymphocytes
Pan et al. (20)	Understand SMC transdifferentiation during atherosclerosis and to identify molecular targets for disease therapy	SMC phenotypic switching	Mouse and human	Single-cell RNA-sequencing	Smooth muscle cells
Gu et al. (21)	Aim to perform scRNA-seq of aortic adventitial cells from WT and ApoE-deficient mice to explore their heterogeneity, diverse functional states, dynamic cellular communications, and altered transcriptomic profiles in disease	Adventitial transcriptome	Mouse	Single-cell RNA-sequencing	Adventitial cells
Cochain et al. (22)	Aim to determine and classify macrophage heterogeneity in both healthy and atherosclerotic aortas of mice using single-cell RNA-sequencing technology	Macrophage heterogeneity in atherosclerosis	Mouse and human	Single-cell RNA-sequencing	Macrophages
Kokkinopoulos et al. (23)	Aim to clarify the role of AdvSCA-1+ progenitor cells in native atherosclerosis, *via* elucidation of their differential gene expression profile between atherosclerosis-resistant and atherosclerosis-susceptible mice	Adventitial progenitor cells	Mouse	Single-cell RNA-sequencing	Stem cells
Rahman et al. (24)	Aim to investigate the source of, and functional requirement for, M2 macrophages in atherosclerosis regression, using a mouse aortic transplantation model	M2 macrophages in atherosclerosis regression	Mouse	Single-cell RNA-sequencing	Macrophages
Gu et al. (25)	Aim to further elucidate the role of perivascular adipose tissue (PVAT), with specific interest in characterizing the transcriptomic profile of PVAT-derived mesenchymal stem cells (PV-ADSCs) and their role in vascular remodeling	Adventitial cells in vascular remodeling	Mouse and cell culture	Single-cell RNA-sequencing	Adventitial cells and stem cells
Winkels et al. (26)	Aim to define an atlas of the immune cell landscape in atherosclerotic lesions, using single-cell RNA-sequencing and mass cytometry (cytometry by time of flight), *via* comparison of healthy and diseased arteries in mouse and human	Immune cells in atherosclerotic lesions	Mouse and Human	Single-cell RNA-sequencing	Lymphocytes
Wirka et al. (27)	Aim to determine: (1) which cell type(s) express Tcf21 during lesion development, (2) how does Tcf21 affect the phenotype of these cells, and (3) how does Tcf21 affect disease risk	SMC phenotypic switching	Mouse, cell culture, and human	Single-cell RNA-sequencing (CITE-Seq and ChIP-Seq) AND next-generation RNA-sequencing	Smooth muscle cells
Kim et al. (28)	Further investigate the specific effects of environment-sensing aryl hydrocarbon receptors (AHR) on the vascular SMC phenotype in atherosclerotic disease	SMC phenotypic switching	Mouse and cell culture	Single-cell RNA-sequencing AND next-generation RNA-sequencing (ChIP-Seq AND ATAC-Seq)	Smooth muscle cells

(Continued)

TABLE 1 | Continued

References	Aims and objectives	Target	Subject examined	Sequencing methodology	Cell type examined
Kim et al. (29)	Aim to examine the transcriptomic profiles of foamy and non-foamy macrophages isolated from atherosclerotic intima, to determine their functional role and contribution to the disease	Transcriptome difference of foamy and non-foamy macrophages	Mouse and human	Single-cell RNA-sequencing AND next-generation RNA-sequencing	Macrophages
Steffen et al. (30)	Scrutinize the identity of sca1+/flk1+ cells, establish a phenotype for these cells, to amend the current hypothesis of vascular regeneration by circulating cells and gain understanding of their role in atherosclerotic disease	Vascular (endothelial) regeneration	Mouse	Next-generation RNA-sequencing	Stem cells
Mendez-Barbero et al. (31)	Aim to further elucidate the role of TWEAK/Fn14 in vascular remodeling, by identifying the downstream molecular mediators of this relationship, and how this has a functional effect on vascular smooth muscle cells (VSMCs)	SMC proliferation and migration	Mouse, cell culture, and human	Next-generation RNA-sequencing	Smooth muscle cells
Lai et al. (32)	Aim to explore the dynamic expression of EndMT genes in vascular endothelial cells under atheroprotective pulsatile shear stress and atheroprone oscillatory shear stress using RNAseq	Endothelial-to-mesenchymal transition	Mouse and cell culture	Next-generation RNA-sequencing	Endothelial cells
Karere et al. (33)	Aim to determine miRNA expression profile differences in baboons with low and high serum low-density lipoprotein cholesterol in response to diet. Aim to establish if any of these miRNAs are relevant to dyslipidemia and risk of atherosclerosis	MicroRNA relevance in dyslipidemia	Baboons	Next-generation microRNA-sequencing	Blood (micro RNAs in low/high LDL-C baboons with HCHF diet)
Depuydt et al. (34)	Aim to utilize single-cell transcriptomics and chromatin accessibility to gain a better understanding of the cellular heterogeneity and pathophysiology underlying human atherosclerosis	Atherosclerotic plaque composition	Human	Single-cell RNA-sequencing and single-cell ATAC-sequencing	Atherosclerotic plaques from carotid artery
Li et al. (35)	To study the role of macrophages and monocytes. In the CV system using a cell line model; to study the effect of matrix stiffness on macrophages behavior in atherosclerosis; to determine the synergistic role of ox-LDL and matrix stiffness on macrophage behavior, such as migration, inflammation, and apoptosis	Matrix stiffness on macrophage behavior (inflammation)	Cell culture	Next-generation microRNA-sequencing	Macrophages
Lin et al. (36)	Aim to improve the understanding of the origins and fates of macrophages in progressing and regressing atherosclerotic plaques using a combination of single-cell RNA sequencing and mouse genetic fate mapping	Macrophage heterogeneity in atherosclerosis	Mouse	Single-cell RNA-sequencing	Macrophages
Alencar et al. (37)	Aim to further define SMC subsets within atherosclerotic lesions, with the goal of identifying factors and mechanisms that promote beneficial SMC phenotypic transitions as novel therapeutic targets	SMC phenotypic switching	Mouse and human	Single-cell RNA-sequencing, Next-generation RNA-sequencing and ChiP-Seq	Smooth muscle cells
Li et al. (35)	Aim to clarify the specific functions and regulatory mechanisms of macrophage subsets present in vascular inflammation and atherosclerosis	Macrophage heterogeneity in atherosclerosis	Human and cell culture	Next-generation microRNA-sequencing	Blood (exosome microRNAs effect on macrophages)

(Continued)

TABLE 1 | Continued

References	Aims and objectives	Target	Subject examined	Sequencing methodology	Cell type examined
Wolf et al. (38)	Aim to interrogate the function of autoreactive CD4+ T cells in atherosclerosis, through the use of a novel tetramer of major histocompatibility complex II to track T cells reactive to the mouse self-peptide apo B978-993 (apoB+) at the single-cell level	Immune cells in atherosclerotic lesions (T-cells)	Mouse	Single-cell RNA-sequencing AND next-generation RNA-sequencing	Lymphocytes
Zhou et al. (39)	Aim to investigate how the endothelial glucocorticoid receptor regulates vascular inflammation	Regulation of vascular inflammation *via* endothelial glucocorticoid receptors	Mouse and cell culture	Next-generation RNA-sequencing and ChiP-Seq	Endothelial cells
Bao et al. (40)	Aim to identify the transcriptome and proteome of stable and unstable atherosclerotic plaques	Atherosclerotic plaque transcriptome and proteome (stable vs. unstable)	Human	Next-generation RNA-sequencing	Atherosclerotic plaques (stable vs. unstable)
Gallina et al. (41)	Aim to identify the mechanisms underlying vascular smooth muscle cell phenotypic transitions associated with atherosclerosis and vascular injury, with specific focus on the glutamate receptor signaling	SMC phenotypic switching	Mouse, rat, and human	Single-cell RNA-sequencing and next-generation RNA-sequencing	Smooth muscle cells
Jiang et al. (42)	Aim to investigate the identity and role of CD34+ cells in vascular regeneration	Vascular (endothelial) regeneration	Mouse	Single-Cell RNA-Sequencing	CD34+ progenitor cells
Kan et al. (43)	Aim to characterize the cellular heterogeneity and diverse functional states within the wall of the ascending aorta in healthy and diseased mice using scRNA-seq to better understand the etiology and progression of aortic disease in HFD-induced obesity	Cell composition of healthy and diseased arteries	Mouse	Single-cell RNA-sequencing	Healthy and diseased aortas
Li et al. (44)	Aim to determine the specific contributions of disturbed flow on the heterogeneity of cells within the affected arterial vasculature	Effect of disturbed flow (shear stress) in the cellular and molecular composition of carotid arteries	Mouse	Single-cell RNA-sequencing	Carotid arteries (under disturbed flow)
Liang et al. (45)	Aim to utilize scSeq to examine VSMC phenotype in carotid artery calcified plaque cores and surrounding tissue to determine phenotype switching markers and mechanisms	SMC phenotypic switching	Human	Single-cell RNA-sequencing	Smooth muscle cells
Lin et al. (46)	Aim to understand the origin and phenotypic heterogeneity of smooth muscle cells (SMCs) contributing to intimal hyperplasia, with specific focus on how vascular cells adapt to the absence of elastin (Eln)	SMC phenotypic switching	Mouse	Single-cell RNA-sequencing	Smooth muscle cells
Brandt et al. (47)	Aim to comprehensively characterize the transcriptomic profile of phenotypically modulated VSMCs and identified mediators of VSMC transdifferentiation and their link to plaque rupture in human atherosclerosis	SMC phenotypic switching	Mouse	Single-cell RNA-sequencing	Smooth muscle cells

(Continued)

TABLE 1 | Continued

References	Aims and objectives	Target	Subject examined	Sequencing methodology	Cell type examined
Newman et al. (48)	Aim to identify which cells, factors and mechanisms contribute to the fibrotic cap formation in atherosclerotic lesions	Fibrous cap composition	Mouse	Single-cell RNA-sequencing and next-generation RNA-sequencing	Atherosclerotic plaques (fibrotic cap)
Quiles-Jimenez et al. (47)	Aim to clarify the specific function of DNA glycolase Neil3 in the development of atherosclerosis, specifically in regard to vascular smooth muscle cell phenotypic modulation.	SMC phenotypic switching	Mouse and cell culture	Next-generation microRNA-sequencing	Smooth muscle cells
Burger et al. (49)	Aim to identify heterogeneous leukocyte clusters with distinct atherosclerosis disease-relevant gene expression signatures and to unveil their role in atherosclerosis pathology	Resident macrophage function in atherosclerosis	Mouse	Single-cell RNA-sequencing	Macrophages + smooth muscle cells

further define SMC phenotypes which affect the pathogenesis of atherosclerotic lesions. SMC^{KLF4} and SMC^{OCT4} KO and WT mice models were developed to validate the gene expression modulations of *Klf4* and *Oct4* in lesion development and stability. Next-generation sequencing and ChIP-Seq analysis identified contrasting lesion morphology phenotypes in each KO model, with *Oct4* targets enriched for genes involved in cell pluripotency and migration, whilst *Klf4* targets were enriched in genes involved in leukocyte recruitment and extracellular matrix (ECM) organization, suggesting an opposing role for these transcription factors in lesion phenotype. Single-cell sequencing of brachiocephalic artery lesions identified 14 distinct clusters of which seven were shown to derive from an SMC origin *via* lineage-tracing modeling, with these clusters lacking traditional SMC markers (e.g., *Myh11*) but expressing markers such as *Vcam1*, *Lgals3*, *Spp1*, and *Sox9* among others, suggesting significant cellular plasticity of cells during lesion development. The $Lgals3^+$ cluster was identified as a chondrogenic-like state, which, in *Klf4* KO mice, was markedly reduced in line with a reduction in lesion size and fibrous cap thickness. $Lgals3^+$ SMCs were shown to exist as a unique intermediate stem cell-like, ECM remodeling phenotype, representing up to two thirds of all lesion SMCs, which are further differentiated into several other pro-inflammatory and *Klf4* dependent osteogenic phenotypes which contribute to plaque calcification and destabilization.

Gallina et al. (41) investigated the mechanisms underpinning vascular SMC phenotypic switching specific to AMPA-type glutamate receptors shown to exert an effect on pulmonary vascular remodeling. Microarray transcriptome and single-cell sequencing revealed human atherosclerotic plaque samples contain relevant signaling and receptor components for glutamate turnover and signaling, specifically GRIA1 and GRIA2 which were exclusively detected in cells of mesenchymal (primarily SMC) origin. Expression of these transcripts were associated with phenotypic transition of SMCs, as determined by next-generation sequencing from a rat carotid artery injury and repair model at different timepoints. *Gria1* were significantly inversely correlated with traditional SMC marker expression (e.g., *Myh11*) in injured arteries, whilst *Gria2* was expressed at lower levels, more prevalent in non-injured vessels and had a positive correlation with SMC marker expression.

Liang et al. (45) undertook single-cell sequencing of carotid artery calcified cores and paracellular tissue, compared to patient matched proximal adjacent to carotid artery tissue, to investigate the phenotypic transitions of SMCs during the calcification process in humans. A total of 20 clusters were categorized from this analysis, and comparison between tissue types identified a greater proportion of T-cell and monocytes in the disease-associated tissue, whilst EC and fibroblast populations were higher in the control tissue. However, the largest population difference between control and diseased tissue was attributed to a macrophage-like SMC cluster, with enrichment analysis identifying gene expression related to inflammatory signaling, immune response, degranulation, and migration. Protein interaction analysis identified significant involvement of MMP9, CXCL8, SPP1, and LGALS3 among others, with upstream transcription factors such as NFkB1, RELA, SP1, JUN, and SPI all associated with downstream gene regulation.

Lin et al. (46) sought to identify SMC subpopulations which arise in response to intimal hyperplasia development through the development of a genetic mouse model of elastin insufficiency. Single-cell sequencing of cells isolated from the ascending aortas of these mice identified 14 distinct clusters, of which three were attributed to SMCs as defined by markers *Myh11*, *Acta2* and *Cnn1*, respectively. $Myh11^+$ cells had the most significant expression of SMC marker genes, with $Acta2^+$ cells exhibiting a more myofibroblast-like expression profile and $Cnn1^+$ cells expressing genes related to cell proliferation. Subsequent pathway analysis indicated a closer transcriptional profile between the $Acta2^+$ and $Cnn1^+$ clusters, suggesting that $Myh11^+$ SMCs are further differentiated. Reactome pathway analysis suggests a less contractile phenotype for the $Acta2^+$ and $Cnn1^+$ clusters, with the former exhibiting a fibroblast,

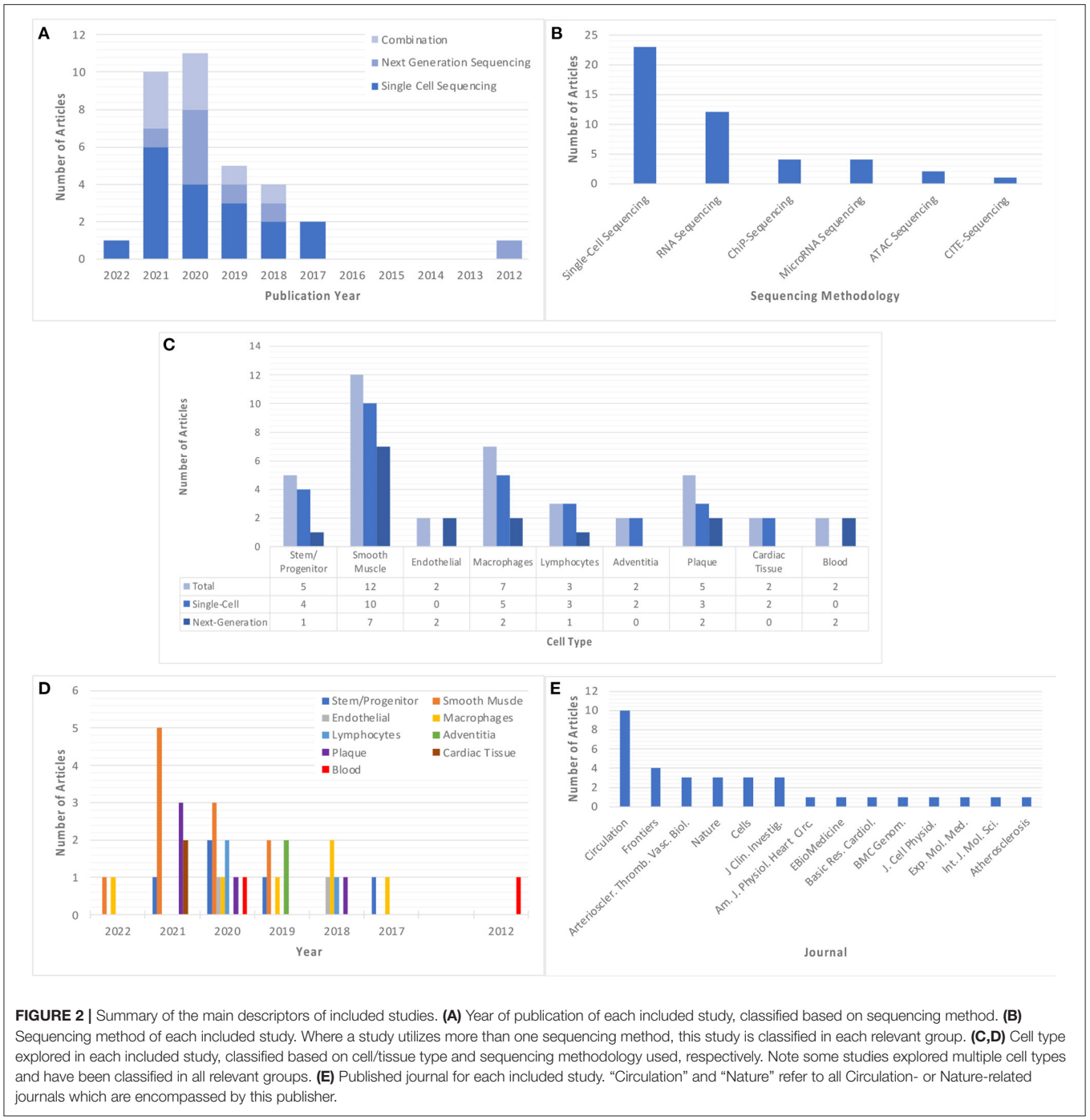


FIGURE 2 | Summary of the main descriptors of included studies. **(A)** Year of publication of each included study, classified based on sequencing method. **(B)** Sequencing method of each included study. Where a study utilizes more than one sequencing method, this study is classified in each relevant group. **(C,D)** Cell type explored in each included study, classified based on cell/tissue type and sequencing methodology used, respectively. Note some studies explored multiple cell types and have been classified in all relevant groups. **(E)** Published journal for each included study. "Circulation" and "Nature" refer to all Circulation- or Nature-related journals which are encompassed by this publisher.

ECM-synthesizing phenotype, and the latter a more proliferative phenotype. Validation experiments in immunostained tissue sections for cluster specific markers were able to localize these two subgroups specifically to the neointimal area.

Brandt et al. (50) undertook single-cell sequencing of CD45^{-} cells isolated from atherosclerotic aortas of *ApoE*$^{-/-}$ mice on a normal and high cholesterol diet, with cells categorized based localization to atheroprone and atheroresistant regions (aortic arch and descending thoracic aorta, respectively). Differential regulation of genes was identified in the atheroprone cell cluster based on diet associated with apoptosis, inflammation, atherogenesis among others. In the atheroresistant cell groups, despite no detectable plaque formation, a high cholesterol diet still resulted in upregulation of genes related to an atherogenic stress response, including regulatory genes for SMC differentiation, apoptosis, and inflammation. Unsupervised clustering of a pooled cell population identified seven distinct CD45– clusters—four of which were exclusive to the atheroprone

region with gene expression indicating a foam cell-like, inflammatory, calcifying, phenotype transitional subgroups with loss of SMC contractile genes (*Myh11* and *Acta2*). Clusters specific to atheroresistant regions maintained and even increased contractile marker expression. In particular, the authors identified a significantly upregulated gene associated with the macrophagic/calcific phenotype—growth differentiation factor 10 (GDF10)—which was validated *in vitro* to modulate ossification, osteoblast differentiation and SMC phenotypic switching.

Quiles-Jimenez et al. (47) investigated the function of a DNA glycolase *Neil3*, associated with a role in atherogenesis, using a $ApoE^{-/-}/Neil3^{-/-}$ mouse model and NEIL3 knockout human aortic SMC cell culture. These mice had significantly increased atherosclerotic lesion areas, without changes in systemic lipid levels. Cell markers in the plaques remained similar to $ApoE^{-/-}$ controls, but with greater SMC medial layer thickness and disorganization, and a clear phenotypic shift was identified in these cells with gene expression associated with increased proliferation, lipid-accumulation and de-differentiation. *In vitro* analysis revealed that NEIL3 deficiency is accompanied by a phenotypic shift to macrophage-like characteristics, with expression of markers CD68, TGFβ, and MMP2. Next-generation messenger RNA (mRNA)-sequencing identified gene enrichment for cell proliferation/apoptosis, differentiation, growth factor response, and others when compared to $ApoE^{-/-}$ controls. This sequencing data, combined with proteomic analysis, revealed that Neil3-deficiency leading to SMC phenotype switching occurs *via* activation of the Akt signaling.

Macrophages

Six articles (22, 24, 29, 36, 49, 51) explored the contributions of macrophages to atherosclerosis and vascular inflammation. Four (22, 24, 36, 49) of these articles utilized single-cell sequencing exclusively, one (51) utilized next-generation microRNA sequencing, and the remaining paper (29) utilized a combination of single-cell and next-generation RNA sequencing methods.

Cochain et al. (22) utilized single-cell sequencing to explore the heterogeneity of aortic macrophages subsets within murine atherosclerosis. $CD45^{+}$ cells were extracted from both non-diseased and atherosclerotic low-density lipoprotein receptor deficient ($Ldlr^{-/-}$) male mice, and unsupervised clustering analysis revealed the presence of 13 unique cell clusters, five of which were exclusively expressed in atherosclerotic models. These included $CD8^{+}$ T-cells, monocyte-derived dendritic cells (MoDC), monocytes and 2 distinct macrophage clusters. Macrophages comprised of three classes from this data: (1) traditionally activated (M1) macrophages expressing atherogenic markers such as chemokine receptor *CCL3* and *IL1B*; (2) M2 macrophages expressing anti-atherogenic markers including *F13A1* and *CCL24*; and (3) a novel, smaller subset with significant expression of *TREM2*. Gene ontology analysis of this cluster highlighted unique functions including lipid metabolism and cholesterol regulation, with expression resembling osteoclasts, indicating a potential role in plaque calcification.

Rahman et al. (24) using single cell sequencing methods tested the hypothesis that M2 macrophages in atherosclerotic plaques derive from newly recruited monocytes. This involved utilizing an aortic transplantation model to assess plaque regression in normolipidemic and $ApoE^{-/-}$ mice against mice deficient in chemokine receptors *CCR2*, *CX3CR1* and *CCR5* involved in inflammatory ($Ly6C^{hi}$) or migratory ($Ly6C^{lo}$) monocyte recruitment. Their results provide strong evidence that plaque regression and inflammation resolution is dependent on the recruitment from the $Ly6C^{hi}$ population, generally considered M1 macrophage precursors, and that $Ly6C^{lo}$ macrophages are unable to fulfill this role. Single-cell sequencing and fate-mapping studies further indicated that polarization of these macrophages to an M2 phenotype is dependent on the action of *STAT6*.

Kim et al. (29) utilized $CD45^{+}$ cells from murine atherosclerotic aortas of WT, $ApoE^{-/-}$ and $Ldlr^{-/-}$ mice, fed a four, eight or 12-week western diet, to investigate the transcriptomic profiles of foamy vs. non-foamy macrophages in the intima of diseased vessels. Unsupervised clustering analysis of single-cell sequencing data identified 11 distinct leukocyte subpopulations, which were subsequently categorized using a fluorescent lipid-labeling flow cytometry method capable of determining lipid-laden foam cells based on granularity. This approach suggested that most lipid-laden, foamy leukocytes originate from clusters with expression heavily indicative of macrophage origin. Next-generation RNA sequencing and gene set enrichment analysis of the foamy and non-foamy macrophage clusters revealed upregulation of markers in non-foamy macrophages corresponding to leukocyte recruitment and inflammation progression. In contrast, foamy macrophage expression markers were more closely correlated with lipid metabolism, lipid transport, and oxidative phosphorylation, suggesting an anti-inflammatory role.

Li et al. (51) identified a role for both matrix stiffness and ox-LDL in modifying the behavior of macrophages in healthy and diseased conditions. Matrix stiffness and ox-LDL were shown to increase adherence of macrophages, coupled with an increased inflammatory response (TNFα and IL1B). Macrophage mobility appeared to be increased in disease conditions, however, the presence of ox-LDL reduced this migratory capacity. Analysis of microRNA differential expression between healthy and disease states indicated upregulation of fatty acid synthesis, MAPK signaling, p53 signaling, and apoptosis.

Lin et al. (36) developed a genetic fate mapping approach specific to circulating $CX3CR1^{+}$ macrophage precursors to assess their contributions in progressing and regressing atherosclerotic plaques. Single-cell sequencing identified 11 clusters of specifically myeloid lineage in a combined progression and regression dataset, which implied a wide range of macrophage activation states existed in these cells beyond the classical M1 and M2 definition, with activation states in greater number during atherosclerosis progression compared to regression. Three clusters were specific to the regression model, with one exhibiting a B-cell like phenotype ($Ebf1^{hi}Cd79a^{hi}$), and another with upregulation of heat shock proteins (HSP^{hi}) suggestive of a protective role. Finally, one cluster had a distinctive transcriptomic profile enriched in cell cycle and cell

proliferative genes, suggesting the existence of a self-renewing monocyte state existing within the inflamed tissue, rather than an immediate differentiation of these cells upon migration to the plaque.

Burger et al. (49) undertook single-cell sequencing of $CD45^+$ cells isolated from descending thoracic aorta and aortic arch from $ApoE^{-/-}$ mice on both a normal chow and high-cholesterol diet. Plaque formation predominated in the aortic arch tissue but was negligible in the descending thoracic aorta. Clustering was defined based on tissue type and diet, which revealed 12 distinct clusters—three clusters specific to the atheroprone aortic arch with expression of inflammatory monocyte/macrophage and resident macrophage related genes and five clusters specific to the atheroresistant descending thoracic aorta. Of note, one unique cluster of resident-like macrophages ($Lyve1^+$ macrophages) which were shown to expand with atherosclerotic plaques progression, and that these cells exhibited a pro-osteogenic action, *via* their high expression of CCL24, which encouraged the phenotypic transition of vascular smooth muscle cells.

Stem/Progenitor Cells

Five articles (18, 23, 25, 30, 42) specifically address the role of stem/progenitor cells in atherosclerosis and vascular inflammation. Four (18, 23, 25, 42) of these articles utilize single-cell sequencing explicitly, whilst the remaining article (30) utilized next-generation RNA sequencing.

Tang et al. (18) used a combination of single-cell RNA sequencing, genetic lineage tracing mouse models, and cell fate mapping to identify a subpopulation of vascular stem cells—$sca1^+PDGFRa^+$—which generate *de novo* SMCs in the media of arteries following severe vascular injury. Interestingly, the data show that these cells only contribute to vessel repair in cases of severe injury, with cell fate mapping and wire-injury modeling indicating that these cells remaining complacent in homeostasis and following minor injury.

Kokkinopoulos et al. (23) explored the role of murine adventitial $sca1^+$ cells in hyperlipidemia-induced atherosclerosis, by means of differential gene expression analysis of WT and $ApoE^{-/-}$ mouse models. Sequencing and gene ontology enrichment depicted these cells to have expression characteristics representative of migratory and locomotive action, as well as cytoskeletal organization and endopeptidase activity, suggesting involvement in epithelial-to-mesenchymal (EMT) transition. Additionally, it seems that both LDL-bound and free cholesterol further enhance this migratory capacity as well as inhibiting their differentiation capacity to ECs and SMCs both *in vitro* and *in vivo*. This increased migratory capacity upon lipid loading appears to be the result of upregulation of microRNA (miRNA) *miR-29b*, which in turn induces *SIRT1* and *MMP9*. Combined, the authors suggest a direct link between adventitial progenitor cells function and blood cholesterol levels.

Gu et al. (25) was able to identify two distinct $sca1^+$ clusters from single cell sequencing of perivascular adipocyte-derived stem cells (PV-ADSCs), the first of which demonstrated marker expression characteristic of endothelial cells (e.g., CD31 and Cadherin 5) with pathway association to VEGF receptor activity and PPAR signaling, indicative of angiogenic potential. Cluster two showed expression of markers associated with SMC differentiation, such as TGFβ signaling, PI3K-AKT signaling, PPAR binding and IGF binding. $Sca1^+$ cells of the second phenotype were shown to significantly contribute to neointimal formation by differentiation toward a functional SMC phenotype in mouse vein graft models.

Steffen et al. (30) utilized next-generation RNA sequencing and PCR array analysis to scrutinize the identity of $sca1^+/flk1^+$ cells, which were hypothesized to exist as endothelial progenitor cells that are upregulated in response to endothelial injury. Mice were subjected to electrical injury to the left common carotid artery and sacrificed 5 days post-injury. Sequencing analysis of purified $sca1^+/flk1^+$ cells from these mice depicted expression levels highly comparable to regulatory B2-cells, including markers such as CD19, CD22, and CD79a/b. Additionally, surface markers such as CD1 and CD86 were shown to be highly expressed, as well as CD38 known to be specific to human regulatory B-cells, indicating that these cells exist as precursors to B2-like cells and not ECs.

Jiang et al. (42) utilized single-cell sequencing to investigate the exact identity and role of $CD34^+$ cells using femoral artery tissue from WT C57BL/6J mice. This data was compared to lineage-traced Cd34-$CreER^{T2}$; R26-tdTomato mice which underwent guide-wire injury, and the analysis revealed a heterogenous perivascular tissue population of $CD34^+$ cells predominantly attributed to mesenchymal and EC origin. These cells were shown to contribute to endothelial regeneration and microvascular remodeling following injury. Bone-marrow transplantation experiments identified that the cells contributing to vascular repair are of non-bone marrow origin, with ablation of these cells aggravating adverse remodeling. Further single-cell sequencing of cells from the vascular injury model was undertaken, which revealed altered frequencies of cells at different repair timepoints in response to injury and marked changes in $CD34^+$ subpopulations, such as SMCs and myofibroblasts of $CD34^+$ origin. Pseudotime analysis of these cells, followed by cell differentiation experiments revealed a possible transition of adventitial $CD34^+$ mesenchymal cells to ECs. Network analysis identified microRNA-21 as a negative modulator of the Smad7-pSmad2/3 pathway resulting in endothelial differentiation.

Endothelial Cells

Only two articles (32, 39) specifically explored the contributions of endothelial cells in atherosclerosis and vascular inflammation, both of which utilized next-generation RNA sequencing with one article (39) also opting to utilize the ChIP-Seq methodology.

Lai et al. (32) utilized a multi-time point approach to explore the effects of pulsatile shear (PS) and oscillatory shear (OS) stress on human primary ECs endothelial-to-mesenchymal (EndMT) genes expression. RNA sequencing of cells exposed to either PS or OS was undertaken at different time points over a 24-h period. Sequencing analysis of the PS-exposed cells showed an increasing expression of endothelial specific markers (e.g., CD31, vWF) over the 24-h period, whilst OS showed no significant change in EC marker expression. However, markers specific to mesenchymal cells were upregulated due to OS, including

VCAM1 and SM22α, as well as inducing reactive oxygen species (ROS)-associated genes.

Zhou et al. (39) looked to explore how glucocorticoid receptors (GR) regulate ECs function using ChIP-Seq for primary mouse lung ECs under several conditions, including. control, GR siRNA-treated, glucocorticoid (dexamethasone) treated and a combination of these methods. These results showed significant binding close to the transcriptional start site, with 65 of the top 1,000 peaks showing both classic glucocorticoid responsive elements and *de novo* binding motif. RNA-Sequencing was undertaken which identified 231 glucocorticoid responsive genes and 203 genes differentially regulated by GRs. Comparison to GR ChIP-seq data in A549 cells revealed similar enrichment profiles for four main pathways: Wnt signaling, cytokine/chemokine signaling, angiogenesis, and cadherin signaling.

Lymphocytes

Three articles (19, 26, 38) investigated the contributions of lymphocytes in atherosclerosis and vascular inflammation, all of which utilized single-cell sequencing with one article (38) also opting to utilize next-generation RNA sequencing.

Sharma et al. (19) used several independent mouse models undergoing atherosclerotic plaque regression to illustrate an expansion in regulatory T-cell (Treg) populations compared to progressing and baseline plaques. Single-cell sequencing of $CD45^+$ cells isolated from mice with progressive and regressive atherosclerosis depicted a greater proportion of thymus-derived Treg markers (e.g., NRP1) in progressing plaques, whilst regressing plaques expressed a higher proportion of markers representative of periphery induced T cells (e.g., Ly6A). Pathway analysis revealed that Treg cells in regressing plaques exhibited expression profiles associated with increased lymphocyte activation and increased metabolic activity, coupled with increases in genes such as TGFβ and IGF1, which were nullified by antibody mediated Treg depletion. Tregs were also shown to alter the macrophage landscape in regressing plaques, with increases in macrophage migration and death, coupled with a decrease in their proliferative capacity.

Winkels et al. (26) utilized single-cell sequencing to specifically elucidate the heterogeneous immune landscape of mouse atherosclerotic lesions, which revealed five distinct T-cell clusters, 2 B-cell clusters and a natural killer cell cluster. Interestingly, T-cells were revealed to accumulate predominantly in the lesion (alongside macrophages), whilst B-cells were localized to the surrounding vessel media and tissue. Comparison of chow and high fat diet mice revealed a transcriptional profile switch from a recruitment to pro-inflammatory phenotype, and mass cytometry analysis in response, particularly in the case of T helper cells (T_H2) which expressed genes related to cytokine/chemokine expression and cell proliferation. Specific analysis of B-cell populations identified three distinct clusters, with the largest cluster ($CD43^{high}B220^{neg}$) associated with antigen presentation, cell adhesion and antibody generation. The $CD43^{low}B220^{high}$ cluster correlated with genes responsible for apoptosis and TNF-signaling, and the smallest cluster $CD43^{neg}B220^{high}$ related to cell division.

Wolf et al. (38) was able to identify the existence of a population of $ApoB^+$ $CD4^+$ T-cells in the lymph nodes of healthy C57BL/6 mice through a combination of *in silico* analysis, coupled with the development of a novel fluorochrome-coupled tetramer of recombinant MHC-II capable of binding and detecting $CD4^+$ T-cells *via* flow cytometry. Interestingly, validation in both WT and $ApoE^{-/-}$ mice revealed ApoB reactive memory T-cells are progressively activated in conditions of hyperlipidemia, and that their existence predates the onset of atherosclerotic disease. These cells appear to co-express markers associated with regulatory and effector T-cell phenotypes *via* flow cytometric analysis, and single-cell sequencing analysis identified a unique phenotype in $ApoB^+$ cells which co-expressed regulatory and helper (T_H1, T_H17, and T_{FH}) markers. Next-generation was subsequently undertaken in old and young $ApoE^{-/-}$ mice, which identified a gradual phenotypic transition of ApoB-Reactive $CD4^+$ regulatory T cells to a pro-inflammatory effector T-cell phenotype, which may be facilitated by the pro-inflammatory environment imposed by atherosclerosis. Further single-cell sequencing specific for aortic T-cells isolated from moderately (chow-fed) and severely (western diet) atherosclerotic $ApoE^{-/-}$ mice. One predominant cluster (expressing markers of T_H1 and T_H17) appears to be selectively expanded with respect to disease severity.

Tissue/Blood-Specific Sequencing

Atherosclerotic Plaque

Five articles (26, 34, 40, 43, 48) explored the total cellular and molecular composition of atherosclerotic plaques to identify heterogeneity within plaque areas and in comparison, to non-diseased controls. Three articles (26, 34, 43) utilized single-cell sequencing [one of which also utilized single-cell ATAC-Seq (34)], one article (40) utilized next-generation RNA sequencing and the last (48) used a combination of both methodologies. The work of both Winkels et al. (26) and Kan et al. (43) address atherosclerotic plaques but have been addressed in other more relevant sections of this review.

Depuydt et al. (34) undertook a combination of single-cell RNA and ATAC-Seq methods to explore the total cellular heterogeneity of atherosclerotic plaques in humans. Fourteen distinct populations were determined, including ECs and SMCs, and a host of immune cell subpopulations for T-cells, B-cells, and macrophages. ECs comprised four subclasses, three of which displayed expression markers indicative of activated endothelium with a role in inflammation progression and cell adhesion. The fourth subset displayed classical smooth muscle markers such as ACTA2 and MYH11, suggesting involvement in endothelial-to-mesenchymal transition (EndMT). Two subclasses of SMCs were also identified—the first with traditionally recognized SMC markers (e.g., ACTA2), whilst the other represented a synthetic class expressing COL1A1, MGP, and KLF4, which suggest a pseudo-macrophage phenotype. T-cell clusters were isolated to identify five $CD4^+$ and three $CD8^+$ clusters, distinguished by activation state. $CD4^+$ T-cells ranged in gene expression, with two associated with cytotoxicity and pro-inflammatory pathways, two associated with a naïve phenotype, and the final with a classical regulatory role. Of the $CD8^+$ clusters, one exhibits a

terminal cytotoxic profile; one displays a quiescent phenotype and the last appears as an effector-memory subset. Macrophages comprised three subclasses, two of which had a M1 phenotype and were believed to be activation states of the same cell type based on how recently they were recruited. One class, expressing foam cell (e.g., MMP9) and pro-fibrotic markers (e.g., TREM2), also expressed ACTA2, suggesting a synthetic stage with gain of smooth muscle cell characteristics.

Bao et al. (40) looked to utilize next-generation RNA sequencing to define the transcriptomic profiles of stable and unstable atherosclerotic plaques in human carotid arteries. Analysis of the sequencing reads from these groups identified 202 mRNAs, 488 long non-coding RNAs and 91 circular RNAs differentially expressed between stable and unstable plaques. Gene ontology and pathway analysis identified these transcripts to have roles in cellular stress responses across both groups. With respect to the genes corresponding to plaque stability, these were identified as having functions related to the immune response, nervous system functions, hematological activity, and endocrine system synthesis and secretion. Comparative analysis of these transcripts to proteomic data generated this group highlighted five key correlated genes in unstable plaques—CD5L, S100A12, CKB, CEMIP, and SH3GLB1. CD5L encodes for CD5, primarily expressed by macrophages, and promotes M2 polarization. S100A12 is known to bind to RAGE and activate NF-kB and ROS pro-inflammatory signaling. CKB and CEMIP are targets of two long non-coding RNAs identified in this study, which play a role in energetic hemostasis and epithelial-mesenchymal transition, respectively. Finally, SH3GLB1 has been implicated in apoptotic and autophagy pathways, but no studies have directly assessed its function in atherosclerosis.

Newman et al. (48) investigated the cellular and molecular contributions responsible for the fibrotic cap formation. Using lineage-traced $Pdgfrb^{SMC-\Delta/\Delta}$ and $Pdgfrb^{SMC-WT/WT}$ mice, initial validation studies revealed a dependence of PDGFRB signaling for SMC to contribute to plaques composition, but that these plaques did not differ in lesion stability or area compared to their control in the absence of SMCs. Next-generation RNA sequencing of brachiocephalic artery lesions from these same mice revealed that absence of PDGFRB signaling results in pathway activation specific to substrate utilization and bioenergetics to maintain plaque stability. Single-cell sequencing of medial SMCs, medial and adventitial SMCs, and lesion cells identified seven SMC specific clusters, with general trends of traditional SMC marker loss (Myh11, Acta2, etc.) across several clusters but increased osteochondrogenic markers (Sox9, Trpv4, etc.), suggestive of a metabolic reprogramming of medial and lesion SMCs during plaque development. Follow-up studies identified a large proportion of $Acta2^{+}$ cells in the fibrotic cap were derived from non-SMC origins, with markers related to EndMT and macrophage-to-myofibroblast transition (MMT).

Adventitial Tissue

Only two articles (21, 25) have explored the cellular and molecular contributions of adventitial and perivascular tissue in atherosclerosis, both of which utilize single-cell sequencing and arise from the same research group.

Gu et al. (21) utilized single-cell sequencing to elucidate the cellular composition of vascular adventitia, using both WT and $ApoE^{-/-}$ mouse models exhibiting early-stage atherosclerosis. Unsupervised clustering analysis identified 15 distinct groupings, with most of these populations representing mesenchymal cells, macrophages, T-cells, B-cells, and innate lymphoid cells (ILCs). Six clusters were characterized as non-immune, with one cluster attributed to adventitial ECs and SMCs, respectively. The remaining four groups (I–IV) were classified as mesenchymal populations, including $sca1^{+}$, $CD34^{+}$, and $Thy1^{+}$ marker-positive cells. Immune clusters comprised of two macrophage groups, identified as pro-inflammatory and resident. The resident cluster expressed the pro-atherogenic chemokine Pf4, suggesting a role in cell activation to hyperlipidemia, attracting immune cells. The follow-up study by (25) specifically investigated the $sca1^{+}$ clusters identified as adipocyte-derived stem cells from their sequencing analysis, has been addressed earlier in this review.

Cardiac Tissue

Two articles (43, 44) specifically explore the cellular and molecular constituents of cardiac tissue using single-cell sequencing, with the first (43) investigating the heterogeneity of aortic tissue in healthy vs. disease states, and the second (44) addressing the impact of hemodynamic factors, specifically disturbed flow, on the composition of carotid arteries.

Kan et al. (43) undertook single-cell sequencing on ascending aorta samples from C57BL/6J mice fed a chow and high-fat diet to determine variations in cell composition and molecular heterogeneity between healthy and diseased aortas. Analysis of these results, once clustered, identified 10 major cell types—fibroblasts, SMCs, ECs, immune cells (B-cell, T-cell, macrophages, and dendritic cells), mesothelial cells, pericytes, and neural cells. In the high-fat diet groups, ECs clusters exhibiting markers characteristic of lipid-transport and inflammation, proliferation, and leukocyte-like properties across three clusters with all showing increased contractile gene expression. SMCs were clustered into five groups, with characteristics defined as synthetic (proliferative), contractile, fibroblast-like, and inflammatory in order of prevalence. Macrophages were clustered into four distinct groups, two of which were resident and had strong proinflammatory and proliferative properties; one blood-derived cluster with pro-inflammatory and ECM degradation characteristics; and one exhibiting a lymphocyte-differentiation and immune-effector profile.

Li et al. (44) addressed the effects of disturbed blood flow at the single-cell level using mice which underwent partial carotid artery ligation. Clustering analysis showed ten distinct cell clusters specific to disturbed flow compared to control: two EC clusters, one SMC, three macrophage, and four immune cell clusters. Of the two disturbed flow specific EC clusters, the first had enrichment of the gene Dkk2 (essential to angiogenesis) and was revealed to be a flow-sensitive cell state that may derive directly from physio-normal (laminar flow) EC phenotype. The other was enriched for CD36 and showed strong association with genes related to lipid metabolism and storage. The SMC cluster

showed significant expression of Spp1, and osteoblastic marker, as well as further markers related to osteoblast differentiation, blood vessel remodeling, collagen biosynthesis, and arterial stiffness. Trajectory analysis implicated disturbed flow as directly influencing the prevalence of this phenotype in SMCs. Two macrophage clusters specific to disturbed flow had expression of markers Trem2 and Birc2, respectively. The first cluster expressed genes associated with chemotaxis and leukocyte migration, whilst the other had functions distinct to proliferation such as cell cycle DNA replication, chromatin segregation and RNA stabilization.

Blood

Two articles (33, 35) investigated blood constituents directly using microRNA sequencing, the first of which addressed the role of microRNAs in dyslipidemia and atherosclerosis risk, and the second identified the role of microRNAs in modulating macrophage heterogeneity.

Karere et al. (33) exposed two groups of baboons to a high cholesterol, high fat (HCHF) challenge diet—one group exhibiting low serum levels of LDL cholesterol and the other with high levels. Next-generation sequencing revealed 517 microRNAs from liver samples, 490 of which were identical to human microRNAs. Comparison of expression between baboon phenotypes indicated 20 novel baboon miRNAs in the low LDL cholesterol group and 29 in the high LDL group. Most of these microRNAs were not diet specific; however, there was increased quantitative expression of microRNAs in high LDL phenotype group in response to the HCHF challenge diet, suggesting diet plays some role in miRNA expression in response to dyslipidemia. Differential expression profiling revealed several unique microRNA expression profiles between the baboon phenotypes, including polycistronic regulation and polymorphism of microRNAs. Target prediction of the differentially expressed microRNAs identified 1,357 targets implicated in atherosclerosis and/or cardiovascular disease related genes, some of which exhibited contrasting or varying expression levels between the groups or in response to the HCHF challenge diet. These included LDLR and VLDLR for fatty acid metabolism, as well as ACVR1B—a receptor belonging to the TGFβ superfamily that is commonly implicated in atherosclerosis and inflammatory processes.

Li et al. (35) undertook microRNA sequencing of plasma exosomes isolated from human control subjects or subjects with significant (>50%) coronary artery stenosis. Three hundred and forty-two known microRNAs were identified, of which 14 (3 downregulated, 10 upregulated) were significantly and abundantly present in the disease samples vs. control. Determination of the targets of these differentially expressed microRNAs identified 38 differentially expressed mRNAs, and subsequent co-expression network analysis identified correlations between five monocyte/macrophage-related cell populations, six microRNAs and 10 mRNAs. RT-PCR validation of these markers showed downregulation of miR-4498 and upregulation of mRNAs CTSS, CCR2, and TREM2 in the disease group, and an inverse relationship was identified for miR-4498 vs. mRNA expression and stenosis severity. *In vitro* validation depicted macrophage uptake of circulating exosomes, and that uptake resulted in increases in the aforementioned mRNAs. Combined, this data suggests exosome-derived microRNAs exosomes may play a causal role in the polarization of macrophages to a chronic inflammatory phenotype.

Risk of Bias and Quality Assessment

Quality and risk of bias assessment was undertaken for all included articles using the ARRIVE checklist and SYRCLE risk of bias tools (**Figures 3A,B**). No studies were excluded from analysis based on these criteria. The most common limitations arising from the ARRIVE checklist was unclear information related to sample size calculation (18, 20, 22–24, 28, 30, 32, 33, 36, 38, 39, 43, 44, 46, 47, 49), blinding (18, 23–25, 31, 32, 41), randomization (20, 23, 24, 30, 31, 33, 34, 41), and specificity of inclusion/exclusion criteria (20, 23, 24, 28, 30, 31, 38, 40, 43). From the SYRCLE assessment, no information was provided regarding random outcome assessment in any of the included articles. Four articles provided unclear information about random housing, with the remaining articles providing no information (22, 26, 42, 43). Six articles had low risk of bias for sequence generation, with the remaining providing no information (19, 22, 26, 29, 42, 48). All other fields showed low or unclear bias for included studies, and factors such as blinding were all considered low due to the necessity for unblinded classification and analysis of sequencing data with respect to specific disease states, subjects, or cell types.

DISCUSSION

Atherosclerosis is well-established as one of the leading burdens to health globally. Despite many decades of research, there remains gaps in our understanding of the complex pathophysiology involved in the development and progression of this chronic inflammatory process. This can be attributed to the involvement of numerous cell types as well as a host of genetic, epigenetic, and environmental factors and the limitation of previously available methodologies for the study of the progression of the disease (1–3). Recent years have witnessed the emergence of new techniques that have contributed to our improved understanding of many complex diseases. In this study, we systematically reviewed studies that employed the techniques of single cell sequencing or/and next generation sequencing to develop better understanding of the processes involved in the development of vascular inflammation and atherosclerosis.

Characteristics of Single-Cell and Next-Generation Sequencing Techniques

Single-cell and next-generation sequencing techniques offer significant yet differing means of further elucidating the onset and progression of atherosclerosis. Single-cell sequencing offers the capacity to examine the transcriptome profiles on a cell-specific level, allowing researchers to define cellular heterogeneity within their samples and determine properties inherent to that cell type (11). Such an approach facilitates the discovery of rare cell subpopulations and cell lineage analyses, as well as providing information on regulatory networks and differential gene expression within and between these cells. This technique

FIGURE 3 | SYRCLE and ARRIVE assessments for quality and risk of bias of included studies.

involves the direct isolation of cells from a sample using microfluidics or flow activated cell sorting (FACS), a step that can present significant technical complexity in achieving the cell quality required for sequencing (52).

Next-generation sequencing, alternatively, is an overarching description for a series of advancements in genome sequencing technologies. In short, the advancements over traditional sequencing approaches include preparation of sequencing libraries without reliance on clonal amplification, and the spatially separated, amplified templates are sequenced in a massively parallel fashion such that thousands to millions of reactions can be undertaken simultaneously, offering comprehensive genome coverage in a faster and more sensitive manner (10, 53). Moreover, the sensitivity of this technology has been utilized to great effect, allowing researchers the opportunity to explore microRNAs, lncRNAs *via* rapid-amplification of cDNA ends (RACE)-sequencing, protein-DNA interactions using ChIP-Seq, and chromatin accessibility (ATAC-Seq) (54–57).

Whilst these technologies offer substantial improvements in both the accuracy and resolution of the transcriptomic data attainable, they are not without their drawbacks in their current state. Challenges facing the current utilization of single-cell sequencing methodologies have been comprehensively summarized by Lahnemann et al. (58), but in short, these challenges include: significant costs, complex computational analysis, lack of unified methodology from cell isolation to data interpretation, and high technical noise which favors highly expressed genes and masks weaker expression (59). Next-generation sequencing also presents some challenges, including a lack of sequencing localization, short read lengths (and the potential for read errors as a result) and a high dependency on reference datasets (60).

Classification in Included Studies

Single cell sequencing found utility in 23 of the 34 included articles, despite greater complexity in operation and analysis, as well as the current technical and financial barrier to entry vs. more traditional sequencing approaches (58). Despite these challenges, its use in the articles reviewed has illustrated the flexibility and novelty of this technique, with research ranging from exploratory whole tissue and plaque composition analysis (26, 43), to investigation of unique cell phenotypes and their roles in disease (20, 22). In comparison, next-generation sequencing was utilized exclusively in a smaller subset of articles (n = 12), with this technique appearing to be specifically adopted for the validation and analysis of the role of pre-defined cell subsets or genetic targets (30, 31), as well validating the effects of biological and physical factors such as lipid levels and vascular hemodynamics and their contributions to disease progression (32, 33).

Of all cell types investigated in the included articles, SMCs featured predominantly (n = 12), with attention paid predominantly to their phenotypic switching capacity in disease states (20, 27, 28, 31, 37, 41, 45–47, 50). Macrophages were another prevalent focus of research (n = 7), with specific attention paid to their phenotypic heterogeneity in plaque progression and regression (22, 24, 29, 36, 49, 51). Stem/progenitor cells were also commonly explored (n = 6), with attention paid to their presence and role in vascular adventitia, or more specifically as a source for vascular cell regeneration or repair (18, 23, 25, 30, 42). Comprehensive sequencing analysis of plaques was also frequently addressed (n = 5), with focus placed on the general composition and transcriptome of the plaque, as well as analysis of variations in expression between stable and unstable plaques (26, 34, 40, 43, 48). Little focus was placed on ECs and lymphocytes by comparison and may present an interesting area of note for future studies (32, 39).

Advancements and Future Recommendations

In addition to single-cell and next-generation sequencing approaches, another unique approach has been described recently—spatial RNA sequencing (61). This technique offers researchers the ability to identify and quantify gene expression from tissue sections, localize that expression with near cell-specific accuracy and provide visual representation of the distribution of RNA transcripts within the section. This technology works by attaching tissue cryosections to spatial transcriptomic slides that have barcoded mRNA capturing reverse transcription primers on the surface. This binding facilitates the synthesis of complementary DNA that contains their unique barcode, allowing researchers to establish the specific tissue region that it originated (62). This technology, if implemented into current atherosclerosis research, could facilitate examination of diseased vessels under various physiological or pharmacological conditions, pinpointing transcript expression to specific lesion areas to determine their compositions and contributions to disease progression.

With increasing adoption of these techniques into research practice, considerations should be made on how best to utilize these methodologies. To date, there is no unified methodological approach best suited for single-cell sequencing, from cell isolation to data interpretation. On the information provided by the articles included in this review, there is significant heterogeneity in the methods of cell isolation and quality control, both of which are key steps in ensuring sequencing reads are representative of a significant, structurally, and transcriptionally intact cell population. Sequencing analysis is also heavily dependent on known cell biomarkers to classify clusters, and variations in pre-processing to account for these may result in significant clustering in some studies, and insignificant in others. To account for these variations, future studies should carefully consider the reproducibility of their workflow, and validation with other alterative sequencing methodologies could inform the reliability of their current approach.

The findings from this work and the implementation of these modern sequencing approaches into atherosclerosis research also provides a blueprint for translatability into other adjacent avenues of research, for instance into understanding the pathophysiology of vein graft disease. To date, no implementation of any of the aforementioned sequencing approaches have been utilized in elucidating the onset and

progression of vein graft disease following coronary artery bypass surgery (CABG), despite its comparability to traditional atherosclerosis and its high prevalence and risk to a substantial patient population worldwide (63, 64). A comprehensive understanding of the pathophysiology of vein graft disease is lacking, particularly with respect to the impact on the phenotypes of vein graft cells and their interactions once implanted into a new hemodynamic environment. The use of single-cell, next-generation and spatial sequencing methodologies may be essential in improving our understanding of this disease and how preventative interventions can be developed for patients who have or will undergo CABG surgery in the future.

Limitations

Whilst this is the first systematic review of the use of these sequencing methodologies in the study of atherosclerosis, it is not without its limitations. Given the heterogeneity of the research models, methods and outcomes of the included studies, a formal meta-analysis was not undertaken. Narrowing the scope of this review may have allowed for a comparative analysis to be undertaken, such as has been recently undertaken by Zernecke et al. (65) and Conklin et al. (66) in the comparison between sequencing datasets and between sequencing and *in vitro* models, however, these analysis approaches garner data which falls outside the scope of this review. Another limitation exists when comparing single-cell sequencing data between publications. Analysis of sequencing data involves complex bioinformatics analysis to denoise data, remove low quality reads and classify cell expression into groups based on known markers, and whilst packages such as 10× Genomics CellRanger (67) exist to simplify this process, the specific criteria can vary between runs to optimize results. As such, novel cell clusters identified by one research group may be considered low quality reads by another and would thus be excluded.

Summary

Single-cell and next-generation sequencing technologies can play a significant role in illuminating the functional role of different cells and subsets of cells in physiological and pathophysiological states. The continued implementation of these technologies, particularly as they become more readily available, may present significant opportunities to explain the exact pathophysiology of atherosclerosis to optimize prevention, management, and treatment. Concurrently, there exists a significant degree of translatability for these findings and sequencing approaches to other vascular complications such as vein graft disease, given the knowledge gap that currently exists in this area, prioritization of these techniques should be considered for any future research to advance the understanding of this disease.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

LM, GM, and MZ: conceptualization and design. LM, SL, RA, GM, and MZ: analysis and interpretation. LM, SL, RA, and MZ: data collection. LM: writing the article and overall responsibility. LM, SL, RA, SG, M-SS, GA, GM, and MZ: critical revision and final approval. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by funding from the British Heart Foundation (CH/12/1/29419) to GM and SL, and the University of Leicester which provides funding matched to this BHF award to LM.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcvm.2022.849675/full#supplementary-material

REFERENCES

1. World Health Organization. *Cardiovascular Diseases (CVDs).* (2020). Available online at: https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds) (accessed November 27, 2020).
2. Benjamin EJ, Blaha MJ, Chiuve SE, Cushman M, Das SR, Deo R, et al. Heart disease and stroke statistics - 2017 update: a report from the American Heart Association. *Circulation.* (2017) 135:e146–e603. doi: 10.1161/CIR.0000000000000491
3. National Heart, Blood and Lungs Institute. *Atherosclerosis*(2020). Available online at: https://www.nhlbi.nih.gov/health-topics/atherosclerosis (accessed November 26, 2020).
4. Wentzel JJ, Chatzizisis YS, Gijsen FJ, Giannoglou GD, Feldman CL, Stone PH. Endothelial shear stress in the evolution of coronary atherosclerotic plaque and vascular remodeling: current understanding and remaining questions. *Cardiovasc Res.* (2012) 96:234–43. doi: 10.1093/cvr/cvs217
5. Tabas I, García-Cardeña G, Owens GK. Recent insights into the cellular biology of atherosclerosis. *J Cell Biol.* (2015) 209:13–22. doi: 10.1083/jcb.201412052
6. Libby P, Buring JE, Badimon L, Hansson GK, Deanfield J, Bittencourt MS, et al. Atherosclerosis. *Nat Rev Dis Primers.* (2019) 5:56. doi: 10.1038/s41572-019-0106-z
7. Liu T, Zhang L, Joo D, Sun SC. NF-κB signalling in inflammation. *Sig Transduct Target Ther.* (2017) 2:17023. doi: 10.1038/sigtrans.2017.23
8. Libby P. Inflammation in atherosclerosis. *Arteriosc Thromb Vasc Biol.* (2012) 32:2045–51. doi: 10.1161/ATVBAHA.108.179705
9. Bentzon JF, Otsuka F, Virmani R, Falk E. Mechanisms of plaque formation and rupture. *Circ Res.* (2014) 114:1852–66. doi: 10.1161/CIRCRESAHA.114.302721
10. Behjati S, Tarpey PS. What is next generation sequencing? *Arch Dis Child Educ Pract Ed.* (2013) 98:236–8. doi: 10.1136/archdischild-2013-304340
11. Wang J, Song Y. Single cell sequencing: a distinct new field. *Clin Transl Med.* (2017) 6:10. doi: 10.1186/s40169-017-0139-4
12. Moher D, Liberati A, Tetzlaff J, Altman DG, The PG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLOS Med.* (2009) 6:e1000097. doi: 10.1371/journal.pmed.1000097
13. Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred reporting items for systematic review and meta-analysis

protocols (PRISMA-P) 2015: elaboration and explanation. *Br Med J.* (2015) 349:g7647. doi: 10.1136/bmj.g7647
14. PROSPERO. *International Prospective Register of Systematic Reviews.* Available online at: https://www.crd.york.ac.uk/prospero (accessed June 3, 2021).
15. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. *Syst Rev.* (2016) 5:210. doi: 10.1186/s13643-016-0384-4
16. Percie du Sert N, Hurst V, Ahluwalia A, Alam S, Avey MT, Baker M, et al. The ARRIVE guidelines 2.0: updated guidelines for reporting animal research. *PLOS Biol.* (2020) 18:e3000410. doi: 10.1371/journal.pbio.3000410
17. Hooijmans CR, Rovers MM, de Vries RB, Leenaars M, Ritskes-Hoitinga M, Langendam MW. SYRCLE's risk of bias tool for animal studies. *BMC Med Res Methodol.* (2014) 14:43. doi: 10.1186/1471-2288-14-43
18. Tang J, Wang H, Huang X, Li F, Zhu H, Li Y, et al. Arterial Sca1(+) vascular stem cells generate *de novo* smooth muscle for artery repair and regeneration. *Cell Stem Cell.* (2020) 26:81–96.e4. doi: 10.1016/j.stem.2019.11.010
19. Sharma M, Schlegel MP, Afonso MS, Brown EJ, Rahman K, Weinstock A, et al. Regulatory T cells license macrophage pro-resolving functions during atherosclerosis regression. *Circ Res.* (2020) 127:335–53. doi: 10.1161/CIRCRESAHA.119.316461
20. Pan H, Xue C, Auerbach BJ, Fan J, Bashore AC, Cui J, et al. Single-cell genomics reveals a novel cell state during smooth muscle cell phenotypic switching and potential therapeutic targets for atherosclerosis in mouse and human. *Circulation.* (2020) 142:2060–75. doi: 10.1161/CIRCULATIONAHA.120.048378
21. Gu W, Ni Z, Tan YQ, Deng J, Zhang SJ, Lv ZC, et al. Adventitial cell atlas of wt (Wild Type) and ApoE (Apolipoprotein E)-deficient mice defined by single-cell RNA sequencing. *Arterioscler Thromb Vasc Biol.* (2019) 39:1055–71. doi: 10.1161/ATVBAHA.119.312399
22. Cochain C, Vafadarnejad E, Arampatzi P, Pelisek J, Winkels H, Ley K, et al. Single-cell RNA-seq reveals the transcriptional landscape and heterogeneity of aortic macrophages in murine atherosclerosis. *Circ Res.* (2018) 122:1661–74. doi: 10.1161/CIRCRESAHA.117.312509
23. Kokkinopoulos I, Wong MM, Potter CMF, Xie Y, Yu B, Warren DT, et al. Adventitial SCA-1(+) progenitor cell gene sequencing reveals the mechanisms of cell migration in response to hyperlipidemia. *Stem Cell Rep.* (2017) 9:681–96. doi: 10.1016/j.stemcr.2017.06.011
24. Rahman K, Vengrenyuk Y, Ramsey SA, Vila NR, Girgis NM, Liu J, et al. Inflammatory Ly6Chi monocytes and their conversion to M2 macrophages drive atherosclerosis regression. *J Clin Invest.* (2017) 127:2904–15. doi: 10.1172/JCI75005
25. Gu W, Nowak WN, Xie Y, Le Bras A, Hu Y, Deng J, et al. Single-cell RNA-sequencing and metabolomics analyses reveal the contribution of perivascular adipose tissue stem cells to vascular remodeling. *Arterioscler Thromb Vasc Biol.* (2019) 39:2049–66. doi: 10.1161/ATVBAHA.119.312732
26. Winkels H, Ehinger E, Vassallo M, Buscher K, Dinh HQ, Kobiyama K, et al. Atlas of the immune cell repertoire in mouse atherosclerosis defined by single-cell RNA-sequencing and mass cytometry. *Circ Res.* (2018) 122:1675–88. doi: 10.1161/CIRCRESAHA.117.312513
27. Wirka RC, Wagh D, Paik DT, Pjanic M, Nguyen T, Miller CL, et al. Atheroprotective roles of smooth muscle cell phenotypic modulation and the TCF21 disease gene as revealed by single-cell analysis. *Nat Med.* (2019) 25:1280–9. doi: 10.1038/s41591-019-0512-5
28. Kim JB, Zhao Q, Nguyen T, Pjanic M, Cheng P, Wirka R, et al. Environment-sensing aryl hydrocarbon receptor inhibits the chondrogenic fate of modulated smooth muscle cells in atherosclerotic lesions. *Circulation.* (2020) 142:575–90. doi: 10.1161/CIRCULATIONAHA.120.045981
29. Kim K, Shim D, Lee JS, Zaitsev K, Williams JW, Kim KW, et al. Transcriptome analysis reveals nonfoamy rather than foamy plaque macrophages are proinflammatory in atherosclerotic murine models. *Circ Res.* (2018) 123:1127–42. doi: 10.1161/CIRCRESAHA.118.312804
30. Steffen E, Mayer von Wittgenstein WBE, Hennig M, Niepmann ST, Zietzer A, Werner N, et al. Murine sca1/flk1-positive cells are not endothelial progenitor cells, but B2 lymphocytes. *Basic Res Cardiol.* (2020) 115:18. doi: 10.1007/s00395-020-0774-6
31. Mendez-Barbero N, Gutierrez-Munoz C, Madrigal-Matute J, Minguez P, Egido J, Michel JB, et al. A major role of TWEAK/Fn14 axis as a therapeutic target for post-angioplasty restenosis. *EBioMedicine.* (2019) 46:274–89. doi: 10.1016/j.ebiom.2019.07.072
32. Lai B, Li Z, He M, Wang Y, Chen L, Zhang J, et al. Atheroprone flow enhances the endothelial-to-mesenchymal transition. *Am J Physiol Heart Circ Physiol.* (2018) 315:H1293–H303. doi: 10.1152/ajpheart.00213.2018
33. Karere GM, Glenn JP, VandeBerg JL, Cox LA. Differential microRNA response to a high-cholesterol, high-fat diet in livers of low and high LDL-C baboons. *BMC Genomics.* (2012) 13:320. doi: 10.1186/1471-2164-13-320
34. Depuydt MAC, Prange KHM, Slenders L, Ord T, Elbersen D, Boltjes A, et al. Microanatomy of the human atherosclerotic plaque by single-cell transcriptomics. *Circ Res.* (2020) 127:1437–55. doi: 10.1161/CIRCRESAHA.120.316770
35. Li X, He X, Wang J, Wang D, Cong P, Zhu A, et al. The regulation of exosome-derived miRNA on heterogeneity of macrophages in atherosclerotic plaques. *Front Immunol.* (2020) 11:2175. doi: 10.3389/fimmu.2020.02175
36. Lin JD, Nishi H, Poles J, Niu X, McCauley C, Rahman K, et al. Single-cell analysis of fate-mapped macrophages reveals heterogeneity, including stem-like properties, during atherosclerosis progression and regression. *JCI Insight.* (2019) 4:e124574. doi: 10.1172/jci.insight.124574
37. Alencar GF, Owsiany KM, Karnewar S, Sukhavasi K, Mocci G, Nguyen AT, et al. Stem cell pluripotency genes Klf4 and Oct4 regulate complex SMC phenotypic changes critical in late-stage atherosclerotic lesion pathogenesis. *Circulation.* (2020) 142:2045–59. doi: 10.1161/CIRCULATIONAHA.120.046672
38. Wolf D, Gerhardt T, Winkels H, Michel NA, Pramod AB, Ghosheh Y, et al. Pathogenic autoimmunity in atherosclerosis evolves from initially protective apolipoprotein B100-reactive CD4(+) T-regulatory cells. *Circulation.* (2020) 142:1279–93. doi: 10.1161/CIRCULATIONAHA.119.042863
39. Zhou H, Mehta S, Srivastava SP, Grabinska K, Zhang X, Wong C, et al. Endothelial cell-glucocorticoid receptor interactions and regulation of Wnt signaling. *JCI Insight.* (2020) 5:e131384. doi: 10.1172/jci.insight.131384
40. Bao MH, Zhang RQ, Huang XS, Zhou J, Guo Z, Xu BF, et al. Transcriptomic and proteomic profiling of human stable and unstable carotid atherosclerotic plaques. *Front Genet.* (2021) 12:755507. doi: 10.3389/fgene.2021.755507
41. Gallina AL, Rykaczewska U, Wirka RC, Caravaca AS, Shavva VS, Youness M, et al. AMPA-Type glutamate receptors associated with vascular smooth muscle cell subpopulations in atherosclerosis and vascular injury. *Front Cardiovasc Med.* (2021) 8:655869. doi: 10.3389/fcvm.2021.655869
42. Jiang L, Chen T, Sun S, Wang R, Deng J, Lyu L, et al. Nonbone marrow CD34(+) cells are crucial for endothelial repair of injured artery. *Circ Res.* (2021) 129:e146–e65. doi: 10.1161/CIRCRESAHA.121.319494
43. Kan H, Zhang K, Mao A, Geng L, Gao M, Feng L, et al. Single-cell transcriptome analysis reveals cellular heterogeneity in the ascending aortas of normal and high-fat diet-fed mice. *Exp Mol Med.* (2021) 53:1379–89. doi: 10.1038/s12276-021-00671-2
44. Li F, Yan K, Wu L, Zheng Z, Du Y, Liu Z, et al. Single-cell RNA-seq reveals cellular heterogeneity of mouse carotid artery under disturbed flow. *Cell Death Discov.* (2021) 7:180. doi: 10.1038/s41420-021-00567-0
45. Liang X, He W, Zhang H, Luo D, Zhang Z, Liu A, et al. Inflammatory cells accelerated carotid artery calcification *via* MMP9: evidences from single-cell analysis. *Front Cardiovasc Med.* (2021) 8:766613. doi: 10.3389/fcvm.2021.766613
46. Lin CJ, Hunkins BM, Roth RA, Lin CY, Wagenseil JE, Mecham RP. Vascular smooth muscle cell subpopulations and neointimal formation in mouse models of elastin insufficiency. *Arterioscler Thromb Vasc Biol.* (2021) 41:2890–905. doi: 10.1161/ATVBAHA.120.315681
47. Quiles-Jimenez A, Gregersen I, Segers FM, Skarpengland T, Kroustallaki P, Yang K, et al. DNA glycosylase Neil3 regulates vascular smooth muscle cell biology during atherosclerosis development. *Atherosclerosis.* (2021) 324:123–32. doi: 10.1016/j.atherosclerosis.2021.02.023
48. Newman AAC, Serbulea V, Baylis RA, Shankman LS, Bradley X, Alencar GF, et al. Multiple cell types contribute to the atherosclerotic lesion fibrous cap by PDGFRbeta and bioenergetic mechanisms. *Nat Metab.* (2021) 3:166–81. doi: 10.1038/s42255-020-00338-8

49. Burger F, Baptista D, Roth A, Brandt KJ, da Silva RF, Montecucco F, et al. Single-cell RNA-seq reveals a crosstalk between hyaluronan receptor LYVE-1-Expressing macrophages and vascular smooth muscle cells. *Cells.* (2022) 11:411. doi: 10.3390/cells11030411
50. Brandt KJ, Burger F, Baptista D, Roth A, Fernandes da Silva R, Montecucco F, et al. Single-cell analysis uncovers osteoblast factor growth differentiation factor 10 as mediator of vascular smooth muscle cell phenotypic modulation associated with plaque rupture in human carotid artery disease. *Int J Mol Sci.* (2022) 23:1796. doi: 10.3390/ijms23031796
51. Li J, Wang S, Li Y, Zhang N, Gribskov M, Zhang X, et al. miRNA-mediated macrophage behaviors responding to matrix stiffness and ox-LDL. *J Cell Physiol.* (2020) 235:6139–53. doi: 10.1002/jcp.29543
52. Hwang B, Lee JH, Bang D. Single-cell RNA sequencing technologies and bioinformatics pipelines. *Exp Mol Med.* (2018) 50:1–14. doi: 10.1038/s12276-018-0071-8
53. Van Dijk EL, Auger H, Jaszczyszyn Y, Thermes C. Ten years of next-generation sequencing technology. *Trends Genet.* (2014) 30:418–426. doi: 10.1016/j.tig.2014.07.001
54. Potla P, Ali SA, Kapoor M. A bioinformatics approach to microRNA-sequencing analysis. *Osteoarthr Cartil Open.* (2021) 3:100131. doi: 10.1016/j.ocarto.2020.100131
55. Adamopoulos PG, Tsiakanikas P, Stolidi I, Scorilas A. A versatile 5' RACE-Seq methodology for the accurate identification of the 5' termini of mRNAs. *BMC Genomics.* (2022) 23:163. doi: 10.1186/s12864-022-08386-y
56. Nakato R, Sakata T. Methods for ChIP-seq analysis: a practical workflow and advanced applications. *Methods.* (2021) 187:44–53. doi: 10.1016/j.ymeth.2020.03.005
57. Buenrostro JD, Wu B, Chang HY, Greenleaf WJ. ATAC-seq: a method for assaying chromatin accessibility genome-wide. *Curr Protoc Mol Biol.* (2015) 109:21.29.1–9. doi: 10.1002/0471142727.mb2129s109
58. Lahnemann D, Koster J, Szczurek E, McCarthy DJ, Hicks SC, Robinson MD, et al. Eleven grand challenges in single-cell data science. *Genome Biol.* (2020) 21:31. doi: 10.1186/s13059-020-1926-6
59. Chen G, Ning B, Shi T. Single-cell RNA-seq technologies and related computational data analysis. *Front Genet.* (2019) 10:317. doi: 10.3389/fgene.2019.00317
60. Mantere T, Kersten S, Hoischen A. Long-read sequencing emerging in medical genetics. *Front Genet.* (2019) 10:426. doi: 10.3389/fgene.2019.00426
61. Marx V. Method of the Year: spatially resolved transcriptomics. *Nat Methods.* (2021) 18:9–14. doi: 10.1038/s41592-020-01033-y
62. Ståhl PL, Salmén F, Vickovic S, Lundmark A, Navarro JF, Magnusson J, et al. Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science.* (2016) 353:78–82. doi: 10.1126/science.aaf2403
63. Guida G, Ward AO, Bruno VD, George SJ, Caputo M, Angelini GD, et al. Saphenous vein graft disease, pathophysiology, prevention, and treatment. A review of the literature. *J Cardiac Surg.* (2020) 35:1314–21. doi: 10.1111/jocs.14542
64. Bakaeen F. CABG: a continuing evolution. *Cleve Clin J Med.* (2017) 84(12 Suppl. 4):e15–e9. doi: 10.3949/ccjm.84.s4.04
65. Zernecke A, Winkels H, Cochain C, Williams JW, Wolf D, Soehnlein O, et al. Meta-analysis of leukocyte diversity in atherosclerotic mouse aortas. *Circ Res.* (2020) 127:402–26. doi: 10.1161/CIRCRESAHA.120.316903
66. Conklin AC, Nishi H, Schlamp F, Ord T, Ounap K, Kaikkonen MU, et al. Meta-analysis of smooth muscle lineage transcriptomes in atherosclerosis and their relationships to *in vitro* models. *Immunometabolism.* (2021) 3:e210022. doi: 10.20900/immunometab20210022
67. 10× Genomics. *Single Cell Gene Expression Software: What is Cell Ranger?* (2021). Available online at: https://support.10xgenomics.com/single-cell-gene-expression/software/ pipelines/latest/what-is-cell-ranger (accessed January 14, 2021).

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Cardiovascular Medicine

TYPE Original Research
PUBLISHED 25 July 2022
DOI 10.3389/fcvm.2022.939972

OPEN ACCESS

EDITED BY
Kunwu Yu,
Huazhong University of Science and Technology, China

REVIEWED BY
Fuchun Fang,
Southern Medical University, China
Naufal Zagidullin,
Bashkir State Medical University, Russia
Ishtiaque Ahammad,
National Institute of Biotechnology, Bangladesh

*CORRESPONDENCE
Dejun Zhang
djzhang@qhu.edu.cn

SPECIALTY SECTION
This article was submitted to Cardiovascular Genetics and Systems Medicine, a section of the journal Frontiers in Cardiovascular Medicine

RECEIVED 10 May 2022
ACCEPTED 05 July 2022
PUBLISHED 25 July 2022

CITATION
Zhang Q, Guo Y, Zhang B, Liu H, Peng Y, Wang D and Zhang D (2022) Identification of hub biomarkers of myocardial infarction by single-cell sequencing, bioinformatics, and machine learning. *Front. Cardiovasc. Med.* 9:939972. doi: 10.3389/fcvm.2022.939972

Identification of hub biomarkers of myocardial infarction by single-cell sequencing, bioinformatics, and machine learning

Qunhui Zhang[1,2], Yang Guo[1,2], Benyin Zhang[2], Hairui Liu[2], Yanfeng Peng[2], Di Wang[2] and Dejun Zhang[1,2]*

[1]Research Center for High Altitude Medicine, Key Laboratory of High-Altitude Medicine (Ministry of Education), Key Laboratory of Application and Foundation for High Altitude Medicine Research in Qinghai Province (Qinghai-Utah Joint Research Key Lab for High Altitude Medicine), Qinghai University, Xining, China, [2]College of Eco-Environmental Engineering, Qinghai University, Xining, China

Background: Myocardial infarction (MI) is one of the first cardiovascular diseases endangering human health. Inflammatory response plays a significant role in the pathophysiological process of MI. Messenger RNA (mRNA) has been proven to play a key role in cardiovascular diseases. Single-cell sequencing (SCS) technology is a new technology for high-throughput sequencing analysis of genome, transcriptome, and epigenome at the single-cell level, and it also plays an important role in the diagnosis and treatment of cardiovascular diseases. Machine learning algorithms have a wide scope of utilization in biomedicine and have demonstrated superior efficiency in clinical trials. However, few studies integrate these three methods to investigate the role of mRNA in MI. The aim of this study was to screen the expression of mRNA, investigate the function of mRNA, and provide an underlying scientific basis for the diagnosis of MI.

Methods: In total, four RNA microarray datasets of MI, namely, GSE66360, GSE97320, GSE60993, and GSE48060, were downloaded from the Gene Expression Omnibus database. The function analysis was carried out by Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG), and Disease Ontology (DO) enrichment analysis. At the same time, inflammation-related genes (IRGs) were acquired from the GeneCards database. Then, 52 co-DEGs were acquired from differentially expressed genes (DEGs) in differential analysis, IRGs, and genes from SCS, and they were used to construct a protein-protein interaction (PPI) network. Two machine learning algorithms, namely, (1) least absolute shrinkage and selection operator and (2) support vector machine recursive feature elimination, were used to filter the co-DEGs. Gene set enrichment analysis (GSEA) was performed to screen the hub-modulating signaling pathways associated with the hub genes. The results were validated in GSE97320, GSE60993, and GSE48060 datasets. The

CIBERSORT algorithm was used to analyze 22 infiltrating immune cells in the MI and healthy control (CON) groups and to analyze the correlation between these immune cells. The Pymol software was used for molecular docking of hub DEGs and for potential treatment of MI drugs acquired from the COREMINE.

Results: A total of 126 DEGs were in the MI and CON groups. After screening two machine learning algorithms and key co-DEGs from a PPI network, two hub DEGs (i.e., IL1B and TLR2) were obtained. The diagnostic efficiency of IL1B, TLR2, and IL1B + TLR2 showed good discrimination in the four cohorts. GSEA showed that KEGG enriched by DEGs were mainly related to inflammation-mediated signaling pathways, and GO biological processes enriched by DEGs were linked to biological effects of various inflammatory cells. Immune analysis indicated that IL1B and TLR2 were correlated with various immune cells. Dan shen, san qi, feng mi, yuan can e, can sha, san qi ye, san qi hua, and cha shu gen were identified as the potential traditional Chinese medicine (TCM) for the treatment of MI. 7-hydroxyflavone (HF) had stable combinations with IL1B and TLR2, respectively.

Conclusion: This study identified two hub DEGs (IL1B and TLR2) and illustrated its potential role in the diagnosis of MI to enhance our knowledge of the underlying molecular mechanism. Infiltrating immune cells played an important role in MI. TCM, especially HF, was a potential drug for the treatment of MI.

KEYWORDS

myocardial infarction, gene, single-cell sequencing, machine learning, immune infiltration, drug prediction, molecular docking

Introduction

Myocardial infarction (MI), a serious cardiovascular disease caused by the rupture of unstable plaque, endangers human health worldwide. The mortality rate and the disease burden of MI are still in rapid growth, especially in China (1, 2). The early diagnosis of MI and the initiation of interventional treatment can decrease apoptotic cardiomyocytes, enhance prognosis, and reduce mortality (3). It is well known that the gold standard for diagnosing MI is evaluating the levels of markers of myocardial injury. However, these biomarkers do not quickly reflect the status of patients with MI due to their lack of sensitivity and specificity, leading patients to miss the best time for treatment (4). Consequently, it is essential to identify the underlying novel biomarkers in the diagnosis and treatment of MI to reduce premature mortality and improve prognosis.

A growing amount of evidence suggests that hereditary factors play an essential role in the development of MI progression. Genome-wide association studies (GWAS) have detected many susceptibility loci of MI (5). However, the outcomes of GWAS have failed to disclose the relevant risk of MI. Only a small fraction of these alterations can be used to explain the pathogenesis and progression of MI. Single-cell sequencing (SCS) technology is a new technology for the high-throughput sequencing analysis of genome, transcriptome, and epigenome at the single-cell level, and it also plays an important role in the diagnosis and treatment of cardiovascular diseases (6, 7). Machine learning plays an important role in screening for hub genes in gene chips.

Inflammation is associated with the occurrence of many cardiovascular diseases, especially MI (8). The inflammatory response cascade is essential in cardiac repair, remodeling, and fibrosis after MI, and non-selective suppression of inflammation after MI may detrimentally favor scarring, promote rupture, and exacerbate adverse remodeling (9). In addition, encouraging results from clinical trials of canakinumab have raised hopes for a therapy targeting inflammation in patients with MI (10, 11). What gene is involved in the regulation of MI through which inflammatory response is induced, and the specific mechanism of regulation are still unclear and need to be further explored.

Consequently, in this study, RNA microarray datasets were downloaded from the Gene Expression Omnibus (GEO)

database to identify differentially expressed genes (DEGs). Subsequently, inflammation-related genes (IRGs) were obtained from the GeneCards online platform. In addition, SCS genes were acquired from the Human Cell Landscape (HCL) online platform. Co-genes were used to construct a protein-protein interaction (PPI) network and perform GSEA. Furthermore, two machine learning algorithms were used to identify the hub genes. At the same time, the validation of hub genes was carried out in three independent MI cohorts. In addition, immuno-infiltration analysis was carried out to investigate the relationship between the immune cell and hub genes. Finally, COREMINE online platform was used to obtain the predictive drugs, and molecular docking was used to verify.

Materials and methods

Myocardial infarction studies collection

The design of this study is shown in **Figure 1**. Four RNA microarray datasets, namely, GSE66360, GSE97320, GSE60993, and GSE48060, were acquired from the National Center for Biotechnology Information's Gene Expression Omnibus (GEO) database. The data of GSE66360, GSE97320, and GSE48060 were collected from the Affymetrix Human Genome U133 Plus 2.0 Array (HG-U133_Plus_2). The data of GSE60993 was collected from the Illumina Human WG-6 v3.0 Expression Beadchip. Finally, we selected 50 MI samples and 49 healthy control (CON) samples in the GSE66360 for subsequent analysis.

Differentially expressed genes collection

The DEGs were acquired using the R package "limma" (12). Samples with $P < 0.05$ and $|\log FC| > 1.5$ were regarded as the threshold for DEGs. PCA was used to evaluate the efficiency of DEGs (13).

Single-cell data analysis

The human cell landscape online platform[1] (14) was used to investigate the expression landscape of target genes in human normal adult heart tissue from published results of SCS (**Supplementary Table 1**).

1 http://bis.zju.edu.cn/HCL/index.html

Inflammation-related genes collection

The IRGs (**Supplementary Table 2**) were acquired from the GeneCards database[2] (15).

Enrichment analysis

Subsequently, Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG), and Disease Ontology (DO) enrichment analyses were performed using the R package "clusterProfiler" (16) or Metascape[3] online platform.

Protein-protein interaction network construction

The intersection genes of DEGs, SCS, and IRGs were considered co-DEGs expressed in patients with MI. These results were used for a subsequent study. A protein-protein interaction network of co-DEGs was constructed using the STRING online platform[4] (17). According to the degree of the PPI network, the significance of intersection genes was acquired (18). The correlation among co-DEGs was evaluated using the Pearson's correlation analysis, which was calculated using Rstudio (Version 1.4.1717) (19). This result was visualized using Cytoscape (Version 3.6.1) (20).

Diagnostic hub differentially expressed genes screening

To find significant prognostic variables, two different algorithms were used to predict the status of the disease. The least absolute shrinkage and selection operator (LASSO) is a kind of regression analysis algorithm used to enhance the prediction accuracy (21). The LASSO regression analysis was conducted using the R package "glmnet" to identify the crucial genes connected with the discrimination of MI and CON (22). Support vector machine (SVM) is also a technique of supervised machine learning that is widely used for classification and regression (23). Recursive feature elimination (RFE) is an algorithm used to choose the optimal genes from samples (24). So, SVM-RFE was applied to select the superior features. The candidate diagnostic overlapping genes were acquired from two algorithms and the above PPI network.

2 https://www.genecards.org/
3 https://metascape.org/gp/index.html#/main/step1
4 https://cn.string-db.org/

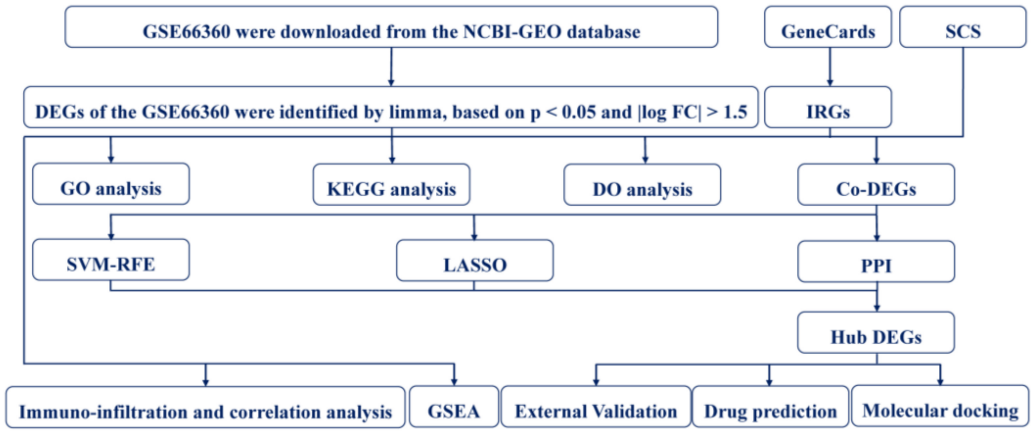


FIGURE 1
The flowchart of this study.

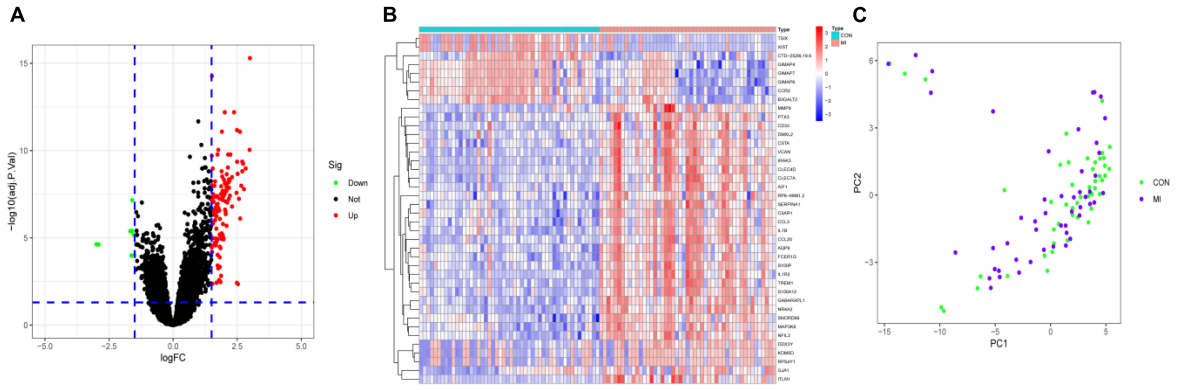


FIGURE 2
Identification of DEGs. **(A)** The volcano plot presented the DEGs between MI and CON groups. The upregulated genes are labeled in red, and the downregulated genes are labeled in green. **(B)** The heatmap visualized the alterations in the expression of DEGs. **(C)** PCA of DEGs showed good differentiation power.

Gene set enrichment analysis

The GSEA function was performed to identify the biological functions of the DEGs using the GO gene set (c5.go.v7.4.symbols.gmt) and KEGG gene set (c2.cp.kegg.v7.4.symbols.gmt). The threshold was set as $P < 0.05$. The enrichment diagrams were plotted using the R packages "clusterProfiler," "enrichplot," and "org.Hs.eg.db."

Diagnostic value of hub differentially expressed genes in myocardial infarction

To detect the predictive value of the ascertained biomarkers, we produced a receiver operator characteristic (ROC) curve using the expression data of mRNA from 50 MI and 49 CON samples. The area under the curve (AUC) from ROC was used to assess the effectiveness in the discrimination of MI and CON groups and further validated in the GSE97320, GSE60993, and GSE48060 datasets.

Immuno-infiltration analysis

To better understand the proportions of infiltrating immune cells in the MI and CON groups, the CIBERSORT[5] algorithm was used to calculate the infiltrating immune cells (25). A comparison of the infiltrated immune cell fractions in the MI and CON groups was performed using the Wilcoxon test with a significant $P < 0.05$.

5 https://cibersortx.stanford.edu/

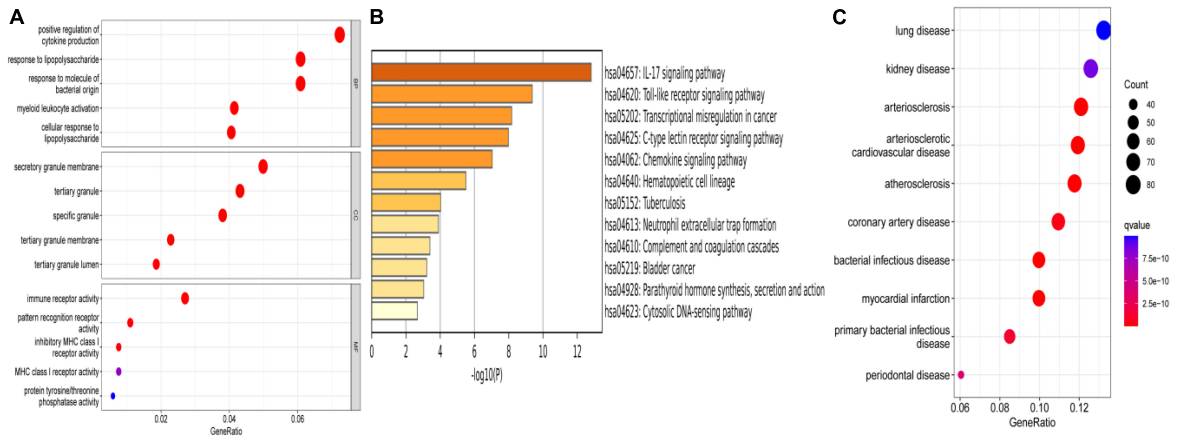


FIGURE 3
Enrichment analysis. **(A)** GO enrichment analysis of DEGs. **(B)** KEGG enrichment analysis of DEGs. **(C)** DO enrichment analysis of DEGs.

Immune correlation analysis

The correlation of infiltrating immune cells was assessed using Pearson's correlation coefficient test (26). The relationship between hub DEGs and vital infiltrated immune cells was identified by Pearson's correlation analysis of the R package "corrplot." $P < 0.05$ was regarded as a statistically significant difference.

Drug prediction

The hub DEGs were input into the COREMINE[6] online platform to acquire the underlying treatment of MI. The underlying drugs were acquired. The threshold was set as $P < 0.05$.

Molecular docking

Protein target receptors for macromolecules were acquired from the RCSB[7] online platform. AutoDock Vina (1.5.6) was used to obtain a binding model prediction. PyMOL (1.7.x) was used to visualize the results.

Statistical analysis

All statistical analysis was performed using Rstudio (Version 1.4.1717). $P < 0.05$ and $|\log FC| > 1.5$ were used as the filtering criteria for DEGs. Pearson's correlation analysis was carried out to find the gene co-expression. $P < 0.05$ was regarded as statistical significance.

Results

Identification of differentially expressed genes in patients with myocardial infarction

Differential analysis was performed to identify DEGs in patients with MI and CON. The threshold was set as $P < 0.05$ and $|\log FC| > 1.5$. A total of 126 DEGs were identified. Among them, 118 were upregulated and 8 were downregulated (**Figures 2A,B**). PCA showed that these DEGs could allow differentiation between MI and CON groups (**Figure 2C**).

Enrichment analyses

To explore the function of these DEGs, GO, KEGG, and DO enrichment analyses were conducted using the R package "clusterProfiler." GO analysis showed that the GO biological processes included positive regulation of cytokine production, response to lipopolysaccharide, and response to molecule of bacterial origin. GO cellular component contained secretory granule membrane, tertiary granule, and specific granule. GO molecular function included immune receptor activity, pattern recognition receptor activity, and inhibitory MHC class I receptor activity (**Figure 3A**). KEGG analysis showed that these DEGs took part in the IL-17 signaling pathway, toll-like receptor signaling pathway, and chemokine signaling pathway (**Figure 3B**). According to DO enrichment, these DEGs were involved in arteriosclerotic

6 https://coremine.com/medical/#search
7 https://www.rcsb.org/

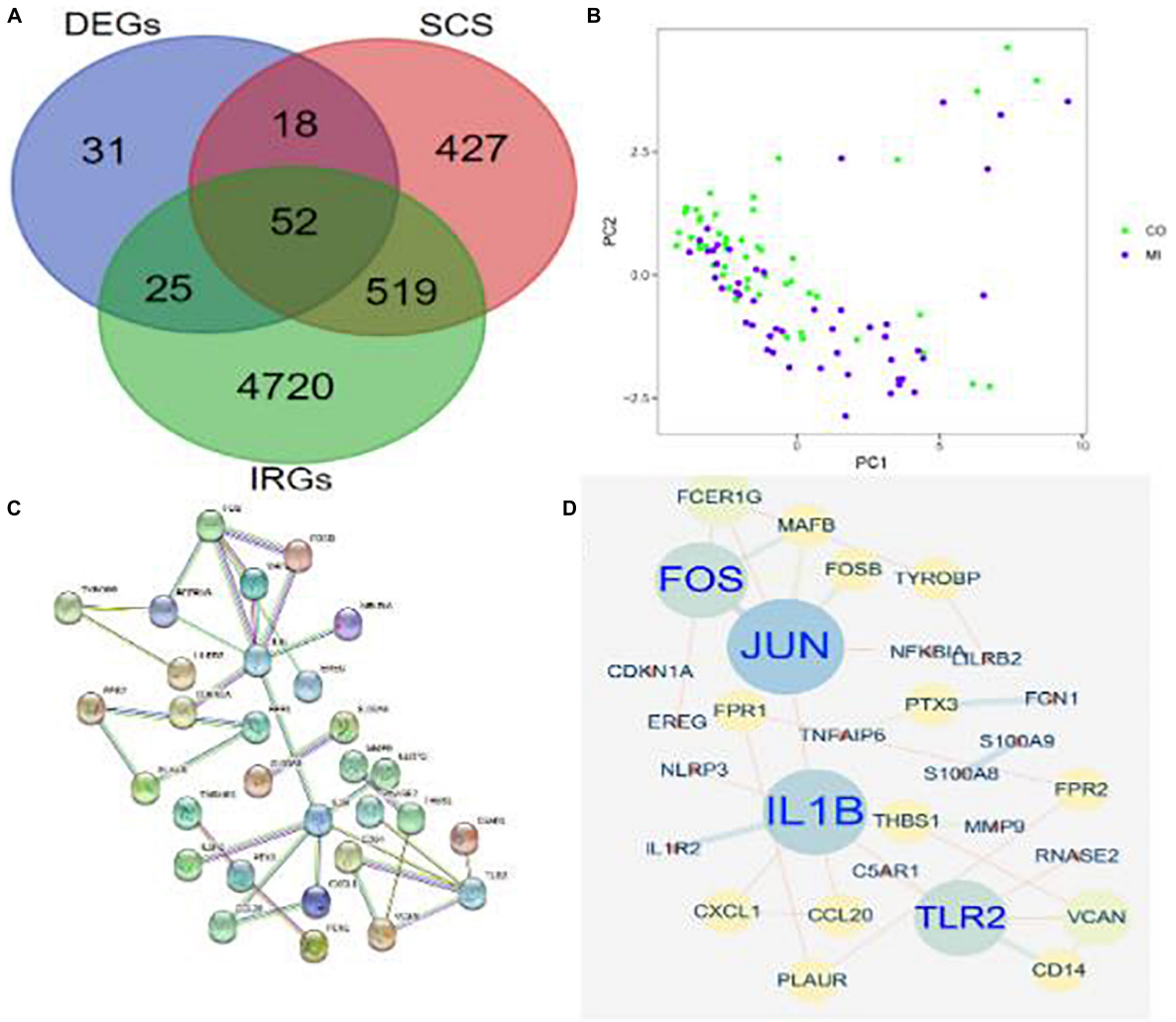


FIGURE 4
Identification of co-DEGs in patients with MI. **(A)** A Venn diagram showed the intersection of obtained DEGs, SCS, and IRGs, where 52 co-DEGs were identified. **(B)** PCA of co-DEGs showed good differentiation power. **(C)** A PPI network diagram of 52 co-DEGs was obtained from the STRING online platform. **(D)** Another PPI network diagram of 52 co-DEGs was constructed using Cytoscape 3.6.1.

cardiovascular disease, coronary heart disease, and MI (Figure 3C). These results demonstrated a tight correlation between DEGs and MI and that DEGs primarily mediated inflammatory responses.

Identification of co-differentially expressed genes and protein-protein interaction network construction

The DEGs in differential analysis, DEGs in SCS, and acquired IRGs were integrated. Subsequently, 52 co-DEGs were obtained in patients with MI (Figure 4A). In addition, PCA showed that these genes could differentiate between MI and CON groups (Figure 4B). A PPI network diagram of 52 co-DEGs was constructed using the STRING online platform with combined scores of > 0.9 points (Figure 4C and Supplementary Table 3). The former result was put into Cytoscape to construct another PPI diagram, consisting of 30 nodes and 32 edges (Figure 4D).

Functional analysis of differentially expressed genes by gene set enrichment analysis

The results of GSEA of GOBP were enriched in the acute inflammatory response, ameboidal type cell migration, and cellular response to external stimulus in the MI group and mitochondrial gene expression, mitochondrial translation, and

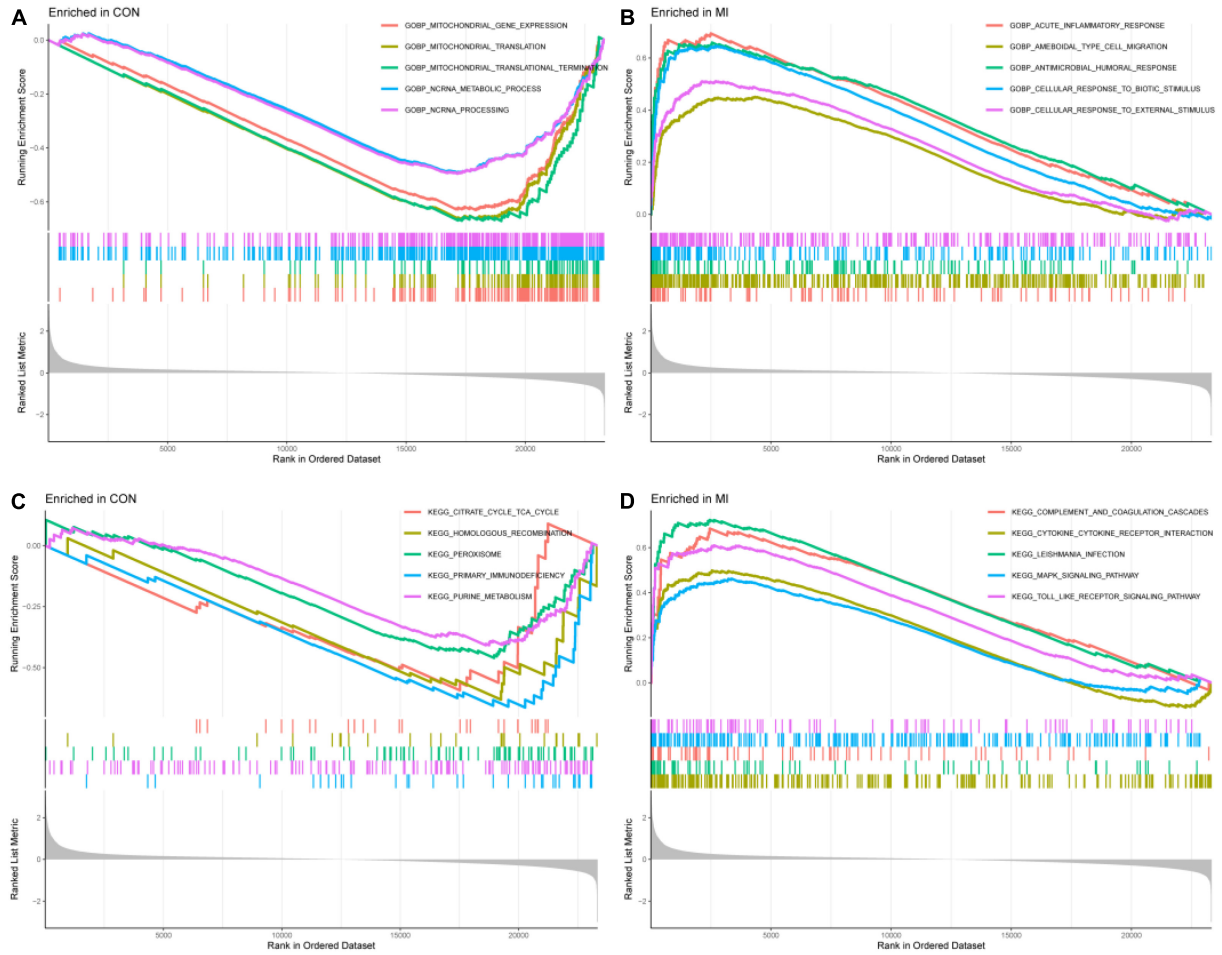


FIGURE 5
The results of GSEA. **(A,B)** The GSEA results from the GSE66360 dataset in GO analysis. **(C,D)** The GSEA results from the GSE66360 dataset in KEGG analysis.

NC RNA metabolic process in the CON group ($P < 0.05$, Figures 5A,B). The results of GSEA of KEGG were enriched in complement and coagulation cascades, MAPK signaling pathway, toll-like receptor signaling pathway, non-homologous end joining, proteasome, and protein export in the MI and CON groups ($P < 0.05$, Figures 5C,D).

Identification and validation of hub differentially expressed genes in patients with myocardial infarction

Two different algorithms (SVM-RFE and LASSO) were conducted to screen the underlying feature biomarkers, yielding 10 and 14 genes, respectively (Figures 6A,B and Supplementary Table 4). A Venn diagram showed the intersection of obtained two hub DEGs (i.e., IL1B and TLR2) of SVM-RFE, LASSO, and PPI network (Figure 6C). The expression levels of IL1B and TLR2 in the MI group were remarkably higher than those in the CON group (Figures 6D,E).

Diagnostic efficiency of IL1B, TLR2 and IL1B+TLR2 in patients with myocardial infarction

The diagnostic efficiency of the IL1B, TLR2, and IL1B + TLR2 in the discrimination of MI and CON groups indicated a good diagnostic value, with the AUC of 0.614 (95%CI: 0.499–0.722), 0.788 (95%CI: 0.693–0.877), and 0.847 (95%CI: 0.781-0.920) in the GSE66360 (Figures 7A–C). In addition, a good discrimination ability was testified in the GSE97320, GSE60993, and GSE48060 datasets. The diagnostic efficiency of the IL1B, TLR2, and IL1B + TLR2 in the discrimination of MI and CON groups indicated a good diagnostic value, with the AUC of 1.000 (95%CI: 1.000–1.000),

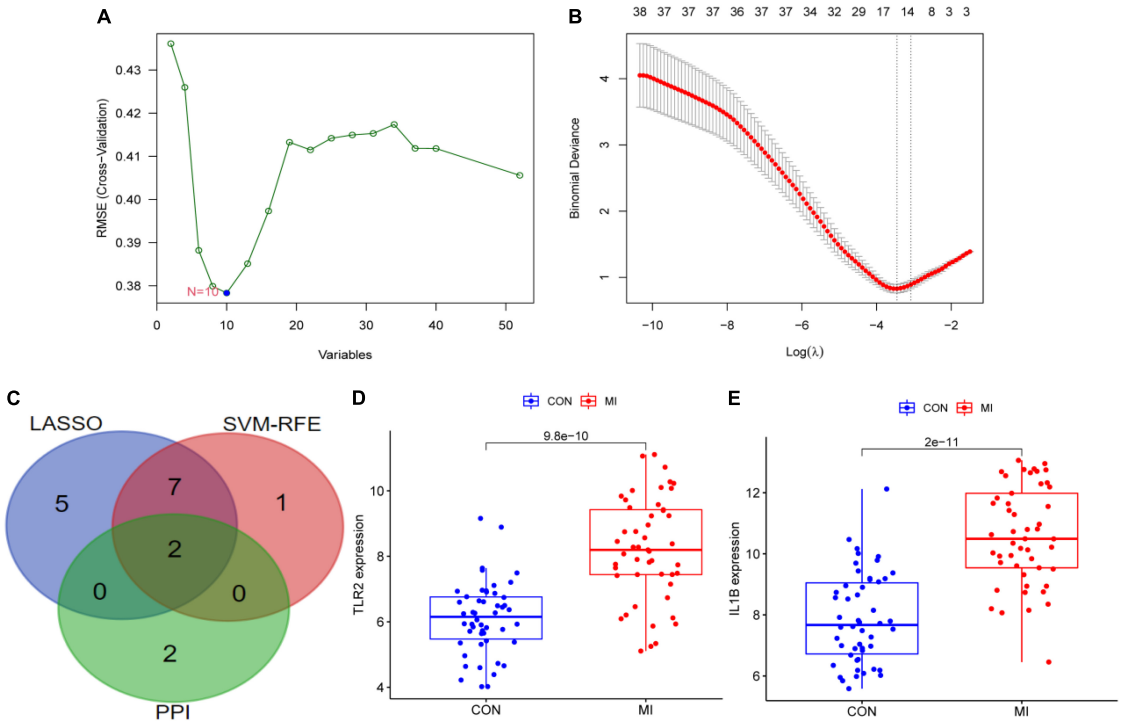


FIGURE 6
The screening of hub DEGs using machine learning and PPI. **(A)** 10 DEGs were obtained using the SVM-RFE algorithm. **(B)** 14 DEGs were obtained using the LASSO regression algorithm. **(C)** A Venn diagram presented the hub DEGs. **(D)** The expression levels of TLR2 in GSE66360. **(E)** The expression levels of IL1B in GSE66360.

1.000 (95%CI: 1.000–1.000), and 1.000 (95%CI: 1.000–1.000) in the GSE97320 (Figures 7D–F). The diagnostic efficiency of the IL1B, TLR2, and IL1B + TLR2 in the discrimination of MI and CON groups indicated a good diagnostic value, with the AUC of 0.681 (95%CI: 0.494–0.857), 0.769 (95%CI: 0.582–0.918), and 0.775 (95%CI: 0.593–0.929) in the GSE60993 (Figures 7G–I). The diagnostic efficiency of the IL1B, TLR2, and IL1B + TLR2 in the discrimination of MI and CON groups indicated a good diagnostic value, with the AUC of 0.638 (95%CI: 0.482–0.780), 0.725 (95%CI: 0.473–0.856), and 0.727 (95%CI: 0.573–0.860) in the GSE48060 (Figures 7J–L).

Immune infiltration analysis

IL1B and TLR2 played an important role in the process of MI. So, it is necessary to investigate the relationship between hub DEGs and immune cells. CIBERSORT algorithm was used to explore the immune cell infiltration. Reasonably, the results indicated that there was a marked difference in the proportions of immune cells (Figure 8A). Then, we compared the difference in immune cell infiltration between the MI and CON groups. As depicted in the picture, compared with the CON group, the proportions of "monocytes," "activated dendritic cells," "activated mast cells," "activated NK cells," "resting NK cells," "T cells follicular helper," "regulatory T cells (Tregs)," and "neutrophils" ($P < 0.001$) were highly expressed in the MI group (Figure 8B). However, the proportions of "CD 4 resting memory T cells," "gamma delta T cells," "CD8 T cells," and "resting mast cells" ($P < 0.001$) were significantly lower than that in the CON group (Figure 8B).

Immune correlation analysis

The correlation between IL1B, TLR2, and infiltrating immune cells was assessed. IL1B was positively correlated with activated mast cells, neutrophils, activated NK cells, monocytes, M2 macrophages, resting NK cells, and eosinophils ($r > 0$, $P < 0.05$), but significantly negatively correlated with resting mast cells, CD 4 resting memory T cells, M0 macrophages, gamma delta T cells, and CD8 T cells ($r < 0$, $P < 0.05$, Figure 9A). TLR2 was positively correlated with neutrophils, monocytes, activated mast cells, activated NK cells, resting NK cells, activated dendritic cells, M2 macrophages, and native B cells ($r > 0$, $P < 0.05$), but significantly negatively correlated with CD 4 resting memory T cells, gamma delta T cells, resting mast cells, and M0 macrophages ($r < 0$, $P < 0.05$, Figure 9B).

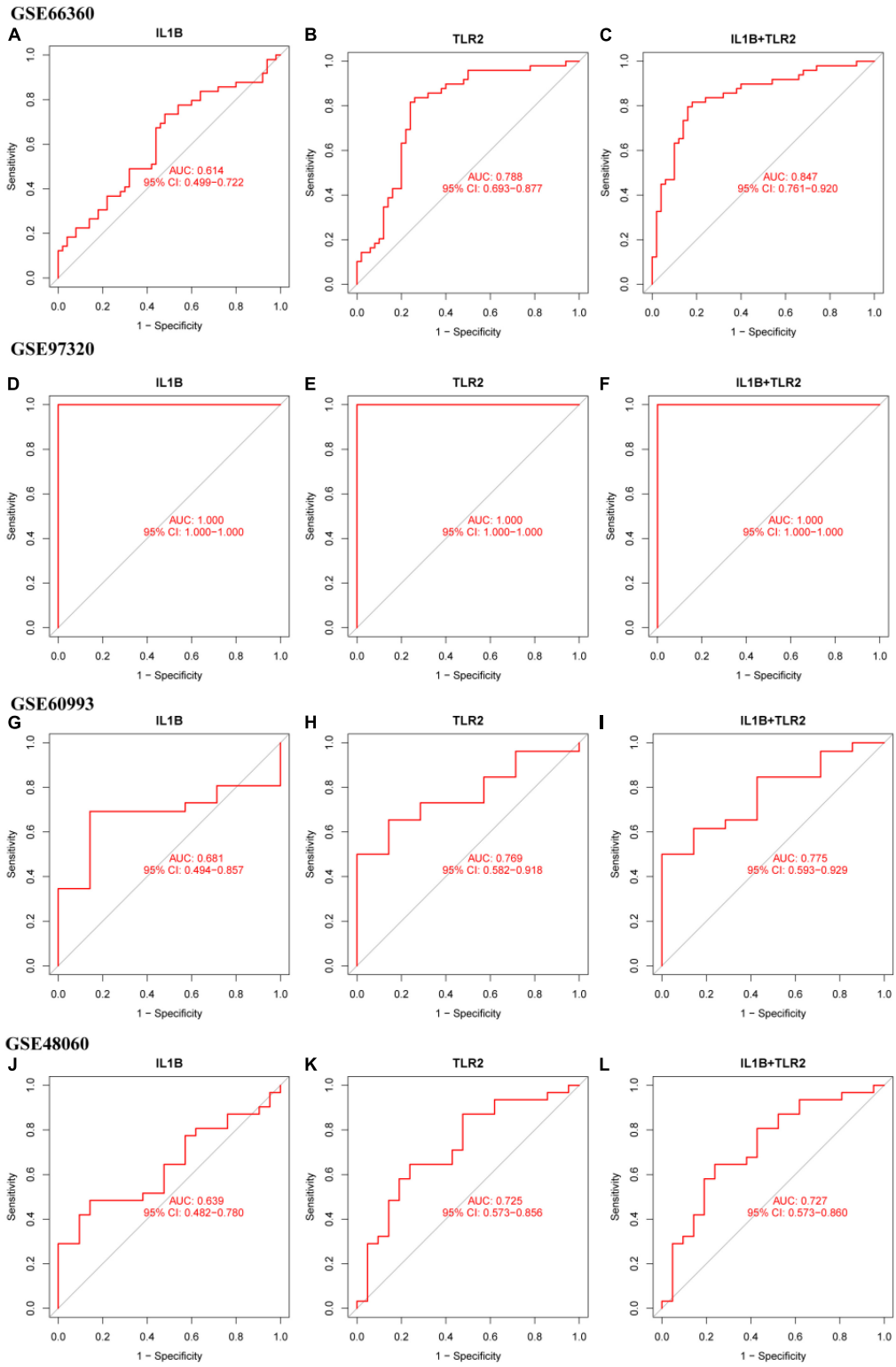


FIGURE 7
The ROC curve of the diagnostic efficiency of IL1B, TLR2, and IL1B + TLR2. **(A)** ROC curve of IL1B in the GSE66360. **(B)** ROC curve of TLR2 in the GSE66360. **(C)** ROC curve of IL1B + TLR2 in the GSE66360. **(D)** ROC curve of IL1B in the GSE97320. **(E)** ROC curve of TLR2 in the GSE97320. **(F)** ROC curve of IL1B + TLR2 in the GSE97320. **(G)** ROC curve of IL1B in the GSE60993. **(H)** ROC curve of TLR2 in the GSE60993. **(I)** ROC curve of IL1B + TLR2 in the GSE60993. **(J)** ROC curve of IL1B in the GSE48060. **(K)** ROC curve of TLR2 in the GSE48060. **(L)** ROC curve of IL1B + TLR2 in the GSE48060.

FIGURE 8
Distribution of infiltrating immune cells. **(A)** A heatmap of the infiltrating immune cells. **(B)** Comparison of 22 immune cells between MI and CON groups. Red and green colors denote CON and MI groups, respectively.

Drug prediction

The IL1B, TLR2, and MI were input into the COREMINE online platform. The threshold was set as $P < 0.05$. The results demonstrated that dan shen, san qi, feng mi, yuan can e, can sha, san qi ye, san qi hua, and cha shu gen were identified as the potential traditional Chinese medicine (TCM) for the treatment of MI.

Molecular docking

In our previous study, *Oxytropis falcata* Bunge (*O. falcata*), a Tibetan medicine/TCM, has played a role in anti-myocardial ischemia-reperfusion injury (MI/RI), and 7-hydroxyflavone (HF) was likely to be proven as a pharmacodynamic substance against MI/RI. Thus, we provided an experimental basis for its early clinical application by simulating molecular docking. The results showed that HF and IL-1B had stable combinations, and the binding affinity was −6.7 kcal/mol. Furthermore, HF and TLR2 had stable combinations, and the binding affinity was −7.4 kcal/mol.

Discussion

Myocardial infarction, a common epidemic coronary heart disease in the world, can lead to sudden major adverse cardiovascular events (MACE), causing high mortality and disability (4). Thus, patients with MI show poor clinical

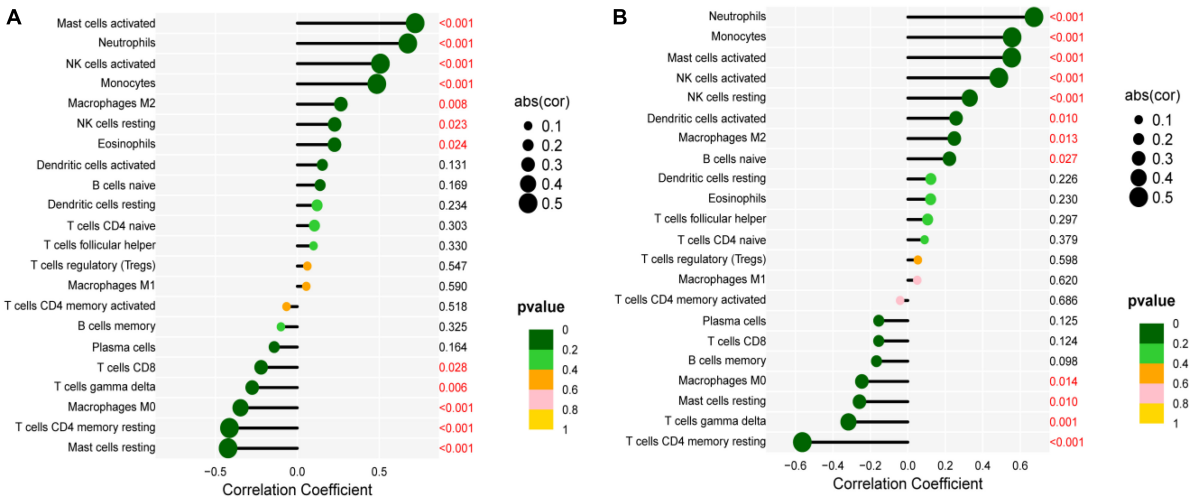


FIGURE 9
Correlation between IL1B, TLR2, and infiltrating immune cells. **(A)** Correlation between IL1B and infiltrating immune cells. **(B)** Correlation between TLR2 and infiltrating immune cells.

prognosis. Due to the lack of a valid early diagnosis, MI patients often lose effective treatment, contributing to a poor prognosis. Early diagnosis can effectively improve the prognosis and reduce mortality in MI patients (2). The boom in microarray technology offers opportunities for the treatment of MI (27). More importantly, machine learning algorithms have a wide scope of utilization in biomedicine and have demonstrated superior efficiency in clinical trials (28). Furthermore, infiltrating immune cells have been proved to play a significant role in MI (29). However, few studies have paid attention to integration with these methods to investigate the association between biomarkers and infiltrating immune cells in patients with MI. Consequently, two machine learning algorithms were used to screen candidate diagnostic biomarkers for MI and explore the role of infiltrating immune cells in MI patients.

In this study, we systematically filtered two hub DEGs specifically expressed in MI patients using two machine learning algorithms and a PPI network to build a superior MI diagnostic model. Furthermore, the diagnostic efficiency of the diagnostic model was assessed in three independent cohorts. The outcomes of enrichment analysis showed that diseases enriched by DEGs were primarily associated with MI and arteriosclerotic cardiovascular disease, KEGG enriched by DEGs were mainly related to inflammation-mediated signaling pathways, and GO biological processes enriched by DEGs were linked to biological effects of various inflammatory cells. These results are in agreement with the former study that leukocyte-mediated inflammatory responses are involved in the pathophysiological process of MI (30). It is well known that MI is caused by unstable and vulnerable plaques and is regarded as a long-term chronic inflammatory disorder (31). In the acute phase of MI, a massive inflammatory response is triggered by the sudden cessation of blood flow, leading to MI. When MI occurs, abnormal accumulation of inflammatory cells and cytokines, platelet aggregation, rupture of vulnerable plaques, and apoptosis of cardiac myocytes have been implicated in the mechanism (32). TNF has been proved to participate in inflammatory response by the combination with its specific receptors in the development of MI. This evidence is consistent with our findings, supporting that the results of this study are precise, as well as indicating that immune response plays a crucial role in MI. The inflammatory immune response plays a crucial role in MI. It is necessary to control all kinds of immune cells accurately to accomplish a much more safe and more effective treatment (33). Consequently, the acquisition of novel biomarkers of MI same as the magnitude of infiltrating immune cells and IRGs by bioinformatics and SCS will make contributions to its effective treatment.

Inflammation is associated with the occurrence of MI. Thus, based on the three machine learning algorithms, one diagnostic hub IRG was identified. Interleukin-1B (IL1B), a core proinflammatory cytokine, is a member of the interleukin 1 (IL-1) cytokine family and is involved in a variety of inflammatory responses and cellular activities, including cell proliferation, differentiation, and apoptosis (34). Recently, studies have shown that high expression of IL-1β in serum and tissues is associated with poor prognosis of patients. IL-1-mediated inflammation leads to the pathology of many diseases, especially MI (35). An article published in the *New England Journal of Medicine* has demonstrated that canakinumab, as a fully human monoclonal antibody targeting interleukin-1β, has been proven to lead to a significantly lower rate of the occurrence of MI (36). This evidence demonstrated that IL1B

plays a key role in MI, which also supported the results of this study. In addition, TLR2 played a significant role in the process of atherosclerotic lesions and MI (37). Li et al. found that the increased expression of TLR2 supported the point that proinflammatory TLR was involved in the pathogenesis of MI. TLR2 was a predictive biomarker in elderly patients with MI (38). Consequently, IL1B and TLR2 were important biomarkers of MI.

From the results of immuno-infiltration analysis, multiple immune cell subtypes were discovered to be intimately involved in the critical biological processes of MI. A growth infiltration of activated mast cells, monocytes, and neutrophils and a reduced infiltration of resting mast cells, Tregs, and gamma delta T cells were detected to be potentially associated with the occurrence and development of MI. Furthermore, immune correlation analysis demonstrated that IL1B was positively correlated with activated mast cells, monocytes, and neutrophils and negatively correlated with resting mast cells, Tregs, and gamma delta T cells. Inflammatory and immune cells, including neutrophils, lymphocytes, and macrophages, have previously been proven to play an essential role in the development of cardiovascular disease (39, 40). In the early stages of MI, cardiomyocytes undergo necrosis, with severe sterile inflammation and a significant increase in monocytes and neutrophils in the blood. This also implies the initiation of intrinsic immunity (41). Infiltrated neutrophils in the infarct ozone can release reactive oxygen species and matrix-degrading enzymes, exacerbating myocardial damage (42). During the proliferative repair phase, $CD4^+$ and $CD8^+$ T cells, Tregs, and NK T cells can enter the infarcted area and promote the maturation and differentiation of these cells. The activation of Tregs is likely to be a promising treatment for MI to enhance the ability of cardiac repair and reduce MACE. In total, these types of infiltrating immune cells play an invaluable role in AMI and should be the focal point of future research.

The limitations of this study need to be recognized. To begin with, the design of this study was a retrospective cohort study. Thus, crucial clinical information could not be acquired. Finally, the function and the immune cell infiltration of IL1B and TLR2 in MI by bioinformatics analysis, SCS, and machine learning algorithms require the designation of prospective studies to validate our outcomes in the following days.

Conclusion

This study identified two hub DEGs (i.e., IL1B and TLR2) and illustrated their potential roles in the diagnosis of MI to enhance our knowledge of the underlying molecular mechanism. Infiltrating immune cells played an important role in myocardial infarction. TCM, especially HF, was a potential drug for the treatment of MI.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**.

Ethics statement

Ethical review and approval was not required for this study with human participants, in accordance with the local legislation and institutional requirements.

Author contributions

QZ and DZ designed and wrote the original manuscript. QZ, YG, YP, and DW performed the experiments and wrote the original manuscript. DZ designed the experiments. DZ, QZ, BZ, and HL administered and coordinated the whole study project. All authors have read and agreed to the published version of the manuscript.

Funding

This study was supported by the Science and Technology Department of Qinghai Province (2020-ZJ-922).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcvm.2022.939972/full#supplementary-material

References

1. GBD 2015 Mortality and Causes of Death Collaborators. Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980-2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet.* (2016) 388:1459–544. doi: 10.1016/S0140-6736(16)31012-1

2. Zhou Z, Zhu B, Fan F, Yang F, Fang S, Wang Z, et al. Prognostic value of coronary angiography-derived fractional flow reserve immediately after stenting. *Front Cardiovasc Med.* (2022) 9:834553. doi: 10.3389/fcvm.2022.834553

3. Al-Mukhtar O, Vogrin S, Lampugnani ER, Noaman S, Dinh DT, Brennan AL, et al. Temporal changes in pollen concentration predict short-term clinical outcomes in acute coronary syndromes. *J Am Heart Assoc.* (2022) 11:e023036. doi: 10.1161/jaha.121.023036

4. Bhatt DL, Lopes RD, Harrington RA. Diagnosis and treatment of acute coronary syndromes: a review. *JAMA.* (2022) 327:662–75. doi: 10.1001/jama.2022.0358

5. Deloukas P, Kanoni S, Willenborg C, Farrall M, Assimes TL, Thompson JR, et al. Large-scale association analysis identifies new risk loci for coronary artery disease. *Nat Genet.* (2013) 45:25–33. doi: 10.1038/ng.2480

6. Ni SH, Xu JD, Sun SN, Li Y, Zhou Z, Li H, et al. Single-cell transcriptomic analyses of cardiac immune cells reveal that Rel-driven CD72-positive macrophages induce cardiomyocyte injury. *Cardiovasc Res.* (2022) 118:1303–20. doi: 10.1093/cvr/cvab193

7. Tombor LS, John D, Glaser SF, Luxán G, Forte E, Furtado M, et al. Single cell sequencing reveals endothelial plasticity with transient mesenchymal activation after myocardial infarction. *Nat Commun.* (2021) 12:681. doi: 10.1038/s41467-021-20905-1

8. Chen J, Song Y, Wang Q, Li Q, Tan H, Gao J, et al. Targeted neutrophil-mimetic liposomes promote cardiac repair by adsorbing proinflammatory cytokines and regulating the immune microenvironment. *J Nanobiotechnol.* (2022) 20:218. doi: 10.1186/s12951-022-01433-6

9. Huse C, Anstensrud AK, Michelsen AE, Ueland T, Broch K, Woxholt S, et al. Interleukin-6 inhibition in ST-elevation myocardial infarction: immune cell profile in the randomised ASSAIL-MI trial. *EBioMedicine.* (2022) 80:104013. doi: 10.1016/j.ebiom.2022.104013

10. Svensson EC, Madar A, Campbell CD, He Y, Sultan M, Healey ML, et al. TET2-Driven clonal hematopoiesis and response to canakinumab: an exploratory analysis of the CANTOS randomized clinical trial. *JAMA Cardiol.* (2022) 2022:386. doi: 10.1001/jamacardio.2022.0386

11. Hettwer J, Hinterdobler J, Miritsch B, Deutsch MA, Li X, Mauersberger C, et al. Interleukin-1β suppression dampens inflammatory leukocyte production and uptake in atherosclerosis. *Cardiovasc Res.* (2021):[Online ahead of print]. doi: 10.1093/cvr/cvab337

12. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* (2015) 43:e47. doi: 10.1093/nar/gkv007

13. Popoviæ BM, Blagojeviæ B, Ždero Pavloviæ R, Miæiæ N, Bijeliæ S, Bogdanoviæ B, et al. Comparison between polyphenol profile and bioactive response in blackthorn (Prunus spinosa L.) genotypes from north Serbia-from raw data to PCA analysis. *Food Chem.* (2020) 302:125373. doi: 10.1016/j.foodchem.2019.125373

14. Han XP, Zhou ZM, Fei LJ, Sun HY, Wang RY, Chen Y, et al. Construction of a human cell landscape at single-cell level. *Nature.* (2020) 581:303–9. doi: 10.1038/s41586-020-2157-4

15. Barshir R, Fishilevich S, Iny-Stein T, Zelig O, Mazor Y, Guan-Golan Y. GeneCaRNA: a comprehensive gene-centric database of human non-coding RNAs in the genecards suite. *J Mol Biol.* (2021) 433:166913. doi: 10.1016/j.jmb.2021.166913

16. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics.* (2012) 16:284–7. doi: 10.1089/omi.2011.0118

17. Szklarczyk D, Gable AL, Nastou KC, Lyon D, Kirsch R, Pyysalo S. The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res.* (2021) 49:D605–12. doi: 10.1093/nar/gkaa1074

18. Sriroopreddy R, Sajeed R, Raghuraman P, Sudandiradoss C. Differentially expressed gene (DEG) based protein-protein interaction (PPI) network identifies a spectrum of gene interactome, transcriptome and correlated miRNA in nondisjunction Down syndrome. *Int J Biol Macromol.* (2019) 122:1080–9. doi: 10.1016/j.ijbiomac.2018.09.056

19. Narváez-Bandera I, Suárez-Gómez D, Isaza CE, Cabrera-Ríos M. Multiple criteria optimization (MCO): a gene selection deterministic tool in RStudio. *PLoS One.* (2022) 17:e0262890. doi: 10.1371/journal.pone.0262890

20. De Marinis I, Lo Surdo P, Cesareni G, Perfetto L. SIGNORApp: a Cytoscape 3 application to access SIGNOR data. *Bioinformatics.* (2021):[Online ahead of print]. doi: 10.1093/bioinformatics/btab865

21. Colombani C, Legarra A, Fritz S, Guillaume F, Croiseau P, Ducrocq V, et al. Application of Bayesian least absolute shrinkage and selection operator (LASSO) and BayesCπ methods for genomic selection in French Holstein and Montbéliarde breeds. *J Dairy Sci.* (2013) 96:575–91. doi: 10.3168/jds.2011-5225

22. Engebretsen S, Bohlin J. Statistical predictions with glmnet. *Clin Epigenet.* (2019) 11:123. doi: 10.1186/s13148-019-0730-1

23. Ma H, Ding F, Wang Y. A novel multi-innovation gradient support vector machine regression method. *ISA Trans.* (2022):[Online ahead of print]. doi: 10.1016/j.isatra.2022.03.006

24. Han Y, Huang L, Zhou F. A dynamic recursive feature elimination framework (dRFE) to further refine a set of OMIC biomarkers. *Bioinformatics* (2021):[Online ahead of print]. doi: 10.1093/bioinformatics/btab055

25. Kawada JI, Takeuchi S, Imai H, Okumura T, Horiba K, Suzuki T, et al. Immune cell infiltration landscapes in pediatric acute myocarditis analyzed by CIBERSORT. *J Cardiol.* (2021) 77:174–8. doi: 10.1016/j.jjcc.2020.08.004

26. Kuo CD, Chen GY, Wang YY, Hung MJ, Yang JL. Characterization and quantification of the return map of RR intervals by Pearson coefficient in patients with acute myocardial infarction. *Auton Neurosci.* (2003) 105:145–52. doi: 10.1016/s1566-0702(03)00049-3

27. Huang GS, Hong MY. Genomic expression for rat model of damp obstruction in Chinese medicine: application of microarray technology. *Am J Chin Med.* (2005) 33:459–74. doi: 10.1142/s0192415x05003065

28. Zhang Y, Xia R, Lv M, Li Z, Jin L, Chen X, et al. Machine-learning algorithm-based prediction of diagnostic gene biomarkers related to immune infiltration in patients with chronic obstructive pulmonary disease. *Front Immunol.* (2022) 13:740513. doi: 10.3389/fimmu.2022.740513

29. Shi HT, Huang ZH, Xu TZ, Sun AJ, Ge JB. New diagnostic and therapeutic strategies for myocardial infarction via nanomaterials. *EBioMedicine.* (2022) 78:103968. doi: 10.1016/j.ebiom.2022.103968

30. De Rosa S, Eposito F, Carella C, Strangio A, Ammirati G, Sabatino J, et al. Transcoronary concentration gradients of circulating microRNAs in heart failure. *Eur J Heart Fail.* (2018) 20:1000–10. doi: 10.1002/ejhf.1119

31. Sun J, Hartvigsen K, Chou MY, Zhang Y, Sukhova GK, Zhang J, et al. Deficiency of antigen-presenting cell invariant chain reduces atherosclerosis in mice. *Circulation.* (2010) 122:808–20. doi: 10.1161/circulationaha.109.891887

32. Xiao J, Moon M, Yan L, Nian M, Zhang Y, Liu C, et al. Cellular FLICE-inhibitory protein protects against cardiac remodelling after myocardial infarction. *Basic Res Cardiol.* (2012) 107:239. doi: 10.1007/s00395-011-0239-z

33. Tobin SW, Alibhai FJ, Weisel RD, Li RK. Considering cause and effect of immune cell aging on cardiac repair after myocardial infarction. *Cells.* (2020) 9:9081894. doi: 10.3390/cells9081894

34. Zaaber I, Mestiri S, Hammedi H, Marmouch H, Mahjoub S, Tensaout BB, et al. Association of interleukin-1B and interleukin-4 gene variants with autoimmune thyroid diseases in Tunisian population. *Immunol Invest.* (2016) 45:284–97. doi: 10.3109/08820139.2016.1153650

35. Stegger JG, Schmidt EB, Tjønneland A, Kopp TI, Sørensen TI, Vogel U, et al. Single nucleotide polymorphisms in IL1B and the risk of acute coronary syndrome: a Danish case-cohort study. *PLoS One.* (2012) 7:e36829. doi: 10.1371/journal.pone.0036829

36. Ridker PM, Everett BM, Thuren T, MacFadyen JG, Chang WH, Ballantyne C, et al. Antiinflammatory therapy with canakinumab for atherosclerotic disease. *N Engl J Med.* (2017) 377:1119–31. doi: 10.1056/NEJMoa1707914

37. Schoneveld AH, Oude Nijhuis MM, van Middelaar B, Laman JD, de Kleijn DP, Pasterkamp G. Toll-like receptor 2 stimulation induces intimal hyperplasia and atherosclerotic lesion development. *Cardiovasc Res.* (2005) 66:162–9. doi: 10.1016/j.cardiores.2004.12.016

38. Li X, Guo D, Chen Y, Hu Y. Toll-Like receptors/TNF-α pathway crosstalk and impact on different sites of recurrent myocardial infarction in elderly patients. *Biomed Res Int.* (2022) 2022:1280350. doi: 10.1155/2022/1280350

39. Kimura Y, Tsukui D, Kono H. Uric Acid in inflammation and the pathogenesis of atherosclerosis. *Int J Mol Sci.* (2021) 22:222212394. doi: 10.3390/ijms222212394

40. Barale C, Melchionda E, Morotti A, Russo I. PCSK9 biology and its role in atherothrombosis. *Int J Mol Sci.* (2021) 22:22115880. doi: 10.3390/ijms22115880

41. Weil BR, Neelamegham S. Selectins and immune cells in acute myocardial infarction and post-infarction ventricular remodeling: pathophysiology and novel treatments. *Front Immunol.* (2019) 10:300. doi: 10.3389/fimmu.2019.00300

42. Carbone F, Nencioni A, Mach F, Vuilleumier N, Montecucco F. Pathophysiological role of neutrophils in acute myocardial infarction. *Thromb Haemost.* (2013) 110:501–14. doi: 10.1160/th13-03-0211

frontiers | Frontiers in Cardiovascular Medicine

TYPE Review
PUBLISHED 26 July 2022
DOI 10.3389/fcvm.2022.905151

OPEN ACCESS

EDITED BY
Michel Puceat,
Institut National de la Santé et de la Recherche Médicale (INSERM), France

REVIEWED BY
June-Wha Rhee,
City of Hope National Medical Center, United States
Feiyang Ma,
University of Michigan, United States

*CORRESPONDENCE
Zhuo Zhao
zhaozhuo1230@163.com

SPECIALTY SECTION
This article was submitted to Cardiovascular Genetics and Systems Medicine, a section of the journal Frontiers in Cardiovascular Medicine

RECEIVED 04 April 2022
ACCEPTED 05 July 2022
PUBLISHED 26 July 2022

CITATION
Hu Y, Zhang Y, Liu Y, Gao Y, San T, Li X, Song S, Yan B and Zhao Z (2022) Advances in application of single-cell RNA sequencing in cardiovascular research.
Front. Cardiovasc. Med. 9:905151.
doi: 10.3389/fcvm.2022.905151

Advances in application of single-cell RNA sequencing in cardiovascular research

Yue Hu[1], Ying Zhang[2], Yutong Liu[1], Yan Gao[3], Tiantian San[1], Xiaoying Li[3,4], Sensen Song[2], Binglong Yan[2] and Zhuo Zhao[1,2]*

[1]Department of Cardiology, Jinan Central Hospital, Shandong University, Jinan, China, [2]Department of Cardiology, Central Hospital Affiliated Shandong First Medical University, Jinan, China, [3]Department of Research Center of Translational Medicine, Central Hospital Affiliated Shandong First Medical University, Jinan, China, [4]Department of Emergency, Central Hospital Affiliated Shandong First Medical University, Jinan, China

Single-cell RNA sequencing (scRNA-seq) provides high-resolution information on transcriptomic changes at the single-cell level, which is of great significance for distinguishing cell subtypes, identifying stem cell differentiation processes, and identifying targets for disease treatment. In recent years, emerging single-cell RNA sequencing technologies have been used to make breakthroughs regarding decoding developmental trajectories, phenotypic transitions, and cellular interactions in the cardiovascular system, providing new insights into cardiovascular disease. This paper reviews the technical processes of single-cell RNA sequencing and the latest progress based on single-cell RNA sequencing in the field of cardiovascular system research, compares single-cell RNA sequencing with other single-cell technologies, and summarizes the extended applications and advantages and disadvantages of single-cell RNA sequencing. Finally, the prospects for applying single-cell RNA sequencing in the field of cardiovascular research are discussed.

KEYWORDS

single-cell RNA sequencing, cardiovascular, heart development, stem cells, precision medicine

Introduction

Cardiovascular disease is considered to be the leading cause of human mortality worldwide. The latest data show that cardiovascular disease morbidity and mortality are gradually increasing (1). The pathogenesis related to the cardiovascular system has not been fully elucidated due to the diverse cellular interactions and complex regulatory mechanisms. Although there has been great progress in the diagnosis and treatment of cardiovascular disease owing to the development of anatomy, genomics, and other disciplines and the development of inspection technology (2), understanding the cellular heterogeneity and gene interactions in cardiac tissue development and cardiac cell differentiation trajectories and in disease states requires further research. The limitations of action hinder the in-depth study of the pathogenesis and the effectiveness of diagnosis and treatment plans.

Therefore, how to use more powerful tools to research cardiovascular diseases and thereby promote cardiovascular health is still an important problem for researchers and clinicians to solve.

To achieve the goal of accurate disease prevention, diagnosis, and treatment, precision medicine not only involves interpreting big data about the living environment and habits of people but also strives to develop the fields of genomics and proteomics to reveal key differences related to the occurrence and development of diseases. The emergence and rapid development of single-cell technologies and omics analyses have facilitated a major leap in cardiovascular precision medicine. With the completion of the Human Genome Project (3), single-cell sequencing technologies have developed rapidly and have brought about progress in experimental biological research. Among the technologies, single-cell RNA sequencing (scRNA-seq) is of great significance in cardiac precision medicine. It captures the transcriptional profiles of thousands of single cells from complex tissues in high resolution. It thereby helps researchers to identify and target abnormal pathways and genes in cardiovascular disease and also assists clinicians in formulating effective treatment strategies (4). The research application of scRNA-seq is very extensive. In addition to revealing cardiovascular cell diversity, identifying stem cell differentiation trajectories, and clarifying cell-to-cell communication (5, 6), scRNA-seq has been used to characterize cardiac development and genetics (7, 8), which has promoted advances in cardiovascular research. This article reviews the application of scRNA-seq in relation to the cardiovascular system and provides an overview of future research strategies and directions.

Single-cell transcriptome sequencing

Since Tang et al. established the initial scRNA-seq process in 2009 (9), a variety of single-cell transcriptomics platforms have emerged one after another, with different degrees of optimization regarding single-cell capture methods, amplification methods, and transcript coverage. The development and improvement of scRNA-seq technologies will help to fully understand the changes at the gene expression level in both physiological and pathological conditions, thereby improving the precision of disease diagnosis and treatment. As shown in Figure 1, the basic scRNA-seq process involves: (1) sample preparation, (2) single-cell capture, (3) amplification and library preparation, and (4) sequencing and analysis (10).

Sample preparation

Single-cell sequencing is based on the use of mechanical dissociation and enzymatic digestion to isolate single-cell samples from tissues. Different tissues or cell types, cell culture conditions, and cell viability require different tissue processing and digestion methods. For dense tissues such as the heart, the maximum number of cardiomyocytes is usually obtained by a combination of sectioning and enzymatic digestion (11). Langendorff method and Enzymatic bulk digestion are two major methods for isolation of cardiomyocytes. Enzyme digestion of heart tissue usually does not yield a sufficient number of viable cells due to insufficient tissue exposure (12). Langendorff method depends on tissue structure and has a complex operation process (13). So, considering the limitations of the two methods, researchers have made efforts to isolate single-cell samples from heart tissues. Based on Langendorff Method, Ackers-Johnson et al. (14) presented a novel method to isolate viable cardiac myocytes and non-myocytes from the adult mouse heart with only standard surgical tools and equipment, without the prerequisite of heparinization. Similarly, Guo et al. (11) developed a simplified method to isolate human cardiomyocytes from tissue slicing. It involves three steps: slice the tissue, $Ca2^{+}$-free perfusion for 9 min, and enzymatic digestion for 45–90 min. This method can not only ensure the survival rate of cardiomyocytes, but also the integrity of morphology and morbid metabolic characteristics, the physiological function of cardiomyocytes. For soft tissues such as lymph nodes, single-cell suspensions can be obtained directly by mechanical separation (15). Researchers should optimize sample preparation procedures to maximize the number of single cells and viability while minimizing cell mortality.

Single-cell capture

The single-cell capture method varies with the cell volume. Mainstream single-cell capture methods include laser capture, fluorescence-activated cell sorting (FACS), microfluidic capture, and microdroplet-based capture (16–18). FACS uses the principle of light scattering and fluorescence signals to sort cells into 96- or 384-well plates. Cell samples can be stored for a long time, increasing the flexibility of the research on the cells and reducing damage to the cells, but the plate-based limitations of FACS restrict the number of cells that can be sorted. The Fluidigm C1 system is one of the main technologies that use microfluidic cell capture. Its advantages are that it uses a smaller volume of cell suspension and reduces the risk of external contamination of cells (10). However, it requires a large number of live cells and cell size uniformity is limited in terms of the scale of analysis, and is expensive (10). Microdroplet-based platforms such as Drop-seq and Chromium (10x Genomics), which use DNA barcoding technology to analyze individual cells encapsulated in oil droplets, can rapidly analyze thousands of individual cells, greatly reducing the amount of time required for each analysis and cost. However, it still has the disadvantages of a low mRNA capture rate and low gene detection efficiency of

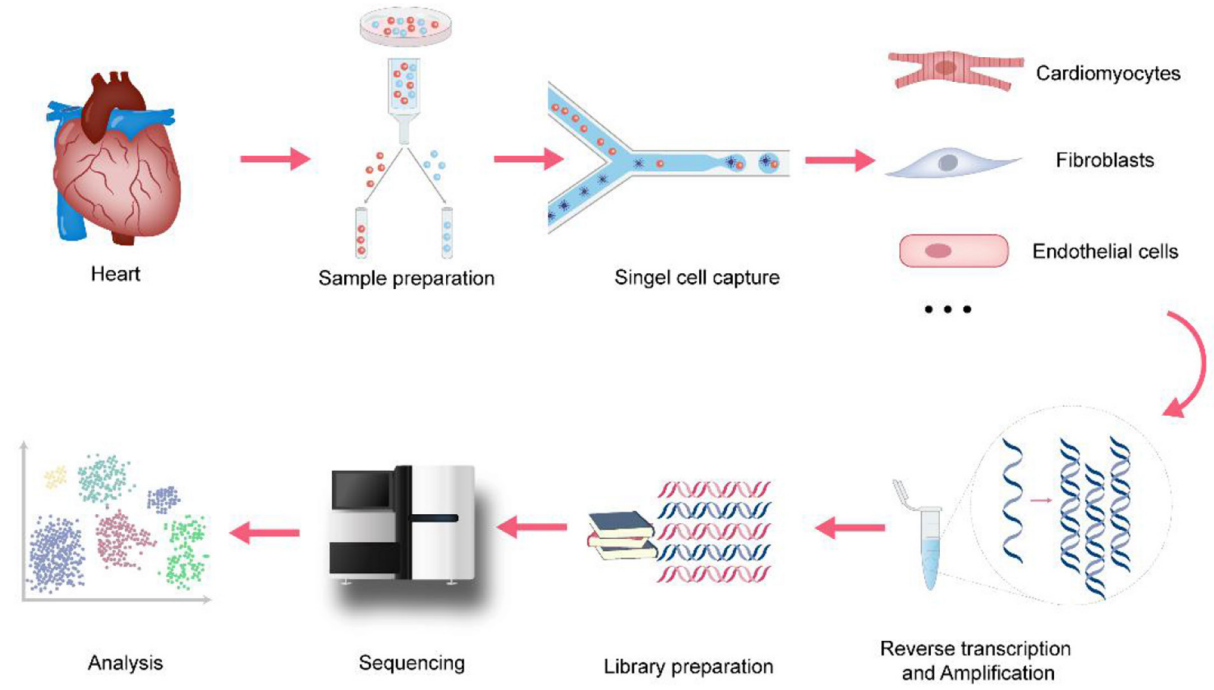


FIGURE 1
The process of single-cell RNA sequencing in cardiovascular research. In cardiovascular research, the basic scRNA-seq process involves: (1) sample preparation, Langendorff method and Enzymatic bulk digestion are two major methods for isolation of cardiomyocytes (2) single-cell capture, snRNA-seq holds advantages when applied to heart tissues because of the size of cardiomyocytes (3) amplification and library preparation, amplification methods are divided into two categories: single-cell whole-genome amplification (WGA) and single-cell whole transcriptome amplification (WTA) (4) sequencing, it involves single-cell genome sequencing, or single-cell transcriptome sequencing and single-cell epigenetic sequencing (5) analysis, cell clustering analysis, Differential expression analysis, Pseudotime analysis, Gene Ontology analysis and RNA velocity are the major analysis methods.

cardiomyocytes due to their large cell size (4). However, Seq-wall and SPLiT-seq can effectively solve the problem. In Seq-wall, barcoded mRNA capture beads and single cells are sealed in an array of subnanoliter wells using a semipermeable membrane, enabling efficient cell lysis and transcript capture (19). So this technology is applicable to almost any cellular suspension for which a reference transcriptome exists. Similarly, In SPLiT-seq, individual transcriptomes are uniquely labeled by passing a suspension of formaldehyde-fixed cells or nuclei through four rounds of combinatorial barcoding. It eliminated the need for single-cell isolation because of the index information of DNA barcodes (20). These two methods have unique technical advantages to achieve a high mRNA capture rate of cardiomyocytes and are inexpensive, portable, and efficient. In addition, researchers have widely used Single-nucleus RNA-seq (snRNA-seq) in cardiovascular research (21). snRNA-seq holds several advantages when applied to heart tissues. First, it can be used to study frozen and archived primary tissues without the preparation of single-cell suspension. Second, it will minimize the bias of cell capture of platform-specific that has a certain size optimal and reduce the impact of aberrant transcriptional changes induced by enzymatic dissociation.

Amplification and library preparation

Single-cell sequencing technology avoids the shortcomings of traditional technology that takes cell populations as the research object and ignores the heterogeneity between cells. Using a single cell as the research object, single-cell sequencing technology lyses a single cell and can amplify ultra-trace amounts hundreds of thousands of times, and establishes a sequencing-level cDNA library for sequencing and data analysis (22).

To avoid amplification bias, researchers should choose the most optimized method based on sensitivity, precision, and accuracy. The commonly used amplification methods are mainly divided into two categories: single-cell whole-genome amplification (WGA) and single-cell whole transcriptome amplification (WTA). The classification of these methods is shown in Table 1. The difference between WTA and WGA is that WTA first reverse-transcribes RNA into cDNA, and then amplifies the cDNA. Among the types of WTA, CEL-seq uses a unique molecular identifier (UMI) in the cDNA synthesis process and linearly amplifies mRNA, thereby reducing amplification bias (32), and it is more cost-effective

TABLE 1 Comparison of single-cell sequencing technology gene amplification methods.

Amplification type	Amplification method	Advantages	Limitations	References
WGA	DOP-PCR	CNV detection in a larger genome.	Low genome coverage and high error rate for SNV detection	(23)
	MDA	high genome coverage	Large bias and susceptible to contamination	(23)
	MLBAC	Small sequence-dependent bias, high CNV detection accuracy, and low SNV false negative rate	Low fidelity and high SNV false positive rate	(24)
	eWGA	Good amplification uniformity, strong sensitivity for both CNV and SNV, and low contamination rate	To be studied	(25)
	LIANTI	High gene coverage, low allele loss, and good amplification uniformity	Less accurate for very small CNVs	(26)
	SISSOR	High sequencing accuracy for undivided cells	to be studied	(27)
	PicoPLEX	Low amplification error rate, sensitive for CNV, good repeatability and amplification uniformity	To be studied	(28)
WTA	CEL-seq	High reproducibility and sensitivity and short amplification times	Low Specificity For mRNA amplification	(29)
	Smart-seq/Smart-seq2	Low amplification bias, high coverage, low variability, and low noise	No analytical ability for polyA ribonucleic acid	(30)
	Drop-seq	Low cost and fast amplification	Low mRNA capture rate	(31)

WGA, single cell whole gene group amplification; WTA, single-cell transcriptome amplification; CNV, copy number variation; SNV, single-nucleotide variation; DOP-PCR, degenerate oligonucleotide primed PCR; MDA, multiple strand displacement amplification; MLBAC, multiple annealing circular cyclic amplification; eWGA, emulsion whole genome amplification; LIANTI, transposon insertion linear amplification; SISSOR, microfluidic reactor single strand sequencing; CEL-seq, Cell Expression by Linear amplification and Sequencing; Smart-seq, Switching mechanism at 5' end of the RNA transcript.

for transcriptome quantification of a large range of expression levels. Another type of WTA, Smart-seq2, uses Moloney murine leukemia virus (MMLV) reverse transcriptase, which preferentially selects full-length cDNA as a substrate for its terminal transferase activity, allowing the recovery of full-length cDNAs. It thereby improves the reverse transcription process and increases the yield and length of single-cell cDNA libraries. It achieves high sensitivity and low amplification bias. Additionally, it can analyze all exons of each single-cell transcript, detect different splice variants, and has a wider range of applications (30).

Sequencing and analysis

At present, the third-generation single-cell transcriptome sequencing technology has been developed. The representative technologies are HeliScope single-molecule sequencing (33), single-molecule real-time sequencing (34), Oxford nanopore sequencing (35). Compared with the second-generation sequencing technology, the third-generation sequencing technology uses useful information such as nanotechnology and modern optics to capture the base sequence, which has the advantages of fast, real-time sequencing and longer base sequence reading. However, there is still room for improvement in the cost and accuracy of sequencing (36).

In addition to the single-cell genome or transcriptome sequencing, single-cell sequencing methods also include single-cell epigenetic sequencing. The innovation regarding single-cell epigenetic sequencing is that the methylation level of the whole genome can be obtained, which is of great significance for the study of epigenetics as it allows the spatiotemporal specificity of epigenetic changes to be determined.

Currently, multi-omics sequencing technologies (37, 38), including G&T-seq (39), TARGET-seq (40), scCAT-seq (41), scM&T-seq (42), and PLAYR (43) technologies, are booming. Compared to single-omics sequencing technologies, multi-omics sequencing technologies can provide more systematic insights into understanding biological heterogeneity and diversity. Multi-omics sequencing analysis that involves both spatial data and time trajectory data avoids some of the shortcomings of single-omics sequencing and enriches the

experimental results. Multi-omics sequencing will certainly be a potential method in the field of sequencing research in the future.

Regarding the analysis of single-cell sequencing data, cell clustering analysis is a key type of data analysis. It is the basis of further analysis of the data (44), such as differential expression analysis, pseudotime analysis, and Gene Ontology analysis. And the downstream analysis such as RNA velocity is also important. A comparison of these methods is shown in Table 2. In addition, the multiplexing method increases the number of samples for scRNA-seq, facilitates library construction, and lowers the reagent costs by relying on DNA-based barcoding that enables the pooling of all barcoded samples into a single mixed sample for analysis. It will provide great insight into high-throughput perturbation screening and tracking the dynamic process of cell differentiation (55). Researchers need to select a data analysis method according to their research purposes and conditions.

Application of single-cell transcriptome sequencing in the cardiovascular system

Cellular heterogeneity in the developing and mature heart

With the development of scRNA-seq, it has become possible to reveal new cell types and cell subpopulation heterogeneity in the heart, and to describe the spatial and temporal expression patterns of different cell types. For example, by performing scRNA-seq on NKX2.5$^+$ and Isl1$^+$ mouse cardiac progenitor cells (CPCs) and developmental trajectory analysis, Jia et al. (56) revealed that Isl1$^+$ CPCs undergo an attractor state before entering various developmental branches, while extended expression of NKX2.5 may promote the unidirectional development of CPCs into cardiomyocytes. This demonstrated that Nkx2.5 expression plays a decisive role in the differentiation of pluripotent CPCs into cardiomyocytes. Similarly, Xiao et al. (57) performed scRNA-seq on mouse embryonic hearts and found that *Lats1* and *Lats2* (in the Hippo signaling pathway) caused subepicardial cells to differentiate into cardiac fibroblasts by inhibiting the YAP target gene *Dhrs3*. After conditional knockout of *Lats1* and *Lats2*, aberrant signaling from C20, a transitional cell subset between the epicardium and cardiac fibroblasts, led to a disordered coronary configuration. *Lats1* and *Lats2* also play important roles in controlling extracellular matrix composition and vascular remodeling while promoting the transition of epicardial progenitors to differentiated cardiac fibroblasts. Moreover, by performing scRNA-seq on cardiac differentiation from human embryonic stem cells and human embryonic/fetal hearts. Sahara et al. (58) identified a unique cell subset that emerges specifically in the proximal outflow tract of human embryonic hearts, marked by LGR5. And the LGR5$^+$ cells promote cardiogenesis through expansion of the ISL1$^+$TNNT2$^+$ intermediates. This study will help to gain insight into human cardiac development and reveal the possible origin of congenital heart disease. Thus, scRNA-seq is important for characterizing the heterogeneity of developing and mature cardiac cells and for identifying the molecular mechanisms of cardiac development by allowing individual cardiac cells to be studied.

Induced pluripotent stem cells

iPSCs have been widely used in research on the molecular mechanisms of cardiovascular-related diseases. However, the endothelial cells and cardiomyocytes derived from iPSCs have differences in physiological structure and gene expression compared to mature cardiac cells, which is a barrier to the application of iPSCs in clinical treatment (59). scRNA-seq can capture the regulatory processes of stem cell differentiation and development with unprecedented sensitivity, providing technical support and a theoretical basis for stem cell transplantation (60). Paik et al. (61) performed large-scale droplet-based scRNA-seq on thousands of human iPSC-derived endothelial cells. They found that *CLDN5*, *APLNR*, *GJA5*, and *ESM1* enrichment characterized four cell subsets, and they elucidated the unique physiological role of each cell subset, providing a foundation for sorting IPSC with specific biological function and identity. In addition, by performing scRNA-seq on human iPSC-derived cardiomyocytes, Churko et al. (62) revealed that *NR2F2* and *HEY2* promote atrial and ventricular-like effects, respectively. The fact that these genes have different regulatory roles that lead to different functional cell states was validated in mouse models (63). They also found that notch signaling may be regulating the atrial vs. ventricular development within different heart chambers. These findings will help to promote the clinical translation of iPSCs research into cardiovascular disease treatments.

Cardiovascular disease

Congenital heart disease

Compared to traditional sequencing technology, scRNA-seq provides more insight into congenital heart disease caused by genetic changes at the single-cell level. By studying autoimmune congenital heart block (CHB) fetal hearts and healthy fetal hearts, Suryawanshi et al. (64) revealed that overexpression of interferon-stimulated genes and activation of interferon signaling were involved in the occurrence of CHB. They also found that stromal cells may contribute to stromal deposition and fibrosis, leading to CHB. Likewise, by using scRNA-seq to study mouse cardiac progenitors, de Soysa et al. (65) demonstrated that *Hand2* is a transcriptional determinant

TABLE 2 Comparison of single-cell sequencing data analysis methods.

	Principles of data analysis	Core algorithm	Advantage	References
Differential expression analysis	Aggregation analysis in which the effects of different samples or treatment methods on gene expression levels are compared	SCDE, MAST Census, BCseq	Identifies cell-specific markers and distinguishes among various cell subsets	(45–47)
Pseudotime analysis	Dynamic pathways of cell development or differentiation are inferred based on gene expression patterns in single cells	Monocle	Identifies key genes in cell differentiation in a non-purified state	(48–51)
Gene Ontology analysis	Controlled word sets, which comprehensively describe the properties of genes and gene products, are identified	Gene Ontology	Determines accurate descriptions of cells and molecular functions	(52)
Cell clustering analysis	Define cell types through unsupervised clustering on the basis of transcriptome similarity	k-means	Identifies putative cell types in any samples	(53)
RNA velocity	Recovers directed information by distinguishing unspliced and spliced mRNAs	velocyto, scVelo	Grants access to the descriptive state of a cell, and its direction and speed of movement in transcriptome space	(54)

of individual CPC fate and differentiation direction. They found that *Hand2*-knockout mice have abnormal right ventricle formation due to defective ventricular outflow tract development, resulting in congenital heart disease phenotypes. In addition, scRNA-seq also provides a higher discriminative power regarding genetic alterations in specific cell types that trigger congenital heart disease. Hu et al. (21) compared the single-cell transcriptome sequencing results of nuclei from the hearts of mice with mitochondrial cardiomyopathy and normal mouse hearts. They found that the growth factor *GDF15* was up-regulated in the former, revealing the cardiac cell type-specific gene regulatory network of *GDF15*. These studies demonstrate that the application of scRNA-seq not only paints a new picture of cardiac development from a completely new perspective, but also provides new information on genes and mechanisms for further elucidation of congenital heart disease.

Heart disease in adults

There is a growing interest in conducting research on cardiac disease pathogenesis and novel therapeutic targets, and scRNA-seq is revolutionizing our understanding of common cardiac diseases. Using a pressure overload-induced heart failure model, scRNA-seq provides insights into cardiac cellular heterogeneity, laying the foundation for therapeutic strategies for the disease. Nomura et al. (66) constructed trajectories related to cardiomyocyte remodeling by subjecting individual cardiomyocytes to scRNA-seq, elucidating the genetic programs underlying the morphological and functional characteristics of cardiac hypertrophy and failure. In the early stage of cardiac hypertrophy, cardiomyocytes activate mitochondrial metabolic genes, which are related to cardiomyocyte size and connect to the *ERK1/2* and *NRF1/2* transcriptional networks to initiate myocardial remodeling. In the late stage of cardiac hypertrophy, the p53 signaling pathway is activated, followed by cardiomyocyte morphological elongation and heart failure. Furthermore, scRNA-seq has been used to track cell fate transitions in the early, developing, and convalescent phases of myocardial infarction, providing new insights into the pathogenesis of myocardial infarction (67). In the early phases of myocardial infarction, fibroblasts are activated, creating a collagen scar that reduces further cardiac tissue rupture. But in the convalescent phases of myocardial infarction, fibroblasts will eventually lead to heart failure (68). Therefore, studying activated fibroblasts helps us to better understand their role in the pathogenesis of myocardial infarction. By performing scRNA-seq on collagen 1α1 green fluorescent protein $(GFP)^{+}$ fibroblasts after myocardial infarction, Ruiz-Villalba et al. (69) found that activated fibroblasts up-regulate collagen triple helix repeat containing 1 (*Cthrc1*) in the early stage of myocardial infarction and participate in collagen synthesis through TGF-β signaling, to promote healing and prevent heart rupture. *Cthrc1* was down-regulated during the myocardial infarction recovery period to avoid pathological ventricular remodeling.

TABLE 3 Research on single-cell sequencing of aortic cells in atherosclerosis (AS).

Species	Cell type	Platform	Main findings	References
Mouse	$CD45^+$ leukocytes	Fluidigm C1	Depiction of immune cell composition and transformation trends within atherosclerotic arterial vessels	(73)
Mouse	Th1-like $IFN\gamma^+CCR5^+$ Treg subset (Th1/Tregs), T regulatory (Treg) cells, and Th1 cells	Fluidigm C1	AS involves Treg plasticity, accumulation of interferon $gamma^+$ Th1/Tregs, Treg subpopulation dysfunction, and further promotes arterial inflammation	(74)
Mouse/human	Smooth muscle cells	10x Genomics	TCF21 regulates the transition of smooth muscle cells to fibromuscular cells in AS, and the latter protects against AS by infiltrating lesions	(75)
Mouse	Adventitial cells	10x Genomics	Descriptive cellular atlas of heterogeneous cell populations in the adventitia, revealing dynamic interactions between adventitial macrophages and stroma in AS	(76)
Mouse	Macrophages	10x Genomics	Non-foaming macrophages promote inflammation in AS	(77)
Human	smooth muscle cells	ICELL8	Histone H2A variant H2A.Z was down-regulated in AS smooth muscle cells, and its overexpression inhibited VSMC dedifferentiation and neointima formation caused by injury, and played a protective role in AS	(78)

Vascular lesions

Atherosclerosis (AS) is a chronic inflammatory disease associated with vascular endothelial cell dysfunction, myeloid cell enrichment, lipid deposition and foam cell formation (70–72). scRNA-seq has advanced the understanding of AS as a multifactorial disease involving diverse cellular interactions (see Table 3 for details). In addition, the application of scRNA-seq to abdominal aortic aneurysm (AAA) research and the development of spatial transcriptomics have led to a deeper understanding of the cellular heterogeneity involved in AAA and the functional status of the various cells. Yang et al. (79) performed scRNA-seq (based on the 10x Genomics) platform on aortic cells in the early stage of AAA. They observed the number of macrophages was increased. And Mϕ-2, an inflammatory macrophage, was found to be the main enriched macrophage type in the early stage of the disease. Thus, scRNA-seq revealed the types of smooth muscle cells and macrophages involved in the development of AAA, and predicted the functional status of cells, in order to improve research on AAA pathogenesis, targeted drug development, and personalized vascular medicine.

Cardiac precision medicine

The purpose of precision medicine is to deliver individualized treatment in order to improve health. This involves constructing information repositories and infrastructure to improve the efficiency of clinical research, and to provide more precise information in clinical settings (80). In cardiovascular system applications, scRNA-seq has enabled the identification of disease-causing gene expression and mutations at the single-cell level, and it has provided unprecedented insights into cell developmental trajectories and differentiation directions. Hulin et al. (81) used droplet-based transcriptome sequencing to show that the gene expression and function of endothelial and immune cell subsets remained relatively constant during aortic and mitral valve development in neonatal d7 (primitive values) and d30 (mature valves) mice, while the interstitial cell subsets exhibited significant changes. This study was the first to reveal the cellular diversity during heart valve remodeling, opening up new avenues for studying heart valve homeostasis and the molecular mechanisms of valve diseases. Due to its unique technical advantages, scRNA-seq has played an important role in revealing cell-to-cell heterogeneity during disease development and drug treatment. Its role in identifying and thereby targeting complex cardiovascular pathological phenotypic pathways has also promoted the development of cardiac precision medicine, providing effective therapeutic strategies for cardiovascular diseases. However, scRNA-seq still has shortcomings such as excessive raw data noise and low gene coverage (82). It is believed that with continuous scRNA-seq innovations related to the current technical barriers, its contribution to cardiac precision medicine will become more significant.

Coronavirus disease 2019 associated heart injury

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) binds to the ACE2 receptor and transmembrane

protease serine 2 (TMPRSS2) cleaves the viral S protein (83). This activates the process of entry into host cells to allow the virus to infect these cells, which eventually leads to novel coronavirus-associated pneumonia (COVID-19) and cardiac damage (84). ACE2 and TMPRSS2 double positivity is a key condition for SARS-CoV-2 to enter cells. Liu et al. (85) compared the scRNA-seq data of fetal and adult human heart tissue with lung tissue data and showed that ACE2 was relatively up-regulated in heart tissue compared to lung tissue, but due to the low TMPRSS2 expression, the heart damage was minor. Thus, a lower percentage of $ACE2^{+}$ $TMPRSS2^{+}$ cells somewhat reduce susceptibility to SARS-CoV-2-induced cardiac damage. Nevertheless, up-regulated cathepsin L (CTSL) and paired basic amino acid-cleaving enzyme (furin) in the heart may compensate for TMPRSS2, mediating SARS-CoV-2 infection of the heart (85). Sequencing single adult coronary artery cells showed that compared to endothelial cells from the lung, ACE2 was up-regulated in cardiac endothelial cells, but TMPRSS2 was down-regulated, so the endothelial cells may act as a barrier that protects the cardiac tissue from circulating SARS-CoV-2 (85). The above studies show that single-cell sequencing, as a powerful technical means, can help to further understand the cardiovascular damage caused by SARS-CoV-2 and its underlying mechanisms, thereby promoting the identification of treatment targets and reducing mortality.

Expansion of the application of scRNA-seq

Cell–cell interactions

Recognition and signal transmission between cells is carried out through receptor-ligand binding. Accurate CCIs based on this process are an important prerequisite for organisms to maintain complex life activities (86). Therefore, the study of the CCI process is beneficial to advancing the interpretation of cell biological functions, metabolic states, disease pathological processes, and other aspects.

Deciphering CCIs based on gene expression using scRNA-seq has been greatly developed. Compared to traditional analytical methods, it has greater advantages in quantifying gene expression in rare cell types and identifying cellular sources of proteins that mediate CCIs. For example, Skelly et al. (87) constructed a framework of the extensive networks of intercellular communication in the heart by performing scRNA-seq on mouse non-cardiomyocyte cells. Similarly, Wang et al. (88) compared the interactions of human cardiomyocyte cells and non-cardiomyocytes in healthy and failing states based on scRNA-seq, which revealed the regulation of cardiomyocyte behavior by non-cardiomyocyte cells.

Trajectory analysis

Traditional lineage tracing methods involve tracking the expression levels of specific genetic markers in progeny cells (89). Due to the limited number of genetic marker genes available, it is impossible to accurately track all progeny cell types, so information on cell heterogeneity is incomplete. However, with the development of scRNA-seq technologies and computational methods, trajectory analysis based on scRNA-seq data makes it possible to trace cell development and differentiation lineages without using clear genetic markers (90). Among the methods, pseudo-time trajectory analysis based on the Monocle algorithm is the most representative (91). The principle is to infer the trajectory based on the similarity of expression patterns between the sequenced cells, and then to sort the single cells in one-dimensional space to indicate the trajectory. In this way, information on the complex dynamic differentiation process of cells can be obtained (92).

However, pseudo-time trajectory analysis cannot completely replace the traditional lineage tracing methods due to the disadvantages of non-linear intermediate product interference, unstable sequencing depth, and obvious batch effects. Nevertheless, the calculation methods involved in pseudo-time trajectory analysis are still being continuously improved. Currently, Monocle3, which is an improved version of Monocle, is being used for trajectory analysis of limbal basal epithelial cells (93) and analysis of tumor cell genetic mutations (94). With the continuous advancement of technology, it is expected that developmental trajectory analysis will be combined with spatial transcriptomics analysis and data on cell and molecular properties. These methods will then play greater roles in revealing the gene expression signatures of different cells and developmental stages and in constructing embryonic developmental maps of unexplored species.

Comparison with other single-cell technologies

Single-nucleus RNA sequencing

Compared to scRNA-seq, snRNA-seq has the advantages of less tissue dissociation bias, high compatibility with frozen samples, and elimination of tissue dissociation-induced transcriptional stress responses because it is not limited by the dissociation conditions (95). It has been widely used in many organs including the heart. By performing snRNA-seq on adult mammalian hearts, Wolfien et al. (96) found that mature cardiomyocytes are not of a single origin and identified a cardiomyocyte subpopulation of endothelium-oriented origin with dual roles. Likewise, Galow et al. (97) found that the turnover of the cardiomyocyte pool is generated by cytokinesis of resident cardiomyocytes rather

than being driven by the differentiation of progenitor cells. These studies contribute to the understanding of cardiac cell biology, including cardiomyocyte regeneration. In addition, by performing snRNA-seq on left ventricle samples from hearts with dilated cardiomyopathy and hypertrophic cardiomyopathy as well as non-failing hearts, Chaffin et al. (98) found that in cardiomyopathic hearts, the expression of proliferating resident cardiac macrophages was reduced and that of activated fibroblasts was increased. *PRELP* and *COL22A1* played a plausible role in cardiac fibrosis by encoding the extracellular matrix protein prolargin in fibroblasts. Their research will expand our understanding of the transcriptional and molecular basis of cardiomyopathy.

However, snRNA-seq has strict nuclear extraction conditions and complex procedures. In addition, the level of mRNA in the nucleus is relatively low, it is estimated that only 50% of RNA is present in the nucleus vs. cytoplasm, so snRNA-seq is not highly sensitive for all cells and genes (99). Thrupp et al. (100) compared the performance of snRNA-seq and scRNA-seq for analyzing human cortical microglia and found that snRNA-seq was less sensitive at identifying some genes (such as *APOE* and *CST3*) related to Alzheimer's disease, showing that snRNA-seq provides insufficient insights into some genes in some human tissues and cells. Additionally, Slyper et al. (99) showed that snRNA-seq was less powerful than scRNA-seq for the analysis of immune cells in fresh and frozen human tumor samples. Therefore, researchers should combine snRNA-seq data with cell phenotyping and proteomics to gain accurate and comprehensive information.

In addition, the combination of snRNA-seq and single nucleus assay for transposase-accessible chromatin sequencing (snATAC-seq) has been used in experimental research. By analyzing transcriptomic and epigenomic interactive multimodal atlas by Combining snRNA-seq and snATAC-seq, Muto et al. (101) highlighted functional heterogeneity in the proximal tubule and thick ascending limb. Similarly, Thomas et al. (102) identified transcription-factor-binding motifs and cis-regulatory elements in the human retina and induced pluripotent stem cell-derived retinal organoids through the use of snATAC-seq and snRNA-seq. This combination will provide a unique opportunity to utilize tissues that have been already obtained and help researchers get more information.

Mass cytometry (cytometry by time-of-flight mass spectrometry)

As post-translational modification is a key process in gene regulation and signaling pathway activation, there is a lack of one-to-one correspondence between mRNA and protein levels in cells. Therefore, the application of single-cell technologies in the field of proteomics helps to gain accurate information on disease pathogenesis. Among the technologies, mass cytometry (CyTOF) is a multi-parameter single-cell technology. Owing to its high accuracy in single-cell analysis and wide range of measurement parameters, it can accurately analyze intracellular signaling networks and immunophenotypes, making it a powerful single-cell proteomics tool (103). For example, by performing CyTOF and scRNA-seq on carotid plaques and T cells of AS patients, Fernandez et al. (104) found that $CD4^+$ T cell subsets differed between AS patients and healthy individuals, with some T cell subsets showing signs of exhaustion. Identification of macrophage subsets associated with carotid plaque vulnerability has furthered research on cardiovascular immunotherapies (104). Similarly, Chen et al. (105) performed CyTOF on the aortic smooth muscle cells of a mouse model of the aneurysm and found that *KLF4* reduces TGF-β signaling and reprograms contractile smooth muscle cells into mesenchymal stem cells, thus promoting the generation and progression of aneurysms.

"Spatial transcriptomics" technologies

scRNA-seq has unique advantages regarding providing information on cellular heterogeneity and disease pathogenesis, but there are still some limitations. The extraction of single cells precludes collecting detailed information on their specific location in tissues and this makes scRNA-seq less effective for functional interpretation of gene expression in specific physiological or pathological microenvironments (106). To obtain both cellular location and gene expression data, in 2016 Ståhl et al. (107) developed spatial transcriptomics technologies to capture mRNAs along with location data in tissues (using unique location barcodes) in order to visualize gene expression distribution in tissues.

In recent years, the development of spatial transcriptomics technologies has been rapid. Among them, tissue-based labeling technology represents the latest technological progress in spatial transcriptomics. The ZipSeq technology invented by Hu et al. (108) is a representative living tissue labeling technology. It uses a light control system to map DNA barcodes to living cells to identify real-time gene expression levels in specific cells. It is beneficial to combine spatial information with surface epitope profiling, and thereby promote understanding of the interconnections between cellular locations and transcriptional heterogeneity. Spatial Transcriptomics frequently provides technical support for research on tissue and cell spatial heterogeneity and gene expression across many research fields such as developmental differentiation lineage mapping, cell functional state interpretation, and disease model exploration. It has become a core technical facility in the field of biological research.

Omics analysis

Precision medicine will be improved by combining multi-omics approaches related to the epigenome, transcriptome, proteome, and metabolome. Epigenomics analysis, including the analysis of DNA methylation, histone acetylation, and phosphorylation, is an important component of the multi-omics analysis (109). Asare et al. (110) used bone marrow reconstitution experiments in mice with hyperlipidemia and epigenomics analysis to show that histone deacetylase 9 (HDAC9) binds to inhibitory kappa B kinase (IKK), resulting in IKK deacetylation and activation, and stimulating the inflammatory response of macrophages and endothelial cells. The research revealed that HDAC9 inhibition therapy delays AS progression.

Current research directions in the field of multi-omics involve reducing the experimental cost and combining more omics techniques together. In addition, metabolomic analysis of metabolic changes, which are located downstream of the genome—closer to the phenotype of the organism—can lead to a better assessment of the physiological changes that occur in diseased states (111). Since the appearance of metabolomic in 1999 (112), great progress has been made in understanding the pathological mechanisms underlying cardiovascular disease. For example, Murashige et al. (113) conducted a metabolomic study of arterial, coronary, and femoral venous blood from 110 heart failure and non-heart failure patients and found that fatty acids are the main energy source of the heart in both groups, but the hearts of the heart failure patients exhibited higher ketone and lactate consumption with a high proteolysis rate. The findings contribute to the construction of a framework describing human cardiac energy sources and provide a basis for understanding cardiometabolic abnormalities in disease states. Unlike scRNA-seq, metabolomics analysis can easily integrate upstream genetic data, transcriptional and proteomic variation, and cellular microenvironment information, so it can better reflect the molecular processes involved in disease states (114). At present, combined single-cell multi-omics analysis technologies are developing rapidly. Multi-omics technologies for analyzing the transcriptome, open chromatin, and histone modification has been gradually developed and has become an important tool for promoting the development of precision medicine (115).

Discussion

The use of scRNA-seq in cardiovascular research has been greatly promoted by the innovation of single-cell capture methods, the development of gene amplification methods, the emergence of multiple sequencing platforms, and the diversification of data analysis methods. scRNA-seq has played an important role in exploring cardiac cell heterogeneity during heart development, improving stem cell models, and revealing the molecular mechanisms of diseases and potential therapeutic targets. In particular, it has been useful in research on the mechanisms of COVID-19-related cardiac injury, which may lead to rapid identification of therapeutic targets and reductions in mortality.

Combining single-cell sequencing with other technologies would also provide more useful information to researchers. Ranzoni et al. (116) applied scRNA-seq and scATAC-seq to human immunophenotypic blood cells from fetal liver and bone marrow, they identified transcriptional and functional differences between hematopoietic stem cells from liver and bone marrow. Similarly, Moncada et al. (117) performed microarray-based spatial transcriptomics with scRNA-Seq on pancreatic tumors, they identified cell type subpopulations across tissue regions and cell state relationships in the tumor microenvironment. In addition, Zhao et al. (118) applied pooled CRISPR delivery and single-cell transcriptome analysis to the βcell line MIN6, they confirmed that the lncRNA-enriched cluster of MIN6 was associated with insulin transcription.

However, scRNA-seq still has certain limitations. First, single-cell sorting platforms have limitations in capturing cardiomyocytes due to their large size. Second, compared to snRNA-seq, scRNA-seq has limitations such as low compatibility with frozen samples, inducing strong transcriptional stress responses that affect the results, and the inability to provide nuclear gene expression information. In addition, compared to CyTOF, scRNA-seq cannot take into account proteomics data alongside gene expression data, and it provides poor insights into pathological states caused by abnormal post-translational modification of genes. An enzyme-tethering strategy called Cleavage Under Targets and Tagmentation (CUT&Tag) has emerged (119). In CUT&Tag, a specific antibody binds to the target chromatin protein, which then tethers a protein A-Tn5 transposase. After activation of the transposase and DNA purification because of the addition of Mg^{2+}, the genomic fragments with adapters at both ends are enriched by PCR. The steps from Sample preparation to amplification and library preparation can be performed in a single tube on the benchtop in 1 day. This strategy provides high-resolution and low background sequencing libraries for profiling diverse chromatin components and saves time. Therefore, there is still a lot of room for improvement of scRNA-seq technologies for use in cardiovascular system research. Future research should focus on developing more efficient platforms for cardiomyocyte capture, improving sequencing coverage, and reducing costs in order to conduct large-scale single-cell sequencing experiments to address clinical problems related to the cardiovascular system. For example, breakthroughs have already been made in the fields of metabolic heart diseases, including diabetic cardiomyopathy (120) and Alcoholic-dilated Cardiomyopathy (121). In addition, scRNA-seq should be improved in terms of sequencing technology standardization, and visualization of analysis results in order to meet the needs of integrated multi-omics research. It is believed that with continuous innovation,

scRNA-seq will play a greater role in the field of cardiovascular system research and promote the development of precision medicine, including targeted diagnostic methods and treatments for cardiovascular diseases.

Author contributions

YH collected the literature and wrote the manuscript with guidance from ZZ. YZ, YL, YG, TS, XL, SS, and BY were involved in making figures and tables. All the authors read and approved the manuscript and agreed to its publication.

Funding

This work was supported by a grant from Natural Science Foundation of Shandong Province (No. ZR2021MH279 to ZZ).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Roth GA, Mensah GA, Johnson CO, Addolorato G, Ammirati E, Baddour LM, et al. Global burden of cardiovascular diseases and risk factors, 1990-2019: update from the Gbd 2019 study. *J Am Coll Cardiol.* (2020) 76:2982–3021. doi: 10.1016/j.jacc.2020.11.010

2. Chen Z, Wei L, Duru F, Chen L. Single-cell RNA sequencing: in-depth decoding of heart biology and cardiovascular diseases. *Curr Genomics.* (2020) 21:585–601. doi: 10.2174/1389202921999200604123914

3. Green ED, Watson JD, Collins FS. Human genome project: twenty-five years of big biology. *Nature.* (2015) 526:29–31. doi: 10.1038/526029a

4. Yamada S, Nomura S. Review of single-cell RNA sequencing in the heart. *Int J Mol Sci.* (2020) 21:1–2. doi: 10.3390/ijms21218345

5. Grancharova T, Gerbin KA, Rosenberg AB, Roco CM, Arakaki JE, DeLizo CM, et al. A comprehensive analysis of gene expression changes in a high replicate and open-source dataset of differentiating hipsc-derived cardiomyocytes. *Sci Rep.* (2021) 11:15845. doi: 10.1038/s41598-021-94732-1

6. Alexanian M, Przytycki PF, Micheletti R, Padmanabhan A, Ye L, Travers JG, et al. A transcriptional switch governs fibroblast activation in heart disease. *Nature.* (2021) 595:438–43. doi: 10.1038/s41586-021-03674-1

7. Cui Y, Zheng Y, Liu X, Yan L, Fan X, Yong J, et al. Single-cell transcriptome analysis maps the developmental track of the human heart. *Cell Rep.* (2019) 26:1947, 1934. doi: 10.1016/j.celrep.2019.01.079

8. Asp M, Giacomello S, Larsson L, Wu C, Fürth D, Qian X, et al. A spatiotemporal organ-wide gene expression and cell atlas of the developing human heart. *Cell.* (2019) 179:1647, 1656–7. doi: 10.1016/j.cell.2019.11.025

9. Tang F, Barbacioru C, Wang Y, Nordman E, Lee C, Xu N, et al. mRNA-seq whole-transcriptome analysis of a single cell. *Nat Methods.* (2009) 6:377–82. doi: 10.1038/nmeth.1315

10. Paik DT, Cho S, Tian L, Chang HY, Wu JC. Single-cell RNA sequencing in cardiovascular development, disease and medicine. *Nat Rev Cardiol.* (2020) 17:457–73. doi: 10.1038/s41569-020-0359-y

11. Guo G-R, Chen L, Rao M, Chen K, Song J-P, Hu S-S, et al. Modified method for isolation of human cardiomyocytes to model cardiac diseases. *J Transl Med.* (2018) 16:288. doi: 10.1186/s12967-018-1649-6

12. Dobrev D, Wettwer E, Himmel HM, Kortner A, Kuhlisch E, Schüler S, et al. G-Protein Beta(3)-subunit 825t allele is associated with enhanced human atrial inward rectifier potassium currents. *Circulation.* (2000) 102:692–7. doi: 10.1161/01.CIR.102.6.692

13. Bell RM, Mocanu MM, Yellon DM. Retrograde heart perfusion: the langendorff technique of isolated heart perfusion. *J Mol Cell Cardiol.* (2011) 50:940–50. doi: 10.1016/j.yjmcc.2011.02.018

14. Ackers-Johnson M, Li PY, Holmes AP, O'Brien S-M, Pavlovic D, Foo RS, et al. Simplified, Langendorff-free method for concomitant isolation of viable cardiac myocytes and nonmyocytes from the adult mouse heart. *Circ Res.* (2016) 119:909–20. doi: 10.1161/CIRCRESAHA.116.309202

15. Xu K, Wang R, Xie H, Hu L, Wang C, Xu J, et al. Single-cell RNA sequencing reveals cell heterogeneity and transcriptome profile of breast cancer lymph node metastasis. *Oncogenesis.* (2021) 10:66. doi: 10.1038/s41389-021-00355-6

16. Fei C, Nie L, Zhang J, Chen J. Potential applications of fluorescence-activated cell sorting (facs) and droplet-based microfluidics in promoting the discovery of specific antibodies for characterizations of fish immune cells. *Front Immunol.* (2021) 12:771231. doi: 10.3389/fimmu.2021.771231

17. Lin S, Liu Y, Zhang M, Xu X, Chen Y, Zhang H, et al. Microfluidic single-cell transcriptomics: moving towards multimodal and spatiotemporal omics. *Lab Chip.* (2021) 21:3829–49. doi: 10.1039/D1LC00607J

18. Karamitros CS, Morvan M, Vigne A, Lim J, Gruner P, Beneyton T, et al. Bacterial expression systems for enzymatic activity in droplet-based microfluidics. *Anal Chem.* (2020) 92:4908–16. doi: 10.1021/acs.analchem.9b04969

19. Gierahn TM, Wadsworth MH, Hughes TK, Bryson BD, Butler A, Satija R, et al. Seq-well: portable, low-cost RNA sequencing of single cells at high throughput. *Nat Methods.* (2017) 14:395–8. doi: 10.1038/nmeth.4179

20. Luan M-W, Lin J-L, Wang Y-F, Liu Y-X, Xiao C-L, Wu R, et al. Scsit: a high-efficiency preprocessing tool for single-cell sequencing data from split-seq. *Comput Struct Biotechnol J.* (2021) 19:4574–80. doi: 10.1016/j.csbj.2021.08.021

21. Hu P, Liu J, Zhao J, Wilkins BJ, Lupino K, Wu H, et al. Single-nucleus transcriptomic survey of cell diversity and functional maturation in postnatal mammalian hearts. *Genes Dev.* (2018) 32:1344–57. doi: 10.1101/gad.316802.118

22. Wang Y, Navin NE. Advances and applications of single-cell sequencing technologies. *Mol Cell.* (2015) 58:598–609. doi: 10.1016/j.molcel.2015.05.005

23. Huang L, Ma F, Chapman A, Lu S, Xie XS. Single-cell whole-genome amplification and sequencing: methodology and applications. *Annu Rev Genomics Hum Genet.* (2015) 16:81–2. doi: 10.1146/annurev-genom-090413-025352

24. Blagodatskikh KA, Kramarov VM, Barsova EV, Garkovenko AV, Shcherbo DS, Shelenkov AA, et al. Improved Dop-Pcr (Idop-Pcr): a robust and simple wga method for efficient amplification of low copy number genomic DNA. *PLoS ONE.* (2017) 12:e0184507. doi: 10.1371/journal.pone.0184507

25. Fu Y, Li C, Lu S, Zhou W, Tang F, Xie XS, et al. Uniform and accurate single-cell sequencing based on emulsion whole-genome amplification. *Proc Natl Acad Sci USA.* (2015) 112:11923–8. doi: 10.1073/pnas.1513988112

26. Chen C, Xing D, Tan L, Li H, Zhou G, Huang L, et al. Single-cell whole-genome analyses by linear amplification via transposon insertion (Lianti). *Science.* (2017) 356:189–94. doi: 10.1126/science.aak9787

27. Chu WK, Edge P, Lee HS, Bansal V, Bafna V, Huang X, et al. Ultraaccurate genome sequencing and haplotyping of single human cells. *Proc Natl Acad Sci USA*. (2017) 114:12512–7. doi: 10.1073/pnas.1707609114

28. Imamura H, Monsieurs P, Jara M, Sanders M, Maes I, Vanaerschot M, et al. Evaluation of whole genome amplification and bioinformatic methods for the characterization of leishmania genomes at a single cell level. *Sci Rep*. (2020) 10:15043. doi: 10.1038/s41598-020-71882-2

29. Hashimshony T, Wagner F, Sher N, Yanai I. Cel-Seq: single-cell RNA-seq by multiplexed linear amplification. *Cell Rep*. (2012) 2:666–73. doi: 10.1016/j.celrep.2012.08.003

30. Picelli S, Faridani OR, Björklund AK, Winberg G, Sagasser S, Sandberg R. Full-length RNA-seq from single cells using smart-seq2. *Nat Protoc*. (2014) 9:171–81. doi: 10.1038/nprot.2014.006

31. Ziegenhain C, Vieth B, Parekh S, Reinius B, Guillaumet-Adkins A, Smets M, et al. Comparative analysis of single-cell Rna sequencing methods. *Mol Cell*. (2017) 65:639–41. doi: 10.1016/j.molcel.2017.01.023

32. Grün D, van Oudenaarden A. Design and analysis of single-cell sequencing experiments. *Cell*. (2015) 163:799–810. doi: 10.1016/j.cell.2015.10.039

33. Yu NY-L, Hallström BM, Fagerberg L, Ponten F, Kawaji H, Carninci P, et al. Complementing tissue characterization by integrating transcriptome profiling from the human protein atlas and from the Fantom5 Consortium. *Nucleic Acids Res*. (2015) 43:6787–98. doi: 10.1093/nar/gkv608

34. Yu SCY, Jiang P, Peng W, Cheng SH, Cheung YTT, Tse OYO, et al. Single-molecule sequencing reveals a large population of long cell-free DNA molecules in maternal plasma. *Proc Natl Acad Sci USA*. (2021) 118:1, 9–10. doi: 10.1073/pnas.2114937118

35. Sheka D, Alabi N, Gordon PMK. Oxford nanopore sequencing in clinical microbiology and infection diagnostics. *Brief Bioinform*. (2021) 22:1–4. doi: 10.1093/bib/bbaa403

36. Hu T, Chitnis N, Monos D, Dinh A. Next-generation sequencing technologies: an overview. *Hum Immunol*. (2021) 82:801–11. doi: 10.1016/j.humimm.2021.02.012

37. Song Y, Xu X, Wang W, Tian T, Zhu Z, Yang C. Single cell transcriptomics: moving towards multi-omics. *Analyst*. (2019) 144:3172–89. doi: 10.1039/C8AN01852A

38. Chappell L, Russell AJC, Voet T. Single-cell (multi)omics technologies. *Annu Rev Genomics Hum Genet*. (2018) 19:15–41. doi: 10.1146/annurev-genom-091416-035324

39. Macaulay IC, Haerty W, Kumar P, Li YI, Hu TX, Teng MJ, et al. G&T-seq: parallel sequencing of single-cell genomes and transcriptomes. *Nat Methods*. (2015) 12:519–22. doi: 10.1038/nmeth.3370

40. Chaligne R, Nam AS, Landau DA. Target-seq takes aim at cancer evolution through multi-omics single-cell genotyping and transcriptomics. *Mol Cell*. (2019) 73:1092–4. doi: 10.1016/j.molcel.2019.03.009

41. Liu L, Liu C, Quintero A, Wu L, Yuan Y, Wang M, et al. Deconvolution of single-cell multi-omics layers reveals regulatory heterogeneity. *Nat Commun*. (2019) 10:470. doi: 10.1038/s41467-018-08205-7

42. Angermueller C, Clark SJ, Lee HJ, Macaulay IC, Teng MJ, Hu TX, et al. Parallel single-cell sequencing links transcriptional and epigenetic heterogeneity. *Nat Methods*. (2016) 13:229–32. doi: 10.1038/nmeth.3728

43. Frei AP, Bava F-A, Zunder ER, Hsieh EWY, Chen S-Y, Nolan GP, et al. Highly multiplexed simultaneous detection of RNAs and proteins in single cells. *Nat Methods*. (2016) 13:269–75. doi: 10.1038/nmeth.3742

44. Zhao T, Lyu S, Lu G, Juan L, Zeng X, Wei Z, et al. Sc2disease: a manually curated database of single-cell transcriptome for human diseases. *Nucleic Acids Res*. (2021) 49:D1413–9. doi: 10.1093/nar/gkaa838

45. Finak G, McDavid A, Yajima M, Deng J, Gersuk V, Shalek AK, et al. Mast: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data. *Genome Biol*. (2015) 16:278. doi: 10.1186/s13059-015-0844-5

46. Qiu X, Hill A, Packer J, Lin D, Ma Y-A, Trapnell C. Single-cell mRNA quantification and differential analysis with census. *Nat Methods*. (2017) 14:309–15. doi: 10.1038/nmeth.4150

47. Chen L, Zheng S. Bcseq: accurate single cell RNA-seq quantification with bias correction. *Nucleic Acids Res*. (2018) 46:e82. doi: 10.1093/nar/gky308

48. Shin J, Berg DA, Zhu Y, Shin JY, Song J, Bonaguidi MA, et al. Single-cell RNA-seq with waterfall reveals molecular cascades underlying adult neurogenesis. *Cell Stem Cell*. (2015) 17:360–72. doi: 10.1016/j.stem.2015.07.013

49. Ji Z, Ji H. Tscan: pseudo-time reconstruction and evaluation in single-cell rna-seq analysis. *Nucleic Acids Res*. (2016) 44:e117. doi: 10.1093/nar/gkw430

50. Juliá M, Telenti A, Rausell A. Sincell: An R/bioconductor package for statistical assessment of cell-state hierarchies from single-cell RNA-seq. *Bioinformatics*. (2015) 31:3380–2. doi: 10.1093/bioinformatics/btv368

51. Welch JD, Hartemink AJ, Prins JF. Slicer: inferring branched, nonlinear cellular trajectories from single cell RNA-seq data. *Genome Biol*. (2016) 17:106. doi: 10.1186/s13059-016-0975-3

52. Wang J-J, Ge W, Zhai Q-Y, Liu J-C, Sun X-W, Liu W-X, et al. Single-cell transcriptome landscape of ovarian cells during primordial follicle assembly in mice. *PLoS Biol*. (2020) 18:e3001025. doi: 10.1371/journal.pbio.3001025

53. Mullin S, Zola J, Lee R, Hu J, MacKenzie B, Brickman A, et al. Longitudinal K-means approaches to clustering and analyzing ehr opioid use trajectories for clinical subtypes. *J Biomed Inform*. (2021) 122:103889. doi: 10.1016/j.jbi.2021.103889

54. Bergen V, Soldatov RA, Kharchenko PV, Theis FJ. RNA velocity-current challenges and future perspectives. *Mol Syst Biol*. (2021) 17:e10282. doi: 10.15252/msb.202110282

55. Cheng J, Liao J, Shao X, Lu X, Fan X. Multiplexing methods for simultaneous large-scale transcriptomic profiling of samples at single-cell resolution. *Adv Sci*. (2021) 8:e2101229. doi: 10.1002/advs.202101229

56. Jia G, Preussner J, Chen X, Guenther S, Yuan X, Yekelchyk M, et al. Single cell RNA-seq and atac-seq analysis of cardiac progenitor cell transition states and lineage settlement. *Nat Commun*. (2018) 9:4877. doi: 10.1038/s41467-018-07307-6

57. Xiao Y, Hill MC, Zhang M, Martin TJ, Morikawa Y, Wang S, et al. Hippo signaling plays an essential role in cell state transitions during cardiac fibroblast development. *Dev Cell*. (2018) 45:153–169.e6. doi: 10.1016/j.devcel.2018.03.019

58. Sahara M, Santoro F, Sohlmér J, Zhou C, Witman N, Leung CY, et al. Population and single-cell analysis of human cardiogenesis reveals unique Lgr5 ventricular progenitors in embryonic outflow tract. *Dev Cell*. (2019) 48:2–14. doi: 10.1016/j.devcel.2019.01.005

59. Bedada FB, Wheelwright M, Metzger JM. Maturation status of sarcomere structure and function in human Ipsc-derived cardiac myocytes. *Biochim Biophys Acta*. (2016) 1863(7 Pt B):1829–38. doi: 10.1016/j.bbamcr.2015.11.005

60. Ou M, Zhao M, Li C, Tang D, Xu Y, Dai W, et al. Single-cell sequencing reveals the potential oncogenic expression atlas of human Ipsc-derived cardiomyocytes. *Biol Open*. (2021) 10:3, 6–7. doi: 10.1242/bio.053348

61. Paik DT, Tian L, Lee J, Sayed N, Chen IY, Rhee S, et al. Large-scale single-cell RNA-seq reveals molecular signatures of heterogeneous populations of human induced pluripotent stem cell-derived endothelial cells. *Circ Res*. (2018) 123:443–50. doi: 10.1161/CIRCRESAHA.118.312913

62. Churko JM, Garg P, Treutlein B, Venkatasubramanian M, Wu H, Lee J, et al. Defining human cardiac transcription factor hierarchies using integrated single-cell heterogeneity analysis. *Nat Commun*. (2018) 9:4906. doi: 10.1038/s41467-018-07333-4

63. Wu S-p, Cheng C-M, Lanz RB, Wang T, Respress JL, Ather S, et al. Atrial identity is determined by a coup-tfii regulatory network. *Dev Cell*. (2013) 25:417–26. doi: 10.1016/j.devcel.2013.04.017

64. Suryawanshi H, Clancy R, Morozov P, Halushka MK, Buyon JP, Tuschl T. Cell atlas of the foetal human heart and implications for autoimmune-mediated congenital heart block. *Cardiovasc Res*. (2020) 116:1446–57. doi: 10.1093/cvr/cvz257

65. de Soysa TY, Ranade SS, Okawa S, Ravichandran S, Huang Y, Salunga HT, et al. Single-cell analysis of cardiogenesis reveals basis for organ-level developmental defects. *Nature* (2019) 572:120–4. doi: 10.1038/s41586-019-1414-x

66. Nomura S, Satoh M, Fujita T, Higo T, Sumida T, Ko T, et al. Cardiomyocyte gene programs encoding morphological and functional signatures in cardiac hypertrophy and failure. *Nat Commun*. (2018) 9:4435. doi: 10.1038/s41467-018-06639-7

67. Tombor LS, John D, Glaser SF, Luxán G, Forte E, Furtado M, et al. Single cell sequencing reveals endothelial plasticity with transient mesenchymal activation after myocardial infarction. *Nat Commun*. (2021) 12:681. doi: 10.1038/s41467-021-20905-1

68. Talman V, Ruskoaho H. Cardiac fibrosis in myocardial infarction-from repair and remodeling to regeneration. *Cell Tissue Res*. (2016) 365:563–81. doi: 10.1007/s00441-016-2431-9

69. Ruiz-Villalba A, Romero JP, Hernández SC, Vilas-Zornoza A, Fortelny N, Castro-Labrador L, et al. Single-cell RNA sequencing analysis reveals a crucial role for cthrc1 (collagen triple helix repeat containing 1) cardiac fibroblasts after myocardial infarction. *Circulation*. (2020) 142:1831–47. doi: 10.1161/CIRCULATIONAHA.119.044557

70. Winkels H, Ehinger E, Ghosheh Y, Wolf D, Ley K. Atherosclerosis in the single-cell era. *Curr Opin Lipidol*. (2018) 29:389–96. doi: 10.1097/MOL.0000000000000537

71. Gu L, Saha ST, Thomas J, Kaur M. Targeting cellular cholesterol for anticancer therapy. *FEBS J.* (2019) 286:4192–208. doi: 10.1111/febs.15018

72. Willemsen L, de Winther MP. Macrophage subsets in atherosclerosis as defined by single-cell technologies. *J Pathol.* (2020) 250:705–14. doi: 10.1002/path.5392

73. Cole JE, Park I, Ahern DJ, Kassiteridi C, Danso Abeam D, Goddard ME, et al. Immune cell census in murine atherosclerosis: cytometry by time of flight illuminates vascular myeloid cell diversity. *Cardiovasc Res.* (2018) 114:1360–71. doi: 10.1093/cvr/cvy109

74. Butcher MJ, Filipowicz AR, Waseem TC, McGary CM, Crow KJ, Magilnick N, et al. Atherosclerosis-driven treg plasticity results in formation of a dysfunctional subset of plastic Ifnγ+ Th1/Tregs. *Circ Res.* (2016) 119:1190–203. doi: 10.1161/CIRCRESAHA.116.309764

75. Wirka RC, Wagh D, Paik DT, Pjanic M, Nguyen T, Miller CL, et al. Atheroprotective roles of smooth muscle cell phenotypic modulation and the Tcf21 disease gene as revealed by single-cell analysis. *Nat Med.* (2019) 25:1280–9. doi: 10.1038/s41591-019-0512-5

76. Gu W, Ni Z, Tan Y-Q, Deng J, Zhang S-J, Lv Z-C, et al. Adventitial cell atlas of Wt (wild type) and apoe (apolipoprotein E)-deficient mice defined by single-cell RNA sequencing. *Arterioscler Thromb Vasc Biol.* (2019) 39:1055–71. doi: 10.1161/ATVBAHA.119.312399

77. Kim K, Shim D, Lee JS, Zaitsev K, Williams JW, Kim K-W, et al. Transcriptome analysis reveals nonfoamy rather than foamy plaque macrophages are proinflammatory in atherosclerotic murine models. *Circ Res.* (2018) 123:1127–42. doi: 10.1161/CIRCRESAHA.118.312804

78. Yao F, Yu P, Li Y, Yuan X, Li Z, Zhang T, et al. Histone variant H2aZ is required for the maintenance of smooth muscle cell identity as revealed by single-cell transcriptomics. *Circulation.* (2018) 138:2274–88. doi: 10.1161/CIRCULATIONAHA.117.033114

79. Yang H, Zhou T, Stranz A, DeRoo E, Liu B. Single-cell RNA sequencing reveals heterogeneity of vascular cells in early stage murine abdominal aortic aneurysm-brief report. *Arterioscler Thromb Vasc Biol.* (2021) 41:1158–66. doi: 10.1161/ATVBAHA.120.315607

80. Antman EM, Loscalzo J. Precision medicine in cardiology. *Nat Rev Cardiol.* (2016) 13:591–602. doi: 10.1038/nrcardio.2016.101

81. Hulin A, Hortells L, Gomez-Stallons MV, O'Donnell A, Chetal K, Adam M, et al. Maturation of heart valve cell populations during postnatal remodeling. *Development.* (2019) 146:2, 4–10. doi: 10.1242/dev.173047

82. Yang J, Liao B, Zhang T, Xu Y. Editorial: bioinformatics analysis of single cell sequencing data and applications in precision medicine. *Front Genet.* (2019) 10:1358. doi: 10.3389/fgene.2019.01358

83. Lukassen S, Chua RL, Trefzer T, Kahn NC, Schneider MA, Muley T, et al. SARS-CoV-2 receptor Ace2 and Tmprss2 are primarily expressed in bronchial transient secretory cells. *EMBO J.* (2020) 39:e105114. doi: 10.15252/embj.2020105114

84. Zheng Y-Y, Ma Y-T, Zhang J-Y, Xie X. Covid-19 and the cardiovascular system. *Nat Rev Cardiol.* (2020) 17:259–60. doi: 10.1038/s41569-020-0360-5

85. Liu H, Gai S, Wang X, Zeng J, Sun C, Zhao Y, et al. Single-cell analysis of SARS-CoV-2 receptor Ace2 and spike protein priming expression of proteases in the human heart. *Cardiovasc Res.* (2020) 116:1733–41. doi: 10.1093/cvr/cvaa191

86. Armingol E, Officer A, Harismendy O, Lewis NE. Deciphering cell-cell interactions and communication from gene expression. *Nat Rev Genet.* (2021) 22:71–88. doi: 10.1038/s41576-020-00292-x

87. Skelly DA, Squiers GT, McLellan MA, Bolisetty MT, Robson P, Rosenthal NA, et al. Single-cell transcriptional profiling reveals cellular diversity and intercommunication in the mouse heart. *Cell Rep.* (2018) 22:600–10. doi: 10.1016/j.celrep.2017.12.072

88. Wang L, Yu P, Zhou B, Song J, Li Z, Zhang M, et al. Single-cell reconstruction of the adult human heart during heart failure and recovery reveals the cellular landscape underlying cardiac function. *Nat Cell Biol.* (2020) 22:108–19. doi: 10.1038/s41556-019-0446-7

89. Kester L, van Oudenaarden A. Single-cell transcriptomics meets lineage tracing. *Cell Stem Cell.* (2018) 23:166–79. doi: 10.1016/j.stem.2018.04.014

90. Tam PPL, Ho JWK. Cellular diversity and lineage trajectory: insights from mouse single cell transcriptomes. *Development.* (2020) 147:1, 4–6. doi: 10.1242/dev.179788

91. Trapnell C, Cacchiarelli D, Grimsby J, Pokharel P, Li S, Morse M, et al. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat Biotechnol.* (2014) 32:381–6. doi: 10.1038/nbt.2859

92. Qiu X, Mao Q, Tang Y, Wang L, Chawla R, Pliner HA, et al. Reversed graph embedding resolves complex single-cell trajectories. *Nat Methods.* (2017) 14:979–82. doi: 10.1038/nmeth.4402

93. Li D-Q, Kim S, Li J-M, Gao Q, Choi J, Bian F, et al. Single-cell transcriptomics identifies limbal stem cell population and cell types mapping its differentiation trajectory in limbal basal epithelium of human cornea. *Ocul Surf.* (2021) 20:20–32. doi: 10.1016/j.jtos.2020.12.004

94. Kołat D, Kałuzińska Z, Orzechowska M, Bednarek AK, Płuciennik E. Functional genomics of ap-2α and Ap-2γ in cancers: *in silico* study. *BMC Med Genomics.* (2020) 13:174. doi: 10.1186/s12920-020-00823-9

95. Wu H, Kirita Y, Donnelly EL, Humphreys BD. Advantages of single-nucleus over single-cell rna sequencing of adult kidney: rare cell types and novel cell states revealed in fibrosis. *J Am Soc Nephrol.* (2019) 30:23–32. doi: 10.1681/ASN.2018090912

96. Wolfien M, Galow A-M, Müller P, Bartsch M, Brunner RM, Goldammer T, et al. Single-nucleus sequencing of an entire mammalian heart: cell type composition and velocity. *Cells.* (2020) 9:1, 3–6. doi: 10.3390/cells9020318

97. Galow A-M, Wolfien M, Müller P, Bartsch M, Brunner RM, Hoeflich A, et al. Integrative cluster analysis of whole hearts reveals proliferative cardiomyocytes in adult mice. *Cells.* (2020) 9:1, 5–9. doi: 10.3390/cells9051144

98. Chaffin M, Papangeli I, Simonson B, Akkad A-D, Hill MC, Arduini A, et al. Single-nucleus profiling of human dilated and hypertrophic cardiomyopathy. *Nature.* (2022) 1–6. doi: 10.1038/s41586-022-04817-8

99. Slyper M, Porter CBM, Ashenberg O, Waldman J, Drokhlyansky E, Wakiro I, et al. A single-cell and single-nucleus RNA-seq toolbox for fresh and frozen human tumors. *Nat Med.* (2020) 26:792–802. doi: 10.1038/s41591-020-0844-1

100. Thrupp N, Sala Frigerio C, Wolfs L, Skene NG, Fattorelli N, Poovathingal S, et al. Single-nucleus RNA-seq is not suitable for detection of microglial activation genes in humans. *Cell Rep.* (2020) 32:108189. doi: 10.1016/j.celrep.2020.108189

101. Muto Y, Wilson PC, Ledru N, Wu H, Dimke H, Waikar SS, et al. Single cell transcriptional and chromatin accessibility profiling redefine cellular heterogeneity in the adult human kidney. *Nat Commun.* (2021) 12:2190. doi: 10.1038/s41467-021-22368-w

102. Thomas ED, Timms AE, Giles S, Harkins-Perry S, Lyu P, Hoang T, et al. Cell-specific *cis*-regulatory elements and mechanisms of non-coding genetic disease in human retina and retinal organoids. *Dev Cell.* (2022) 57:1, 4–13. doi: 10.1016/j.devcel.2022.02.018

103. Chattopadhyay PK, Winters AF, Lomas WE, Laino AS, Woods DM. High-parameter single-cell analysis. *Annu Rev Anal Chem.* (2019) 12:411–30. doi: 10.1146/annurev-anchem-061417-125927

104. Fernandez DM, Rahman AH, Fernandez NF, Chudnovskiy A, Amir E-AD, Amadori L, et al. Single-cell immune landscape of human atherosclerotic plaques. *Nat Med.* (2019) 25:1576–88. doi: 10.1038/s41591-019-0590-4

105. Chen P-Y, Qin L, Li G, Malagon-Lopez J, Wang Z, Bergaya S, et al. Smooth muscle cell reprogramming in aortic aneurysms. *Cell Stem Cell.* (2020) 26:2, 11–13. doi: 10.1016/j.stem.2020.02.013

106. van den Brink SC, Sage F, Vértesy Á, Spanjaard B, Peterson-Maduro J, Baron CS, et al. Single-cell sequencing reveals dissociation-induced gene expression in tissue subpopulations. *Nat Methods.* (2017) 14:935–6. doi: 10.1038/nmeth.4437

107. Ståhl PL, Salmén F, Vickovic S, Lundmark A, Navarro JF, Magnusson J, et al. Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science.* (2016) 353:78–82. doi: 10.1126/science.aaf2403

108. Hu KH, Eichorst JP, McGinnis CS, Patterson DM, Chow ED, Kersten K, et al. Zipseq: barcoding for real-time mapping of single cell transcriptomes. *Nat Methods.* (2020) 17:833–43. doi: 10.1038/s41592-020-0880-2

109. Fischer MA, Vondriska TM. Clinical Epigenomics For Cardiovascular Disease: Diagnostics And Therapies. *J Mol Cell Cardiol.* (2021) 154:1–3, 12–13. doi: 10.1016/j.yjmcc.2021.01.011

110. Asare Y, Campbell-James TA, Bokov Y, Yu LL, Prestel M, El Bounkari O, et al. Histone deacetylase 9 activates ikk to regulate atherosclerotic plaque vulnerability. *Circ Res.* (2020) 127:811–23. doi: 10.1161/CIRCRESAHA.120.316743

111. Barba I, Andrés M, Garcia-Dorado D. Metabolomics and heart diseases: from basic to clinical approach. *Curr Med Chem.* (2019) 26:46–59. doi: 10.2174/0929867324666171006151408

112. Nicholson JK, Lindon JC, Holmes E. 'Metabonomics': understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological nmr spectroscopic data. *Xenobiotica.* (1999) 29:1181–9. doi: 10.1080/004982599238047

113. Murashige D, Jang C, Neinast M, Edwards JJ, Cowan A, Hyman MC, et al. Comprehensive quantification of fuel use by the failing and nonfailing human heart. *Science.* (2020) 370:364–8. doi: 10.1126/science.abc8861

114. McGarrah RW, Crown SB, Zhang G-F, Shah SH, Newgard CB. Cardiovascular metabolomics. *Circ Res.* (2018) 122:1238–58. doi: 10.1161/CIRCRESAHA.117.311002

115. Leon-Mimila P, Wang J, Huertas-Vazquez A. Relevance of multi-omics studies in cardiovascular diseases. *Front Cardiovasc Med.* (2019) 6:91. doi: 10.3389/fcvm.2019.00091

116. Ranzoni AM, Tangherloni A, Berest I, Riva SG, Myers B, Strzelecka PM, et al. Integrative single-cell RNA-seq and atac-seq analysis of human developmental hematopoiesis. *Cell Stem Cell.* (2021) 28:1, 13–14. doi: 10.1016/j.stem.2020.11.015

117. Moncada R, Barkley D, Wagner F, Chiodin M, Devlin JC, Baron M, et al. Integrating microarray-based spatial transcriptomics and single-cell RNA-seq reveals tissue architecture in pancreatic ductal adenocarcinomas. *Nat Biotechnol.* (2020) 38:333–42. doi: 10.1038/s41587-019-0392-8

118. Zhao R, Lu J, Li Q, Xiong F, Zhang Y, Zhu J, et al. Single-cell heterogeneity analysis and crispr screens in min6 cell line reveal transcriptional regulators of insulin. *Cell Cycle.* (2021) 20:2053–65. doi: 10.1080/15384101.2021.1969204

119. Kaya-Okur HS, Wu SJ, Codomo CA, Pledger ES, Bryson TD, Henikoff JG, et al. Cut&tag for efficient epigenomic profiling of small samples and single cells. *Nat Commun.* (2019) 10:1930. doi: 10.1038/s41467-019-09982-5

120. Longo M, Scappaticcio L, Cirillo P, Maio A, Carotenuto R, Maiorino MI, et al. Glycemic control and the heart: the tale of diabetic cardiomyopathy continues. *Biomolecules.* (2022) 12:1–7. doi: 10.3390/biom12020272

121. Fernández-Solà J. The effects of ethanol on the heart: alcoholic cardiomyopathy. *Nutrients.* (2020) 12:572. doi: 10.3390/nu12 00572

frontiers | Frontiers in Cardiovascular Medicine

TYPE Original Research
PUBLISHED 09 January 2023
DOI 10.3389/fcvm.2022.1062106

OPEN ACCESS

EDITED BY
Jingyan Han,
Boston University, United States

REVIEWED BY
Roberto Vazquez-Padron,
University of Miami, United States
Yanming Li,
Baylor College of Medicine,
United States
Bhama Ramkhelawon,
Grossman School of Medicine,
New York University, United States

*CORRESPONDENCE
Hairong Zhang
✉ 18838221713@163.com
Ningheng Chen
✉ fccchennh@zzu.edu.cn
Xueli Guo
✉ guoxueli2000@163.com

SPECIALTY SECTION
This article was submitted to General Cardiovascular Medicine, a section of the journal Frontiers in Cardiovascular Medicine

RECEIVED 05 October 2022
ACCEPTED 16 December 2022
PUBLISHED 09 January 2023

CITATION
Wu S, Liu S, Wang B, Li M, Cheng C, Zhang H, Chen N and Guo X (2023) Single-cell transcriptome *in silico* analysis reveals conserved regulatory programs in macrophages/monocytes of abdominal aortic aneurysm from multiple mouse models and human. *Front. Cardiovasc. Med.* 9:1062106. doi: 10.3389/fcvm.2022.1062106

Single-cell transcriptome *in silico* analysis reveals conserved regulatory programs in macrophages/monocytes of abdominal aortic aneurysm from multiple mouse models and human

Shiyong Wu[1], Shibiao Liu[1], Baoheng Wang[1], Meng Li[1], Chao Cheng[2], Hairong Zhang[3]*, Ningheng Chen[1]* and Xueli Guo[1]*

[1]Department of Vascular Surgery, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, Henan, China, [2]Center for Genome Analysis, Wuhan Ruixing Biotechnology Co., Ltd., Wuhan, China, [3]Department of Colorectal and Anal Surgery, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, Henan, China

Abdominal aortic aneurysm (AAA) is a life-threatening disease and there is currently a lack of effective treatment to prevent it rupturing. ScRNA-seq studies of AAA are still lacking. In the study, we analyzed the published AAA scRNA-seq datasets from the mouse elastase-induced model, $CaCl_2$ treatment model, Ang II-induced model and human by using bioinformatic approaches and *in silico* analysis. A total of 26 cell clusters were obtained and 11 cell types were identified from multiple mouse AAA models. Also, the proportion of Mφ/Mo increased in the AAA group and Mφ/Mo was divided into seven subtypes. There were significant differences in transcriptional regulation patterns of Mφ/Mo in different AAA models. The enrichment pathways of upregulated or downregulated genes from Mφ/Mo in the three mouse datasets were different. The actived regulons of Mφ/Mo had strong specificity and the repressed regulons showed high consistency. The co-upregulated genes as well as actived regulons and co-downregulated genes as well as repressed regulons were closely correlated and formed regulatory networks. Mφ/Mo from human AAA dataset was divided into five subtypes. The proportion of three macrophage subpopulations increased but the proportion of two monocyte subpopulations decreased. In the AAA group, the upregulated or downregulated genes of Mφ/Mo were enriched in different pathways. After further analyzing the genes in Mφ/Mo of both mouse and human scRNA-seq datasets, two genes were upregulated in the four datasets,

IL-1B and *THBS1*. In conclusion, *in silico* analysis of scRNA-seq revealed that Mφ/Mo and their regulatory related genes as well as interaction networks played an important role in the pathogenesis of AAA.

KEYWORDS

abdominal aortic aneurysm, single-cell transcriptome analysis, macrophages/monocytes, single-cell RNA sequencing, *in silico* analysis

Introduction

Abdominal aortic aneurysm (AAA), a cardiovascular disease with serious complications, is charactered by permanent local dilation of the abdominal aortic wall that exceeds 50% of the normal blood vessel diameter (1). As the disease worsens and the inner diameter of aorta dilates, the risk of AAA rupture increases (2). Over time, AAA can grow in size and rupture, causing life-threatening bleeding. Patients with AAA are usually asymptomatic until a catastrophic rupture occurs (3). Inflammatory cell infiltration, neovascularization, and the production as well as activation of various proteases as well as cytokines contributed to the development of AAA (4). Inflammatory processes played a critical role in AAA and significantly affected many determinants of aortic wall remodeling (5, 6). Various inflammatory cell types in AAA, such as macrophages, $CD4^+$ T cells, and B cells, had great importance in the pathological aortic wall through phenotypic regulation (7). In addition, continuous crosstalk between various cells also affected the occurrence and development of AAA (8, 9). Therefore, the study of cell heterogeneity in the development of AAA may be a breakthrough to further understand its pathogenesis and develop targeted drugs. However, the relevant studies are still scarce.

The advances in the pathophysiology of AAA partly depend on the development and application of effective animal AAA models that replicate the key aspects of human. The basic premise of these animal models is that they share the same biochemical and cellular mechanisms as human possesses. Mouse AAA models have been widely used to study the occurrence and progression of AAA, including spontaneous aneurysm formation, drug-induced aneurysm, surgically induced aneurysm, genetic manipulation, chemical induction, and dietary models. Chemical methods included intracavity infusion of elastase, perivascular incubation of calcium chloride ($CaCl_2$), and subcutaneous infusion of angiotensin II (Ang II) (10). Validation of mouse AAA models will provide insights into the mechanism of progression of human AAA. However, in mouse AAA models induced by different ways, what are the differences as well as similarities of gene expression patterns in different cell clusters and what molecular mechanisms are common to human AAA remains unclear.

The main pathological features of AAA included extracellular matrix remodeling that related to degeneration and loss of vascular smooth muscle cells (VSMCs), accumulation and activation of inflammatory cells (11). MMP9 derived from macrophages was a key factor in the degradation of extracellular matrix and crucial for the development of AAA (12). Different monocytes and macrophages subpopulations played a key and differential role in the initiation, progression, and healing of AAA process (5). The specific role of macrophages/monocytes (Mφ/Mo) accumulation in AAA remains unclear. The human aneurysm tissue showed numerous infiltrating macrophages (5). Multiple studies have shown that multiple genes mediated the development or suppression of AAA by regulating macrophages (13–15). For example, *IL-1*β and *TNF-*α influenced the formation of AAA through differential effects on macrophage polarization (13). Histone demethylase JMJD3 induced *NFκB*-mediated inflammatory gene transcription in infiltrating aortic macrophages. Targeted inhibition of *JMJD3* significantly reduced AAA amplification and attenuated macrophage-mediated inflammation *in vivo* (14). Chemokine *CCL7* contributed to Ang II-induced AAA by promoting M1 phenotype of macrophages through *CCR1/JAK2/STAT1* signaling pathway (15). Macrophages can also work through crosstalk with other cells. For example, macrophage-derived netrin-1 promoted AAA formation by activating MMP3 in VSMCs (8). However, most of the existing research conclusion were based on a certain AAA model, and there was still a lack of comparison of the differences of macrophage transcriptional regulation modes in different mouse AAA models.

Previous sequencing technologies are to study the aggregation of cells, reflecting the average level of cell clusters, which cannot objectively reflect the information of the occurrence and development of diseases; but the information contained and expression level between cells vary greatly. Single-cell RNA sequencing (scRNA-seq) enabled the amplification and sequencing of whole transcriptome at the single-cell level. The principle was to amplify the trace amounts of whole transcriptome RNA from isolated single cells, then perform high-throughput sequencing,

finally output the gene expression level of each cell through bioinformatics analysis. ScRNA-seq technology can reveal the overall level of gene expression status within a single cell, accurately reflect the heterogeneity among cells (16–18), reveal the diversity of immune cells in tissues (19), and build an interaction network among different cell populations (17). Therefore, single-cell transcriptome sequencing has been widely used to detect gene expression in different cell types during reproduction, development and disease occurrence, and to reveal the molecular mechanisms of the functions and effects of different cells in these processes.

Single-cell RNA sequencing provided a useful tool for studying the heterogeneity and dynamic regulation of AAA cells and much information on cell-specific gene expression profiles during the development and progression of AAA. Davis identified increased *JMJD3* in aortic Mφ/Mo by using scRNA-seq from human AAA tissue, leading to upregulation of inflammatory immune responses (14). Hadi found that macrophage derived netrin-1 promotes the formation of AAA by activating MMP3 in VSMCs through scRNA-seq of mouse AAA (8). Yang induced AAA in C57BL/6J mouse by perivascular application of $CaCl_2$ for scRNA-seq, and analyzed the transcriptional profile and potential functional characteristics of populations in VSMCs, fibroblasts and macrophages (20). Yu demonstrated the key role of *Malat1* VSMCs in the occurrence and progression of AAA by scRNA-seq of Ang II-induced AAA treated with or without the inhibitor (21). Zhao demonstrated the heterogeneity and cellular response of VSMCs and Mφ/Mo in the progression of AAA by scRNA-seq of elastase-induced AAA (22). These scRNA-seq datasets provided insights into the pathogenesis of diseases and were rich resources for developing novel targeted therapy strategies. Based on these published datasets of different AAA mouse models and human AAA tissues, we hoped to uncover the conserved transcriptional regulatory patterns of macrophages during the development of AAA, and thereby elucidated the potential key roles of these genes in the pathogenesis of AAA.

Materials and methods

Retrieval and process of public data

Unique Molecular Identifier (UMI) count matrix for AAA and control sample scRNA-seq data of 4 datasets (14, 20–22) were downloaded from the GEO database. The UMI count matrix was converted from R package Seurat to Seurat objects (23) (version 4.0.4). Cells with UMI number < 500 or detected genes < 200 or those with mitochondrial-derived UMI counts of more than 10% were considered to be of low quality and were removed. Genes detected in less than 3 cells were removed for downstream analysis.

ScRNA-seq data preprocessing and quality control

After quality control, the UMI count matrix was log normalized. Then top 2,000 variable genes were applied to create potential Anchors by using Seurat's FindIntegrationAnchors function. Subsequently, IntegrateData function was applied to integrate data. In order to reduce the dimensionality of the scRNA-seq datasets, principal component analysis (PCA) was performed on the integrated data matrix. With Elbowplot function from Seurat, top 50 principal components (PCs) were applied to perform the downstream analysis. The main cell clusters were recognized with the FindClusters function of Seurat, with resolution set as default (res = 0.4). Finally, the cells aggregated into different cell types. Then they were visualized with t-distributed stochastic neighbor embedding (tSNE) or uniform manifold approximation and projection (UMAP) plots. For the gene markers of each cell clusters, we used the FindMarkers function in the Seurat package (version 4.0.4), and then we annotated cell types using previously published marker genes (24, 25).

Differential gene expression analysis

The Seurat package FindMarkers/FindAllMarkers functions (one-tailed Wilcoxon rank sum test, *p*-values adjusted for multiple testing using the Bonferroni correction) were used to determine the differentially expressed genes (DEGs). When DEGs were calculated, the expression difference of all genes on the natural logarithmic scale was at least 0.5, and the adjusted *p*-value was less than 0.05.

Transcription factor regulatory network analysis

The modules of transcription factor (TF) were recognized by the SCENIC (26) python workflow (version 0.11.2) using default parameters.[1] A list of mouse TF genes was extracted from the resources of pySCENIC.[2] Actived TFs were identified in the AUC matrix and differentially actived TFs were selected by using the FindAllMarkers function of the Seurat package. The networks of the modules with TFs and their target genes were visualized by Cytoscape (version 3.9.1).[3]

1 http://scenic.aertslab.org
2 https://github.com/aertslab/pySCENIC/tree/master/resources
3 https://cytoscape.org

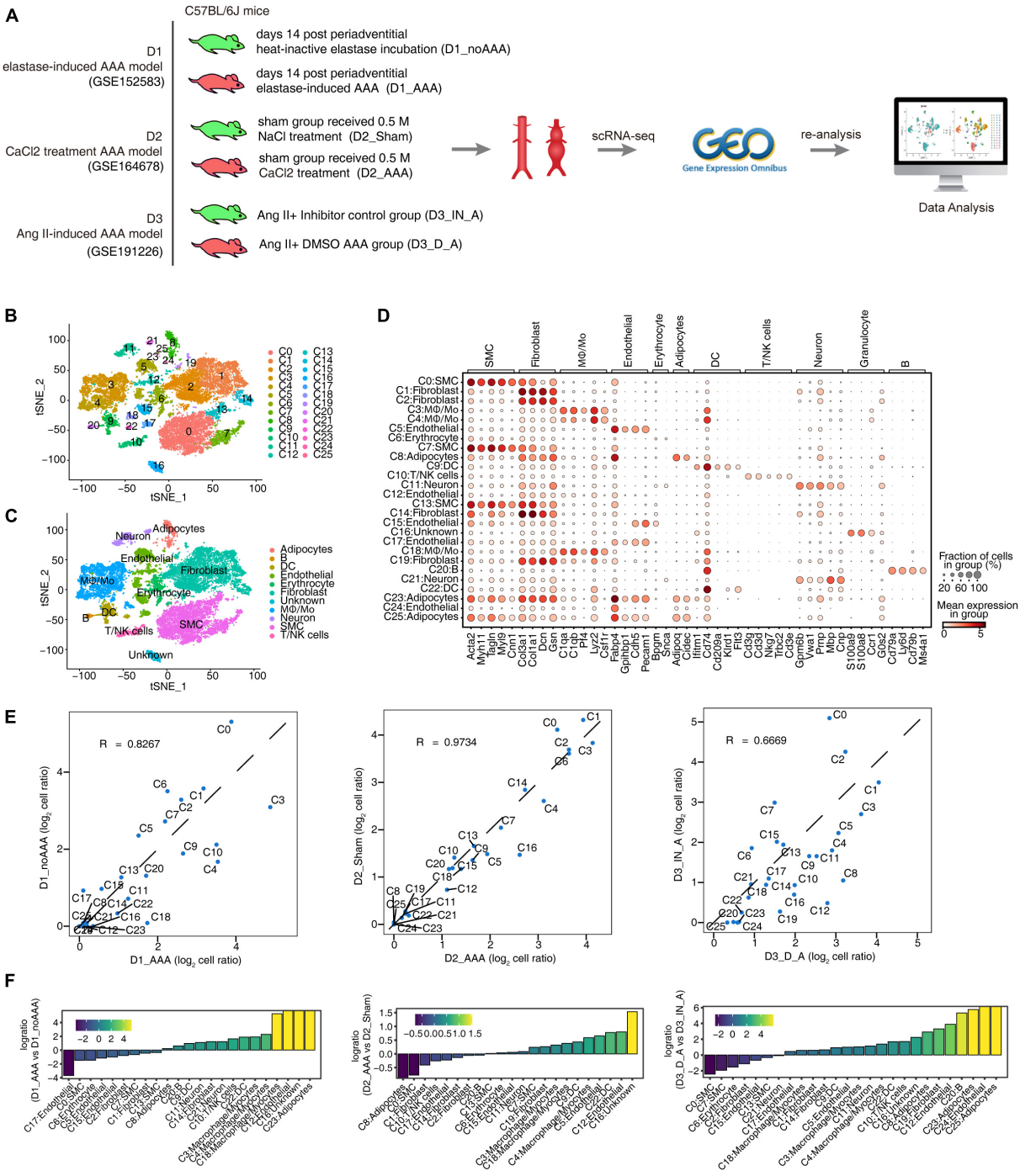


FIGURE 1
Single-cell RNA sequencing (ScRNA-seq) analysis of abdominal aortic tissue from different mouse abdominal aortic aneurysm (AAA) models identified distinct macrophages/monocytes (Mφ/Mo) types. **(A)** Schematic illustration of sample preparation and scRNA-seq data processing. **(B, C)** t-distributed stochastic neighbor embedding (tSNE) plot of composite single-cell transcriptomic profiles from all six abdominal aortic samples from three datasets. Colors indicated cell clusters along with annotations. Mφ/Mo, macrophage/monocyte cells; SMC, smooth muscle cells; DC, dendritic cells. **(D)** Dot plot showing expression of representative genes in each cell type. **(E)** Scatter plot comparing the proportions of cell populations of each cell type in two groups. **(F)** Rank order based on decreasing values of the relative frequency ratio between two sample groups.

Functional enrichment analysis

To sort out functional categories of genes, Gene Ontology (GO) terms and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways were identified using KOBAS 2.0 (27). Hypergeometric test and Benjamini-Hochberg FDR controlling procedure were applied to define the enrichment of each term.

Other statistical analysis

The pheatmap package[4] in R was used for performing the clustering based on Euclidean distance.

Results

ScRNA-seq analysis of abdominal aortic tissue from different mouse AAA models identified distinct cells types

In order to explore the differences and similarities of specifically key regulatory factors of aortic tissue cells in mouse AAA models constructed by different induction methods, we collected single-cell transcriptome datasets of the three published mouse AAA models: elastase induction (D1), $CaCl_2$ induction (D2), and Ang II induction (D3). Each dataset included two samples of AAA group and control group (**Figure 1A**). Through data quality control procedure (**Supplementary Table 1**), the transcriptome map data of 22,391 single cells were obtained. The transcriptome expression matrix of them was normalized and analyzed by principal component dimension reduction. The top 50 PCs were selected for tSNE dimension reduction and visualization. After unbiased clustering analysis, 26 cell clusters were obtained (**Figure 1B** and **Supplementary Figure 1A**). Using the newly published SCTYPE software and combined with the reported cell markers in mouse artery tissue, 11 different cell types were identified (**Figure 1C**). As shown in **Figure 1D** and **Supplementary Figure 1B**, each cell type had specific representative genes and expression of top 3 marker genes. Most cell clusters were detected in three datasets, and only a few cell clusters, such as C24 and C25, were detected only in D3, because D3 had the largest number of cells (**Supplementary Figure 1C**). There was both consistency and difference in the variation trend of each cell subtype proportion in the three datasets (**Figures 1E, F**). The proportion of cell subtypes in AAA increased as follows: C3: Mφ/Mo, C4: Mφ/Mo, C9: DC, C11: Neuron, C12: Endothelial cell, C18: Mφ/Mo; the proportion of C0 (SMC) and C2 (Fibroblast) decreased in AAA; the proportion of other cell subtypes showed different trends in the three datasets (**Figures 1E, F**). The above results provided a panorama of cell compositions and changes in different mouse AAA models, and these cell subgroups also differed in functions (**Supplementary Figure 1D**).

Single-cell analysis revealed complex Mφ/Mo heterogeneity and conserved regulated genes between AAA and control samples

Studies have shown that Mφ/Mo played a unique and important role in the occurrence and development of AAA in both patients and animal models (5, 14, 28). So, it is particularly important to further study the changes of Mφ/Mo in different mouse AAA models and the regulation of its related gene expression. On the whole, the proportion of Mφ/Mo increased in the AAA group compared with the control group, with a larger increase in D1 and D3 datasets and a smaller increase in D2 dataset (**Figure 2A**). It may be related to the shorter construction time of the AAA model in D2 dataset. We conducted secondary clustering of Mφ/Mo. A total of seven subtypes were generated (**Figure 2B** and **Supplementary Figure 2A**) and the top 3 marker genes of these subtypes were shown (**Figure 2C**). Taking Mφ/Mo as the overall background, it's found that the changes of the relative proportions of Mφ/Mo subtypes in the three groups were highly dynamic and heterogeneous (**Figure 2D** and **Supplementary Figure 2B**). For example, the relative proportion of m-M0 type increased in AAA of D1 dataset and decreased in AAA of D2 and D3 datasets; the m-M1 as well as m-M2 types increased and m-M3 as well as m-M6 types decreased in AAA of the three datasets; the proportion of m-M4 and m-M5 varied in the three datasets (**Figure 2D** and **Supplementary Figure 2B**). The marker genes enriched functions of Mφ/Mo subtypes were also varied to different extent (**Figure 2E**). Further, we explored the changes in gene expressions in Mφ/Mo during the development of AAA. Compared with the control group, the enrichment functions of upregulated genes were different. But it was relatively consistent that inflammatory response, immune system and apoptotic process pathways were mainly enriched in the three datasets (**Figure 2F**). For downregulated genes, the enriched pathways were also significantly different in the three datasets (**Supplementary Figure 2C**). These results suggested that there were significant differences in the transcriptional regulation patterns of Mφ/Mo in different mouse AAA models. The intersection of upregulated and downregulated genes in the three datasets showed that six genes, *Gngt2*, *Il-1b*, *Lgals3*, *Spp1*, *Tgm2* as well as *Thbs1*, were upregulated and three genes, *Cbr2*, *Folr2* as well as *Mrc1*, were downregulated (**Figure 2G**). As shown in **Figure 2H**, the expression of these co-DEGs was showed in the three datasets, and their changes displayed certain

4 https://cran.r-project.org/web/packages/pheatmap/index.html

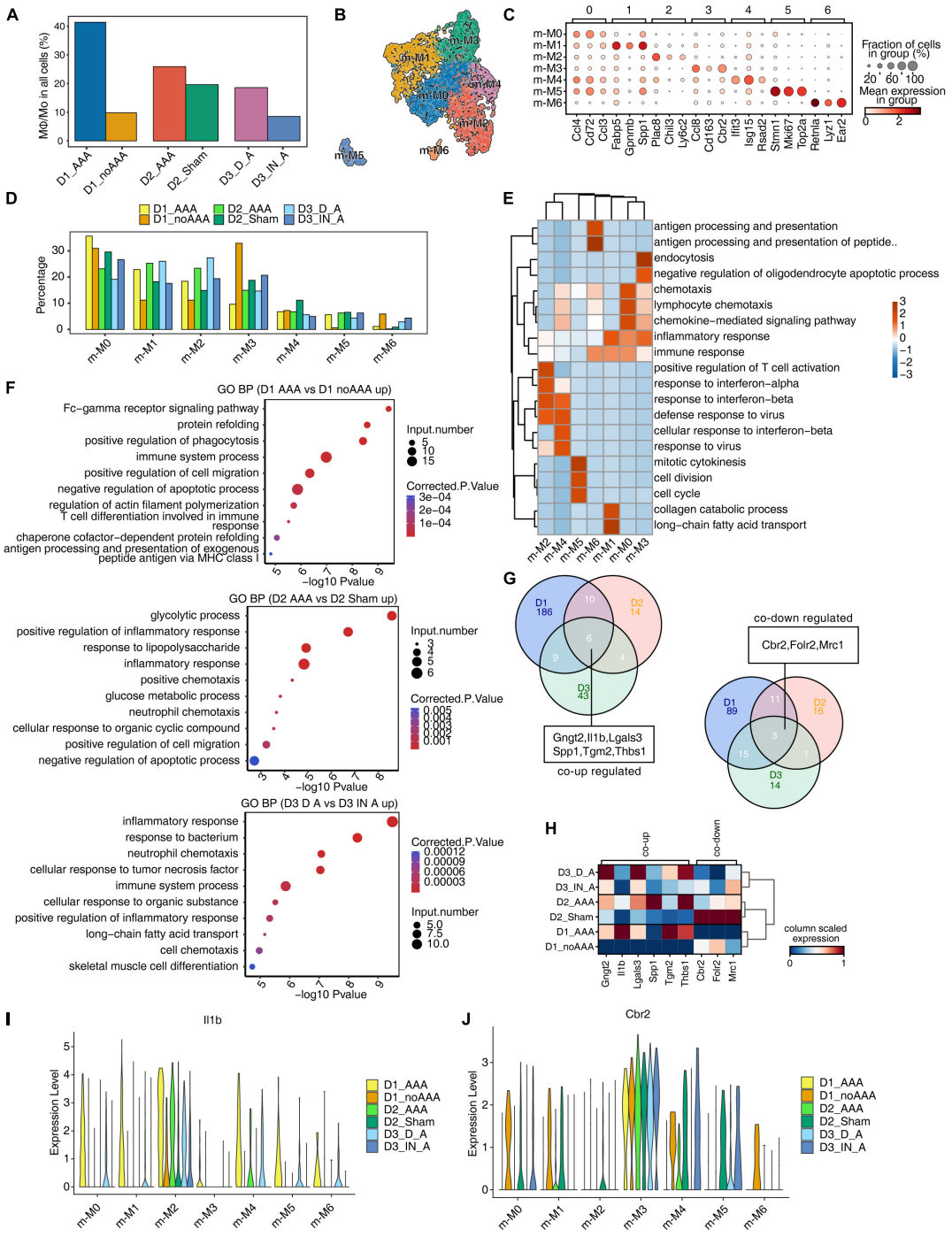


FIGURE 2
Single-cell analysis revealed complex macrophages/monocytes (Mφ/Mo) heterogeneity and conserved regulated genes between abdominal aortic aneurysm (AAA) and control samples. **(A)** Barplot represented the percentage of Mφ/Mo in total cells. **(B)** Uniform manifold approximation and projection (UMAP) plot of single-cell RNA sequencing (scRNA-seq) profile from Mφ/Mo separated into seven subtypes. Cells were colored according to different cell types. **(C)** Dot plot showing expression of top 3 markers in each subtype of Mφ/Mo. Color of dots represented the log fold change in each cluster comparing with other cells and dot size indicated percentage of cells in each cluster expressing the marker genes. **(D)** Bar plot comparing the proportions of cell populations of subtype of Mφ/Mo within each sample group. **(E)** Gene ontology enrichment analysis of biological processes of marker genes of each cell type. Top 3 terms were selected for each cluster and heatmap showed the enrichment *q*-value of these terms (scaled by column). **(F)** Gene ontology terms enriched in AAA versus in control Mφ/Mo for each dataset, respectively. The top 10 terms from upregulated genes are depicted as scatter plots displaying $-\log_{10}$ (*p*-value) and gene number. **(G)** Venn diagram showing the co-up (left) and co-down (right) regulated genes comparing AAA Mφ/Mo with control group from three datasets. **(H)** Unsupervised clustering heatmap showing relative expression (column scaled) levels of co-up genes and co-down genes in Mφ/Mo of each sample group. **(I,J)** Gene expression level of *Il-1b* **(I)** and *Cbr2* **(J)** were represented in the violin plot split by different sample groups.

heterogeneity in different Mφ/Mo subtypes (**Figures 2I, J** and **Supplementary Figures 2D, E**).

Different AAA models had similar inhibition characteristics of transcription factors

To explore the regulons of Mφ/Mo, we used pySCENIC to calculate regulon activity scores (RASs) in all Mφ/Mo and Seurat to construct the Mφ/Mo transcriptional regulation map. It's observed that the clusters displayed by regulons were generally consistent with Mφ/Mo subtypes but showed certain specificity (**Figure 3A**). It was also observed that different subgroups showed highly specific and different activation of regulons (**Supplementary Figure 3A**). By further comparing the actived and repressed regulons in AAA and control groups, it's found that the actived regulons had strong specificity, and only one co-actived regulon *Gm14327* was found in the AAA group of the three datasets (**Supplementary Figure 3B**). However, the repressed regulons were showed high consistency and a total of four regulons were detected in three datasets, including *Dbp*, *Sp1*, *Tcf4*, *Zfp275* (**Figure 3B**). As shown in **Figure 3C**, the co-actived and co-repressed regulons in the intersection of two or three databases showed different expression levels in the control group and the AAA group. In order to further confirm the regulatory functions of these regulons, we constructed regulatory networks of the co-varied regulons and co-DEGs. The co-upregulated genes as well as actived regulons and co-downregulated genes as well as repressed regulons were closely correlated (**Figure 3D**). These results suggested that co-varied regulons of Mφ/Mo had existed in different mouse AAA models and formed regulatory networks with co-DEGs, which played an important role in the occurrence and development of AAA.

ScRNA-seq of human AAA showed the increase of macrophages but not monocytes

It may be difficult to obtain human AAA tissue samples. So far, few article has reported AAA single-cell data in humans. Model comparisons between animals and humans were critical to understand the pathogenicity and cell-specific regulatory factors they shared. We first downloaded the datasets from GEO and reanalyzed it (14). The clustering and annotation results showed that the cell composition of the AAA group and the control group had a very large specificity and divided into 23 cell clusters (**Supplementary Figure 4A**). Each cell type had specific representative genes and expression of top 3 marker genes (**Supplementary Figure 4B**). The relative proportions of cell clusters in human AAA and control group, and the proportions of cell populations within each sample group had obvious differences (**Supplementary Figures 4C, D**). Mφ/Mo, specifically expressing CD14, were extracted for clustering and annotation again (**Figure 4A**). Further, h-M0, h-M1, and h-M4 were annotated as macrophages, while h-M2 and h-M3 were annotated as monocytes (**Figure 4B**). Each subtype of Mφ/Mo had representative marker genes (**Figure 4C**). By comparing the relative proportion of these five subpopulations of Mφ/Mo, it's found that the proportion of three macrophage subpopulations increased in AAA, while the proportion of two monocyte subpopulations decreased (**Figure 4D**). Also, different subpopulations of Mφ/Mo had their own specific GO enrichment of biological processes, which represented a different role in the development of AAA. The h-M0 was mainly enriched in synapse pruning and macrophage migration. The h-M1 was mainly enriched in immune response, inflammatory response and chemokine-mediated signaling pathway. The h-M3 was mainly enriched in immune response and neutrophil degranulation. The h-M2 and h-M4 were mainly enriched in neutrophil degranulation, but h-M4 was also enriched in endocytosis, regulation of macrophage migration (**Supplementary Figure 4E**).

*IL-1*B and *THBS1* were most conserved regulated genes in Mφ/Mo during AAA development

To compare with the mouse model, we firstly analyzed the genes of human AAA in Mφ/Mo that were differentially expressed from the control group (**Figure 5A**). The upregulated genes in the AAA group were significantly enriched in immune, inflammation, proliferation, and apoptosis related pathways (**Figure 5B**). The downregulated genes were mainly concentrated in neutrophil degranulation and translation related pathways (**Figure 5C**). The upregulated and downregulated genes in human AAA dataset and three mouse AAA datasets were intersected. Two genes, *IL-1*B and *THBS1*, were co-upregulated in the four datasets (**Figure 5D**). However, there were no co-downregulated genes, and only two downregulated genes (*APOE*, *FOS*) were found in D1 and human dataset (**Supplementary Figure 5A**). The genes that were upregulated in any two of the four datasets were extracted and displayed their expression in the human AAA and control samples. It's found that most genes showed high expression in AAA group, indicating that the expression trends of these genes were consistent in mouse AAA model and human AAA tissues (**Figures 5E, F**). Similarly, we extracted downregulated genes detected in any two of the four datasets and displayed their expression in human AAA and the control group. Different from the upregulated genes, most downregulated genes showed a trend of high expression in human AAA tissues (**Supplementary Figures 5B, C**). As shown in **Figures 5G, H**, it's found that *IL-1*B in macrophages subtypes h-M0 as well

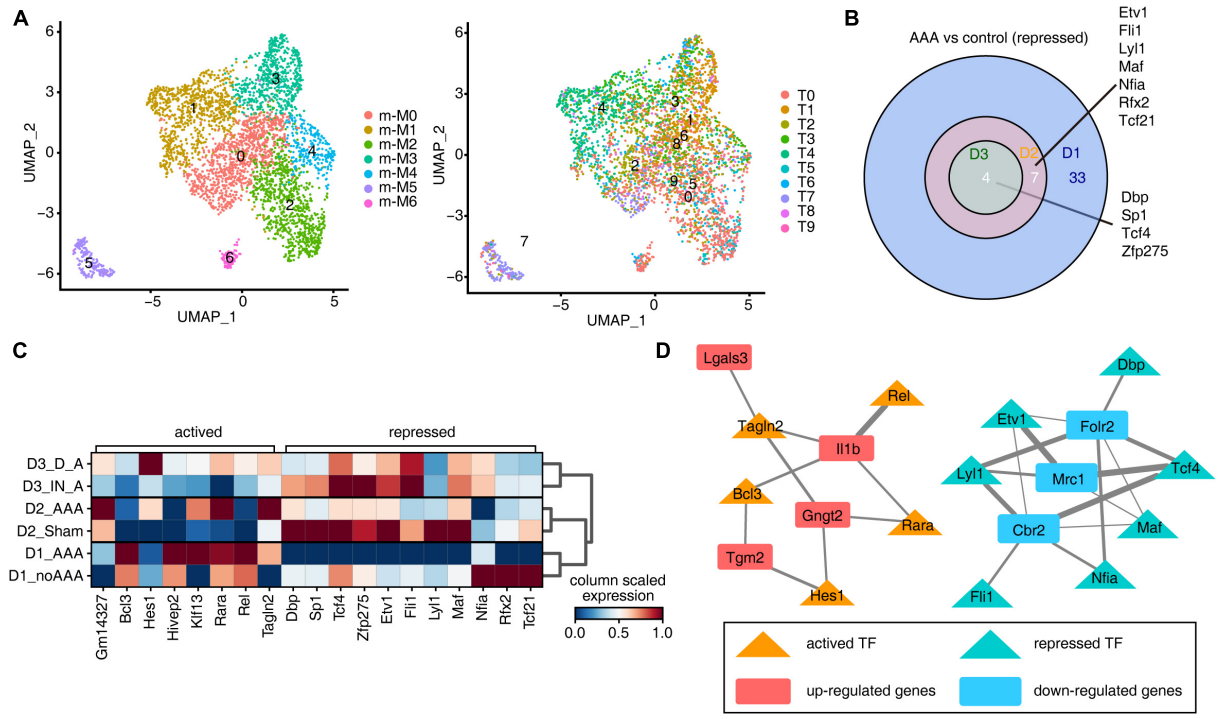


FIGURE 3
Different abdominal aortic aneurysm (AAA) models had similar inhibition characteristics of transcription factors. **(A)** Uniform manifold approximation and projection (UMAP) displayed the distributions of 10 cell types based on regulons (right panel). Colors indicated cell types. **(B)** Venn diagram showing the co-repressed regulons comparing AAA macrophages/monocytes (Mφ/Mo) with control group from three datasets. **(C)** Unsupervised clustering heatmap displayed the active states of co-actived and co-repressed regulons in each sample groups. Red colors indicated that the network was more activated, blue colors indicated that the network was more silenced. **(D)** Cytoscape showed the regulatory networks comprising transcription factors from panel **(C)** and their target co-differentially expressed genes (DEGs) underlying Mφ/Mo. Edges connected transcription factor-target gene pairs while nodes represented genes.

as h-M1, and *THBS1* in h-M0, h-M1 as well as h-M4 were significantly increased in AAA. These results suggested that *IL-1*B and *THBS1* were most conserved regulated genes and played a role in the involvement of Mφ/Mo in promoting the development of AAA both in mouse and humans.

Discussion

Abdominal aortic aneurysm is a life-threatening disease and there is currently a lack of effective treatment to prevent it rupturing. The mammalian abdominal aorta is composed of a large number of multifunctional cell populations, and each cell cluster has a distinct relationship with AAA. Several studies have used the scRNA-seq technique to characterize the heterogeneity of vascular cells, including VSMCs (29), endothelial cells (ECs) (30), macrophages (31), and aortic advection cells (32) in healthy and atherosclerotic arteries. However, studies on the cellular heterogeneity and aneurysm-related transcriptional signatures during AAA development are still deficient. Here, based on the published scRNA-seq datasets of GSE152583, GSE164678, GSE191226, and GSE166676 from different mouse AAA models and human AAA tissues, we explored the cellular heterogeneity and conserved transcriptional regulation patterns of Mφ/Mo during AAA to elucidate potential critical roles of certain genes in the pathogenesis of AAA. In the study, 26 cell clusters were obtained and 11 different cell types were identified by their markers in the three AAA mouse models. The proportion variation and function of each cell subtype were both consistent and different (**Figure 1** and **Supplementary Figure 1**). We further studied the heterogeneity of Mφ/Mo and conserved regulated genes between AAA and control samples. The proportion of Mφ/Mo increased in the AAA group. Mφ/Mo was divided into seven subtypes and the relative proportion changes of these subtypes were highly dynamic and heterogeneous. Moreover, the enrichment function of Mφ/Mo subtypes and DEGs had similarities as well as differences. In addition, there were co-downregulated genes, *Gngt2*, *Il-1b*, *Lgals3*, *Spp1*, *Tgm2* as well as *Thbs1*, and co-upregulated genes, *Cbr2*, *Folr2* as well as *Mrc1*, in the three datasets (**Figure 2** and **Supplementary Figure 2**). Next, we explored the regulons of Mφ/Mo, the cluster groups displayed by

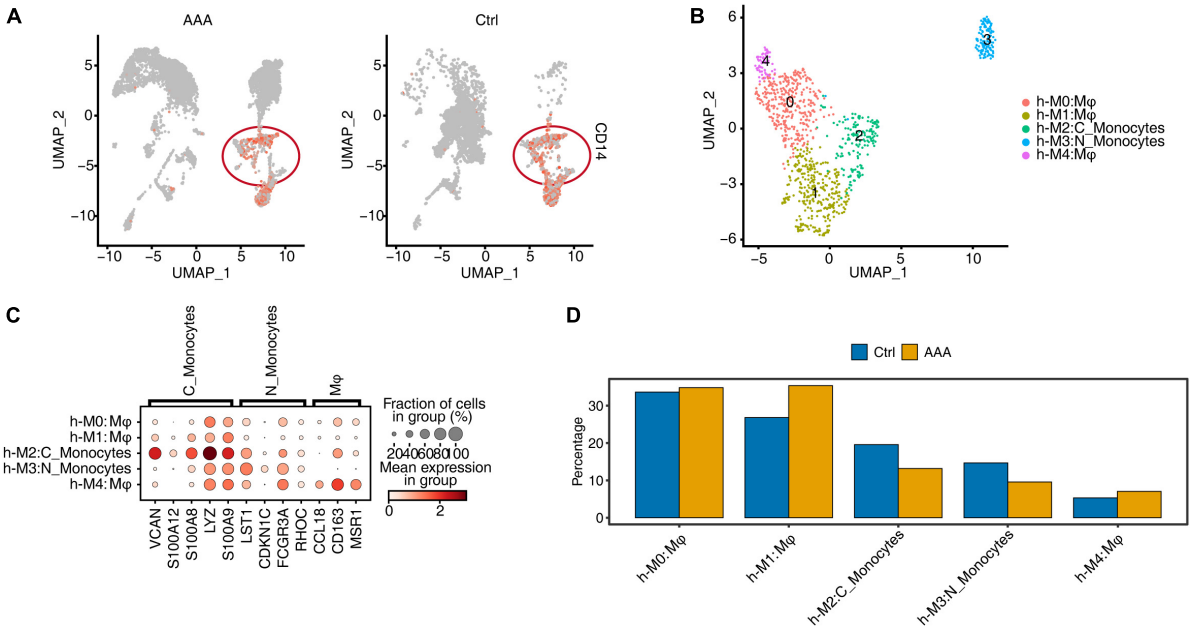


FIGURE 4
Single-cell RNA sequencing (ScRNA-seq) of human abdominal aortic aneurysm (AAA) showed the increase of macrophage cells but not monocytes. **(A)** Gene expression level of CD14 were represented in the uniform manifold approximation and projection (UMAP) plot split by human AAA and control group. **(B)** UMAP plot of scRNA-seq profile from macrophages/monocytes (Mφ/Mo) was separated into five subtypes. Cells were colored according to different cell types. C_Monocytes, classical monocytes; N_Monocytes, non-classical monocytes. **(C)** Dot plot showing expression of representative markers in each subtype of Mφ/Mo. Color of dots represented the log fold change in each cluster comparing with other cells and dot size indicated percentage of cells in each cluster expressing the marker genes. **(D)** Bar plot comparing the proportions of cell populations of subtype of Mφ/Mo within each sample group.

regulons were generally consistent with Mφ/Mo subtypes but showed certain specificity. The activated regulons had strong specificity but the repressed regulons had shown high consistency. The co-upregulated genes as well as actived regulons and co-downregulated genes as well as repressed regulons were significantly correlated (**Figure 3** and **Supplementary Figure 3**). The Mφ/Mo in human AAA tissue was divided into five subtypes. The proportion of three macrophage subpopulations increased in AAA, while the proportion of two monocyte subpopulations decreased. The different subpopulations had their own specific functions, indicating that scRNA-seq of human AAA showed the increase of macrophages but not monocytes (**Figure 4** and **Supplementary Figure 4**). Finally, the upregulated and downregulated genes in the human AAA dataset and the three mouse AAA datasets were intersected. Two genes were co-upregulated in the four datasets, including *IL-1B* and *THBS1* (**Figure 5** and **Supplementary Figure 5**). In summary, *in silico* analysis of scRNA-seq revealed that Mφ/Mo and its regulatory related genes as well as interaction networks played an important role in the pathogenesis of AAA.

Single-cell RNA sequencing is a contemporary and powerful technique for determining transcriptome gene profiles at the cellular level. ScRNA-seq has been recently used by many researchers to study the transcriptome profiles of aortic aneurysm tissues in humans and experimental animals at single-cell resolution (14, 20–22, 33). A comprehensive and unbiased genetic analysis by scRNA-seq can lead to a better understanding of cell-specific molecular signatures of AAA under physiological and pathophysiological conditions (34). At present, no single animal model can accurately reflect the full spectrum of human AAA pathophysiology (20). Therefore, further exploring the scRNA-seq datasets of different AAA models and human specimens will help better understand the pathogenesis of this disease from multiple perspectives and provide reference for clinical intervention. The scRNA-seq datasets (D1, D2, D3) from the three mouse AAA models have consistently demonstrated the heterogeneity of AAA tissue cells. Mφ/Mo infiltration in the adventitia of aneurysm tissue was a significant change during AAA progression. Mφ/Mo in the vessel wall had multiple functions, including amplification of the local inflammatory responses through secretion of proinflammatory cytokines, chemokines, and production of proteases and reactive oxygen species (5). Based on the scRNA-seq data herein, 26 cell clusters were obtained and 11 different cell types were identified. The proportion of different cell

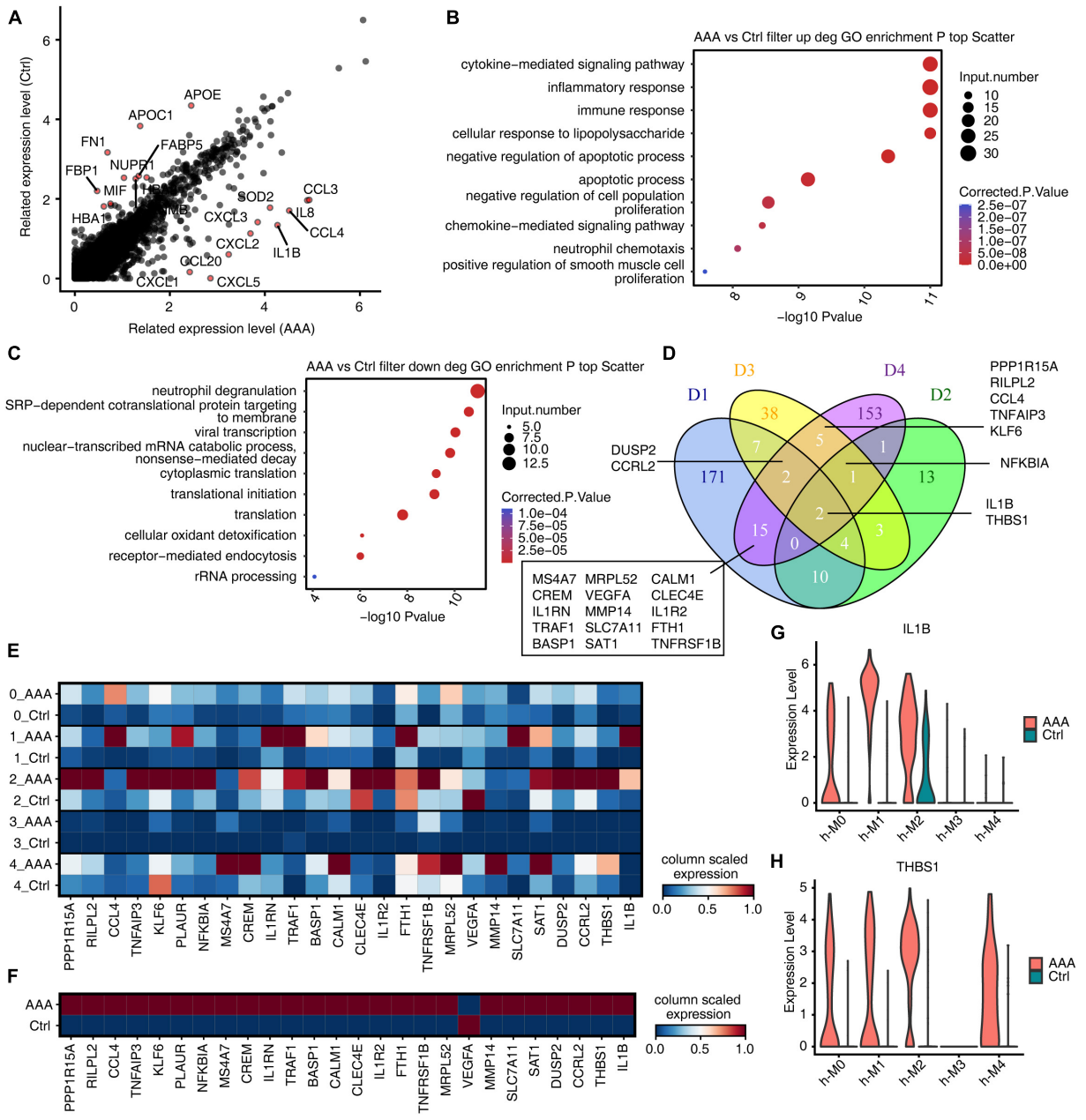


FIGURE 5
IL-1B and *THBS1* were most conserved regulated genes in macrophages/monocytes (Mφ/Mo) during abdominal aortic aneurysm (AAA) development. **(A)** Scatter plot displayed the differentially expressed genes in Mφ/Mo. Red colors indicated that the top 10 upregulated or downregulated genes. **(B)** Gene ontology terms of upregulated genes in AAA versus in control Mφ/Mo from human AAA dataset. The top 10 terms from upregulated genes were depicted as scatter plots displaying $-\log_{10}$ (*p*-value) and gene number. **(C)** Gene ontology terms of downregulated genes in AAA versus in control Mφ/Mo from human AAA dataset. The top 10 terms from upregulated genes were depicted as scatter plots displaying $-\log_{10}$ (*p*-value) and gene number. **(D)** Venn diagram showing the co-upregulated genes comparing AAA Mφ/Mo with control group from four datasets (D1–D3: mouse; D4: human). **(E)** Unsupervised clustering heatmap showing relative expression (column scaled) levels of upregulated genes at least in two datasets (include D4) split by human AAA and control group. **(F)** Same as panel **(E)** except that expression was showed in different subgroups of human AAA and control group. **(G, H)** Gene expression level of *IL-1B* and *THBS1* were represented in the violin plot split by human AAA and control group.

groups between AAA group and control group showed different trends. The proportion of cells such as Mφ/Mo, DC and neuron increased and the proportion of VSMC and fibroblast decreased in AAA. Monocytes and macrophages played a critical role in vascular injury and AAA formation. Macrophages were mainly derived from circulating monocytes

and the main inflammatory cell types in AAA lesions (5). Monocytes adhesion, migration, and MMP-9 production all increased in AAA patients, leading to aneurysm expansion (35). Mφ/Mo were extracted for analysis and a total of 7 subtypes were defined, which expressed diverse factors and were highly dynamic and heterogeneous. For example, m-M0 highly expressed chemokines, such as *Ccl4*, *Cd72*, and *Ccl3*, which was related to chemotactic biological function. m-M3 highly expressed *Ccl8*, *Cd163*, and *Folr2*, which was related to endocytosis, oligodendrocyte apoptotic process, and inflammatory. The subtypes of Mφ/Mo and its expressed varied genes performed different biological functions and participated in mouse AAA. Further analysis revealed that some genes, *Gngt2*, *Il-1b*, *Lgals3*, *Spp1*, *Tgm2* as well as *Thbs1*, were upregulated in three datasets, which was consistent with numerous reports in the literature. *Il-1b* was a proinflammatory cytokine, but it effected AAA formation as well as macrophage polarization (13) and treatment with anti- *Il-1a* or anti-*Il-1b* mAb blocked LCWE-induced AAA formation (36). Transglutaminase 2 (*Tgm2*) expression and activity in AAA formation were enhanced and had a potential role of ECM protector in aortic walls during AAA remodeling (37). *THBSs* overexpression may affect the formation of matrix cells and inhibit the activity of matrix proteins, thus destroying the structure of extracellular matrix and affecting the AAA occurrence (38). Yang predicted commonly altered signaling pathways by using intercellular communication networks in different experimental AAA models and human AAA, with a particular focus on *THBS* signaling among different cell populations (39). These results suggested that these genes were vital in the occurrence and development of AAA and worth further exploring.

The co-upregulated genes as well as actived regulons and co-downregulated genes as well as repressed regulons from Mφ/Mo were significantly correlated and formed regulatory networks. There were differences in the number and regulatory correlation of co-varied regulons and co-DEGs interactions, which played important and different roles in the formation of AAA by affecting the function of Mφ/Mo. To explore variation of macrophages and monocytes in human AAA tissue samples, it's found that the cell composition had a very large specificity between the AAA group and the control group, and immune cells were greatly amplified in the AAA group. The proportion of three macrophage subtypes (h-M0, h-M1, and h-M4) increased in AAA, while the proportion of two monocyte subtypes (h-M2 and h-M3) decreased. The pathways enriched in each subtype were also not same. The h-M0, h-M1, and h-M4 were mainly related to macrophage migration, immune inflammatory response and endocytosis; the h-M2 and h-M3 were mainly related to immune response and neutrophil degranulation, indicating that macrophages played a more important role than monocytes in human AAA. To further analyze the scRNA-seq datasets of Mφ/Mo from both mouse and human, it's found that two genes, *IL-1B* and *THBS1*, were co-upregulated and no gene was co-downregulated in the four datasets. This fully demonstrated the critical role of *IL-1B* and *THBS1* in AAA, which is consistent with the previous discussion and related literature (36, 39). The formation mechanisms of AAA induced by elastase, $CaCl_2$ as well as Ang II were different, and none of them can completely replace human AAA process. However, there were two upregulated DEGs, *IL-1B* and *THBS1*, in all three mouse AAA models and human AAA samples, indicating that they played a vital role in the commonly critical key link of AAA formation, which needed further exploration.

However, there were some limitations to our study. First of all, due to the reanalysis of the original datasets, the quality of them was also evaluated, but it was difficult to closely evaluate the reliability of the original samples, such as modeling quality, specimen collection process, and data sequencing. In the D3 dataset, no healthy control group data was provided, and the analysis of IN_A as a remission/rescue group instead of control group may have an effect on gene variation. Second, human AAA specimens were relatively few and existed individual differences, which required more human data support. Third, although we used the latest algorithms and other tools for evaluation, it may still cause some errors in the actual situation. Therefore, further studies were needed to provide more direct evidence for the role of Mφ /Mo in AAA.

In conclusion, this was the first study to compare the regulation of gene expression of Mφ/Mo in different mouse AAA models at the single-cell level. Our analysis revealed that co-DEGs of Mφ/Mo in the three mouse models played a critical role in the development of AAA. Moreover, we were the first to analyze and compare the transcriptional regulatory networks in different mouse AAA models. The co-varied regulons constituted the closely interaction regulatory networks with co-DEGs, regulating macrophage endocytosis, proliferation, and apoptosis. In addition, we determined the similarities and differences of the genes in the four datasets by comparing scRNA-seq datasets of three mouse AAA models and human AAA sample. In particular, *IL-1*B and *THBS1* were co-upregulated genes obtained from all four datasets and worthy of attention. These comparisons allowed us to show the cell classification, gene expression, and transcriptional regulatory networks in the current AAA models, which made us better grasp the similarities and differences of the models at the molecular level, and

also provided a new idea for the development of animal models in line with human AAA as well as targeted interventions for AAA.

Data availability statement

The original contributions presented in this study are included in the article/**Supplementary material**, further inquiries can be directed to the corresponding authors.

Author contributions

SW, NC, and XG proposed and designed this research. SW wrote this manuscript. SW, HZ, CC, and SL participated in data analysis. CC, BW, and ML participated in the design of the study. SW, HZ, and XG reviewed and edited the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This work was financially supported by the Joint Co-construction Project of Henan Medical Science and Technology Research Plan (No. LHGJ20200342), the Key Research & Development and Promotion of Special Project (Scientific Problem Tackling) of Henan Province (No. 202102310122), and the Medical Science and Technology Research Project (co-constructed by province and ministry) of Henan Province (No. SB201901009).

Conflict of interest

CC was employed by Wuhan Ruixing Biotechnology Co., Ltd.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcvm.2022.1062106/full#supplementary-material

SUPPLEMENTARY FIGURE 1
Single-cell RNA sequencing (ScRNA-seq) analysis of abdominal aortic tissue from different mouse abdominal aortic aneurysm (AAA) models identified distinct macrophages/monocytes (Mφ/Mo) types. **(A)** Uniform manifold approximation and projection (UMAP) plot split by different sample groups. **(B)** Dot plot showing expression of top 3 marker genes in each cell type. **(C)** Bar plot comparing the proportions of cell populations of each cell type within each sample group. **(D)** Gene ontology enrichment analysis of biological processes of top 100 marker genes of each cell type. Top 3 terms were selected for each cluster and heatmap shows the enrichment *q*-value of these terms (scaled by column).

SUPPLEMENTARY FIGURE 2
Single-cell analysis revealed complex macrophages/monocytes (Mφ/Mo) heterogeneity and conserved regulated genes between abdominal aortic aneurysm (AAA) and control samples. **(A)** Uniform manifold approximation and projection (UMAP) visualization of the Mφ/Mo split by different sample groups. **(B)** Rank order based on decreasing values of the relative frequency ratio between AAA and control sample group in three datasets. **(C)** Gene ontology terms of downregulated genes in AAA versus in control Mφ/Mo for each dataset, respectively. The top 10 terms from upregulated genes were depicted as scatter plots displaying $-\log_{10}$ (*p*-value) and gene number. **(D,E)** Gene expression level of *Spp1* **(D)** and *Mrc1* **(E)** were represented in the violin plot split by different sample groups.

SUPPLEMENTARY FIGURE 3
Different abdominal aortic aneurysm (AAA) models had similar inhibition characteristics of transcription factors. **(A)** Unsupervised clustering heatmap displayed the active states of cluster-specific regulons in each subtype of macrophages/monocytes (Mφ/Mo). Red colors indicated that the network was more activated, blue colors indicated that the network was more silenced. **(B)** Venn diagram showing the co-activated regulons comparing AAA Mφ/Mo with control group from three datasets.

SUPPLEMENTARY FIGURE 4
Single-cell RNA sequencing (ScRNA-seq) of human abdominal aortic aneurysm (AAA) showed the increase of macrophage cells but not monocytes. **(A)** Uniform manifold approximation and projection (UMAP) plot of composite single-cell transcriptomic profiles from all human AAA and control groups. Colors indicated cell clusters along with annotations. **(B)** Dot plot showing expression of top 3 marker genes in each cell type. **(C)** Stacked bar plot showing the relative proportions of cell clusters in human AAA and control group. **(D)** Bar plot comparing the proportions of cell populations of cell clusters within each sample group. **(E)** Gene ontology enrichment analysis of biological processes of marker genes of each subgroup of macrophage/monocyte cells. Top 3 terms were selected for each cluster and heatmap shows the enrichment *q*-value of these terms (scaled by column).

SUPPLEMENTARY FIGURE 5
*IL-1*B and *THBS1* were most conserved regulated genes in macrophages/monocytes (Mφ/Mo) during abdominal aortic aneurysm (AAA) development. **(A)** Venn diagram showing the co-downregulated genes comparing AAA Mφ/Mo with control group from four datasets (D1–D3: mouse; D4: human). **(B)** Unsupervised clustering heatmap showing relative expression (column scaled) levels of downregulated genes at least in two datasets split by human AAA and control group. **(C)** Unsupervised clustering heatmap showing relative expression (column scaled) levels of downregulated genes at least in two datasets showed in panel **(A)**, cells were split by different subgroups of human AAA and control group.

SUPPLEMENTARY TABLE 1
Quality control form of datasets.

References

1. Sedrakyan A, Goodney P, Mao J, Beck A, Schermerhorn M. Changes in the long-term risk of adverse outcomes in patients treated with open vs endovascular abdominal aortic aneurysm repair. *JAMA Surg.* (2022) 157:733–5. doi: 10.1001/jamasurg.2022.1070

2. Wu S, Liu S, Chen N, Zhang C, Zhang H, Guo X. Genome-wide identification of immune-related alternative splicing and splicing regulators involved in abdominal aortic aneurysm. *Front Genet.* (2022) 13:816035. doi: 10.3389/fgene.2022.816035

3. Sakalihasan N, Limet R, Defawe O. Abdominal aortic aneurysm. *Lancet.* (2005) 365:1577–89. doi: 10.1016/S0140-6736(05)66459-8

4. Liu B, Granville D, Golledge J, Kassiri Z. Pathogenic mechanisms and the potential of drug therapies for aortic aneurysm. *Am J Physiol Heart Circ Physiol.* (2020) 318:H652–70. doi: 10.1152/ajpheart.00621.2019

5. Raffort J, Lareyre F, Clément M, Hassen-Khodja R, Chinetti G, Mallat Z. Monocytes and macrophages in abdominal aortic aneurysm. *Nat Rev Cardiol.* (2017) 14:457–71. doi: 10.1038/nrcardio.2017.52

6. Katsuki S, Koga J, Matoba T, Umezu R, Nakashiro S, Nakano K, et al. Nanoparticle-mediated delivery of pitavastatin to monocytes/macrophages inhibits angiotensin ii-induced abdominal aortic aneurysm formation in apoe mice. *J Atheroscler Thromb.* (2022) 29:111–25. doi: 10.5551/jat.54379

7. Li H, Bai S, Ao Q, Wang X, Tian X, Li X, et al. Modulation of immune-inflammatory responses in abdominal aortic aneurysm: emerging molecular targets. *J Immunol Res.* (2018) 2018:7213760. doi: 10.1155/2018/7213760

8. Hadi T, Boytard L, Silvestro M, Alebrahim D, Jacob S, Feinstein J, et al. Macrophage-derived netrin-1 promotes abdominal aortic aneurysm formation by activating MMP3 in vascular smooth muscle cells. *Nat Commun.* (2018) 9:5022. doi: 10.1038/s41467-018-07495-1

9. DeRoo E, Stranz A, Yang H, Hsieh M, Se C, Zhou T. Endothelial dysfunction in the pathogenesis of abdominal aortic aneurysm. *Biomolecules.* (2022) 12:509. doi: 10.3390/biom12040509

10. Lysgaard Poulsen J, Stubbe J, Lindholt J. Animal models used to explore abdominal aortic aneurysms: a systematic review. *Eur J Vasc Endovasc Surg.* (2016) 52:487–99. doi: 10.1016/j.ejvs.2016.07.004

11. Lu H, Du W, Ren L, Hamblin M, Becker R, Chen Y, et al. Vascular smooth muscle cells in aortic aneurysm: from genetics to mechanisms. *J Am Heart Assoc.* (2021) 10:e023601. doi: 10.1161/JAHA.121.023601

12. Yao F, Yao Z, Zhong T, Zhang J, Wang T, Zhang B, et al. Imatinib prevents elastase-induced abdominal aortic aneurysm progression by regulating macrophage-derived MMP9. *Eur J Pharmacol.* (2019) 860:172559. doi: 10.1016/j.ejphar.2019.172559

13. Batra R, Suh M, Carson J, Dale M, Meisinger T, Fitzgerald M, et al. IL-1β (interleukin-1β) and TNF-α (tumor necrosis factor-α) impact abdominal aortic aneurysm formation by differential effects on macrophage polarization. *Arterioscler Thromb Vasc Biol.* (2018) 38:457–63. doi: 10.1161/ATVBAHA.117.310333

14. Davis F, Tsoi L, Melvin W, denDekker A, Wasikowski R, Joshi A, et al. Inhibition of macrophage histone demethylase JMJD3 protects against abdominal aortic aneurysms. *J Exp Med.* (2021) 218:e20201839. doi: 10.1084/jem.20201839

15. Xie C, Ye F, Zhang N, Huang Y, Pan Y, Xie X. CCL7 contributes to angiotensin II-induced abdominal aortic aneurysm by promoting macrophage infiltration and pro-inflammatory phenotype. *J Cell Mol Med.* (2021) 25:7280–93. doi: 10.1111/jcmm.16757

16. Birnbaum K. Power in numbers: single-Cell RNA-Seq strategies to dissect complex tissues. *Annu Rev Genet.* (2018) 52:203–21. doi: 10.1146/annurev-genet-120417-031247

17. Shao X, Lu X, Liao J, Chen H, Fan X. New avenues for systematically inferring cell-cell communication: through single-cell transcriptomics data. *Protein Cell.* (2020) 11:866–80. doi: 10.1007/s13238-020-00727-5

18. Potter S. Single-cell RNA sequencing for the study of development, physiology and disease. *Nat Rev Nephrol.* (2018) 14:479–92. doi: 10.1038/s41581-018-0021-7

19. Seumois G, Vijayanand P. Single-cell analysis to understand the diversity of immune cell types that drive disease pathogenesis. *J Allergy Clin Immunol.* (2019) 144:1150–3. doi: 10.1016/j.jaci.2019.09.014

20. Yang H, Zhou T, Stranz A, DeRoo E, Liu B. Single-Cell RNA sequencing reveals heterogeneity of vascular cells in early stage murine abdominal aortic aneurysm-brief report. *Arterioscler Thromb Vasc Biol.* (2021) 41:1158–66. doi: 10.1161/ATVBAHA.120.315607

21. Yu L, Zhang J, Gao A, Zhang M, Wang Z, Yu F, et al. An intersegmental single-cell profile reveals aortic heterogeneity and identifies a novel Malat1 vascular smooth muscle subtype involved in abdominal aortic aneurysm formation. *Signal Transduct Target Ther.* (2022) 7:125. doi: 10.1038/s41392-022-00943-x

22. Zhao G, Lu H, Chang Z, Zhao Y, Zhu T, Chang L, et al. Single-cell RNA sequencing reveals the cellular heterogeneity of aneurysmal infrarenal abdominal aorta. *Cardiovasc Res.* (2021) 117:1402–16. doi: 10.1093/cvr/cvaa214

23. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol.* (2018) 36:411–20. doi: 10.1038/nbt.4096

24. Wilson P, Wu H, Kirita Y, Uchimura K, Ledru N, Rennke H, et al. The single-cell transcriptomic landscape of early human diabetic nephropathy. *Proc Natl Acad Sci USA.* (2019) 116:19619–25. doi: 10.1073/pnas.1908706116

25. Chung J, Goldstein L, Chen Y, Lee J, Webster J, Roose-Girma M, et al. Single-cell transcriptome profiling of the kidney glomerulus identifies key cell types and reactions to injury. *J Am Soc Nephrol.* (2020) 31:2341–54. doi: 10.1681/ASN.2020020220

26. Aibar S, Gonzalez-Blas C, Moerman T, Huynh-Thu V, Imrichova H, Hulselmans G, et al. SCENIC: single-cell regulatory network inference and clustering. *Nat Methods.* (2017) 14:1083–6. doi: 10.1038/nmeth.4463

27. Xie C, Mao X, Huang J, Ding Y, Wu J, Dong S, et al. KOBAS 2.0: a web server for annotation and identification of enriched pathways and diseases. *Nucleic Acids Res.* (2011) 39:316–22. doi: 10.1093/nar/gkr483

28. Davis F, Gallagher K. Epigenetic mechanisms in monocytes/macrophages regulate inflammation in cardiometabolic and vascular disease. *Arterioscler Thromb Vasc Biol.* (2019) 39:623–34. doi: 10.1161/ATVBAHA.118.312135

29. Dobnikar L, Taylor A, Chappell J, Oldach P, Harman J, Oerton E, et al. Disease-relevant transcriptional signatures identified in individual smooth muscle cells from healthy mouse vessels. *Nat Commun.* (2018) 9:4567. doi: 10.1038/s41467-018-06891-x

30. Kalluri A, Vellarikkal S, Edelman E, Nguyen L, Subramanian A, Ellinor P, et al. Single-cell analysis of the normal mouse aorta reveals functionally distinct endothelial cell populations. *Circulation.* (2019) 140:147–63. doi: 10.1161/CIRCULATIONAHA.118.038362

31. Cochain C, Vafadarnejad E, Arampatzi P, Pelisek J, Winkels H, Ley K, et al. Single-cell RNA-seq reveals the transcriptional landscape and heterogeneity of aortic macrophages in murine atherosclerosis. *Circ Res.* (2018) 122:1661–74. doi: 10.1161/CIRCRESAHA.117.312509

32. Gu W, Ni Z, Tan Y, Deng J, Zhang S, Lv Z, et al. Adventitial cell atlas of wt (wild type) and ApoE (apolipoprotein E)-deficient mice defined by single-cell RNA sequencing. *Arterioscler Thromb Vasc Biol.* (2019) 39:1055–71. doi: 10.1161/ATVBAHA.119.312399

33. Boytard L, Hadi T, Silvestro M, Qu H, Kumpfbeck A, Sleiman R, et al. Lung-derived HMGB1 is detrimental for vascular remodeling of metabolically imbalanced arterial macrophages. *Nat Commun.* (2020) 11:4311. doi: 10.1038/s41467-020-18088-2

34. Sawada H, Lu H, Daugherty A. Single-cell transcriptomics as a building block for determining mechanistic insight of abdominal aortic aneurysm formation. *Cardiovasc Res.* (2021) 117:1243–4. doi: 10.1093/cvr/cvab083

35. Pilecki B, de Carvalho P, Kirketerp-Møller K, Schlosser A, Kejling K, Dubik M, et al. MFAP4 deficiency attenuates angiotensin II-induced abdominal aortic aneurysm formation through regulation of macrophage infiltration and activity. *Front Cardiovasc Med.* (2021) 8:764337. doi: 10.3389/fcvm.2021.764337

36. Wakita D, Kurashima Y, Crother T, Noval Rivas M, Lee Y, Chen S, et al. Role of interleukin-1 signaling in a mouse model of kawasaki disease-associated abdominal aortic aneurysm. *Arterioscler Thromb Vasc Biol.* (2016) 36:886–97. doi: 10.1161/ATVBAHA.115.307072

37. Munezane T, Hasegawa T, Suritala, Tanaka A, Okada K, Okita Y. Activation of transglutaminase type 2 for aortic wall protection in a rat abdominal aortic aneurysm formation. *J Vasc Surg.* (2010) 52:967–74. doi: 10.1016/j.jvs.2010.04.049

38. Auguściak-Duma A, Lesiak M, Stępień K, Gutmajster E, Sieroń A. mRNA expression of thrombospondin 1, 2 and 3 from proximal to distal in human abdominal aortic aneurysm–preliminary report. *Acta Biochim Pol.* (2021) 68:745–50. doi: 10.18388/abp.2020_5645

39. Yang H, DeRoo E, Zhou T, Liu B. Deciphering cell-cell communication in abdominal aortic aneurysm from single-cell RNA transcriptomic data. *Front Cardiovasc Med.* (2022) 9:831789. doi: 10.3389/fcvm.2022.831789

frontiers | Frontiers in Cardiovascular Medicine

TYPE Review
PUBLISHED 25 April 2023
DOI 10.3389/fcvm.2023.1057870

OPEN ACCESS

EDITED BY
Sasha A. Singh,
Harvard Medical School, United States

REVIEWED BY
Gregory Quaife-Ryan,
The University of Queensland, Australia
Sarvesh Chelvanambi,
Harvard Medical School, United States

*CORRESPONDENCE
Jun Li
gamyylj@163.com
Xin Yi
yixin_bjxc@163.com
Yanwei Xing
xingyanwei12345@163.com

†These authors have contributed equally to this work and share first authorship

RECEIVED 03 October 2022
ACCEPTED 10 April 2023
PUBLISHED 25 April 2023

CITATION
Su X, Wang L, Ma N, Yang X, Liu C, Yang F, Li J, Yi X and Xing Y (2023) Immune heterogeneity in cardiovascular diseases from a single-cell perspective.
Front. Cardiovasc. Med. 10:1057870.
doi: 10.3389/fcvm.2023.1057870

Immune heterogeneity in cardiovascular diseases from a single-cell perspective

Xin Su[1†], Li Wang[2†], Ning Ma[3†], Xinyu Yang[4], Can Liu[1], Fan Yang[1], Jun Li[1*], Xin Yi[5*] and Yanwei Xing[1*]

[1]China Academy of Chinese Medical Sciences, Guang'anmen Hospital, Beijing, China, [2]Department of Breast Surgery, Xingtai People's Hospital, Xingtai, China, [3]Department of Breast Surgery, Dezhou Second People's Hospital, Dezhou, China, [4]Fangshan Hospital Beijing University of Chinese Medicine, Beijing, China, [5]Department of Cardiology, Beijing Huimin Hospital, Beijing, China

A variety of immune cell subsets occupy different niches in the cardiovascular system, causing changes in the structure and function of the heart and vascular system, and driving the progress of cardiovascular diseases (CVDs). The immune cells infiltrating the injury site are highly diverse and integrate into a broad dynamic immune network that controls the dynamic changes of CVDs. Due to technical limitations, the effects and molecular mechanisms of these dynamic immune networks on CVDs have not been fully revealed. With recent advances in single-cell technologies such as single-cell RNA sequencing, systematic interrogation of the immune cell subsets is feasible and will provide insights into the way we understand the integrative behavior of immune populations. We no longer lightly ignore the role of individual cells, especially certain highly heterogeneous or rare subpopulations. We summarize the phenotypic diversity of immune cell subsets and their significance in three CVDs of atherosclerosis, myocardial ischemia and heart failure. We believe that such a review could enhance our understanding of how immune heterogeneity drives the progression of CVDs, help to elucidate the regulatory roles of immune cell subsets in disease, and thus guide the development of new immunotherapies.

KEYWORDS

cardiovascular diseases, immune cell, heterogeneity, single cell RNA sequencing (scRNAseq), atherosclerosis

1. Introduction

Cardiovascular diseases (CVDs) have been recognized as the leading cause of global mortality and disability in human beings. Usual CVDs include ischemic heart disease, cardiomyopathies, arrhythmic disorders, heart failure, peripheral arterial disease, and other cardiac and vascular conditions (1, 2). Mensah et al. (3) used a vivid metaphor when explaining the global burden of cardiovascular disease in 2017: 17.8 million people died of CVDs is equivalent to reducing 330 million years of life and increasing 35.6 million years of disability. CVDs severely reduced quality of life and increased medical burden for many families. The cardiovascular system is primarily composed of cardiomyocytes, fibroblasts, endothelial cells, smooth muscle cells, pericytes, and stroma (4). Furthermore, various types of immune cells that inhabit or infiltrate cardiac tissue have been identified and characterized. All major leukocyte classes, populations of lymphocyte and myeloid origin, are present in cardiovascular tissue (5). Multiple immune cell subsets exist and occupy different niches in cardiac tissue, involving in communication with the cardiovascular system (6, 7). However, infiltrating immune cells

cause structural and functional changes in the heart and vasculature, and contribute to CVDs progression (8). Targeting the inflammatory cascade in animal models has been shown to be conducive to improve myocardial injury and promote healing. In mice with deletion of interleukin 10 (IL-10), diastolic function was improved. With the occurrence of diastolic dysfunction, myocardial macrophages produce IL-10 (9). Single-cell technology allows for a more precise focus on subtypes of cells, and some research studies have confirmed that resident macrophage-derived IL-10 promotes myocardial fibrosis (10). Blocking this process could inhibit the activation of fibroblasts and reduce collagen deposition. Clinical trial has also demonstrated that inflammation is an effective therapeutic target for secondary prevention of CVDs (11). Though the study of macrophages driving CVDs is highly discussed, the great diversity and abundance of other immune cells infiltrating the damaged areas, such as a large number of lymphocytes, indicate that there is a broader dynamic immune network controlling the dynamic CVDs niche. Due to technical limitations, the effects and molecular mechanisms of these dynamic immune networks on CVDs are not adequately revealed.

Advances in single-cell applications are reshaping the way we understand the integrative behavior of immune populations, no longer inadvertently confusing or concealing the unique contributions of individual cells, especially when behavior is highly heterogeneous or driven by rare cell types (12). The high parameterization and high throughput of single-cell applications allow for in-depth characterization of low-abundance cell populations and unbiased identification (13). The use of single-cell technology to understand how immune populations use cell diversity to achieve this breadth and flexibility, especially in dynamic processes such as transdifferentiation and antigen response, provides inspiring opportunities for studying immune heterogeneity in an unprecedented scope. Given these advances, we reviewed recent work on single-cell technology, discussed data on immune heterogeneity in CVDs, and explored how immune populations generated and exploited cellular heterogeneity at multiple molecular and phenotypic levels. Furthermore, we highlighted the advantages of single-cell applications in uncovering the antecedents and consequences of immune system heterogeneity. Lastly, the latest progress toward clinical application, the remaining challenges, and future perspectives on the development of single-cell technology are briefly discussed.

2. Advances in single-cell applications for immune heterogeneity

The first living cells were discovered in the 17th century, but it took more than two hundred years to realize that cells are not only the structural unit of life, but also the functional unit (14). The continued exploration of cellular heterogeneity then began (15, 16). Human cells carry nearly the same genetic material, but the transcriptome information of a single cell only reflects the unique activity of a subset of genes (17). Immune cells are characterized by their heterogeneity. Studying the immune system solely at the level of cell population may not offer a comprehensive understanding, as it is crucial to explore individual immune cells' communication and masking phenotypes at the single-cell level (18). By analyzing the genomic and proteomic profiles of individual immune cells, researchers can identify previously unrecognized subpopulations and cell-to-cell interactions that contribute to the immune responses. Single-cell technology has revealed the uniqueness of individual cells and addressed questions unobtainable in bulk analysis (19, 20). Single cell RNA sequencing (scRNA-seq) analyzes the transcriptome of millions of cells at single cell resolution, achieve genome-wide characterization of the degree of transcriptional differences between coding and non-coding RNAs and generate a comprehensive gene expression atlas (21). For instance, immune cell subgroups can be disaggregated by scRNA-seq to enable characterization of heterogeneous cell populations (22, 23). This new technology can amend the developmental trajectory and classification of immune cells in the human body. Traditional hematopoietic grading is mostly based on subjectively purified cell populations that are passively segmented into component regions. Dendritic cells (DCs) differentiate, migrate and respond to environmental stimuli with a wide heterogeneity of markers and transcriptome features, suggesting that DC cells are difficult to classify precisely with predetermined markers. scRNA-seq can identify a comprehensive transcriptomic description of transient and transcriptional states. scRNA-seq introduced a paradigm shift in the classification of human DCs and identified the existence of multiple phenotypically distinct subsets in addition to the three well-defined subsets, DC1–3 (24, 25).

scRNA-seq is reshaping the way we define the biological state of cells and is well suited to work with a limited sample size. Since the first technology was developed in 2009 (26), multiple single-cell methods have been iterated, methods for cell capture and amplification, transcript construction, and read depth per cell vary (27, 28). But in general, a similar process for scRNA-seq is: sample preparation (single-cell isolation), single-cell capture, cell lysis, reverse transcription and amplification (conversion of RNA into complementary DNA), library preparation, sequencing, and analysis (29), as shown in **Figure 1**. The point is to capture each individual cell in an isolated reaction mixture in which all transcripts from a single cell will be uniquely barcoded after conversion to complementary DNA (cDNA) (17). At the end of library construction, additional sample-specific barcodes are added enzymatically, allowing multiple biological samples to be sequenced simultaneously (30). This enables unique identification and access to the transcriptome of per single cell in every sample. Stoeckius et al. (31) described cellular indexing of transcriptomes and epitopes by high-throughput sequencing, a oligonucleotide-labeled antibodies phenotyping method, which is conducive to resolve distinct immune cell population with greater clarity.

The scRNA-seq data sets are high-dimensional, and analysis strategies should concentrate on highly variable genes and dimensionality reduction. Assigning biological annotations is a fundamental step that establishes the foundation for subsequent analysis of specific cell populations. In addition, functional

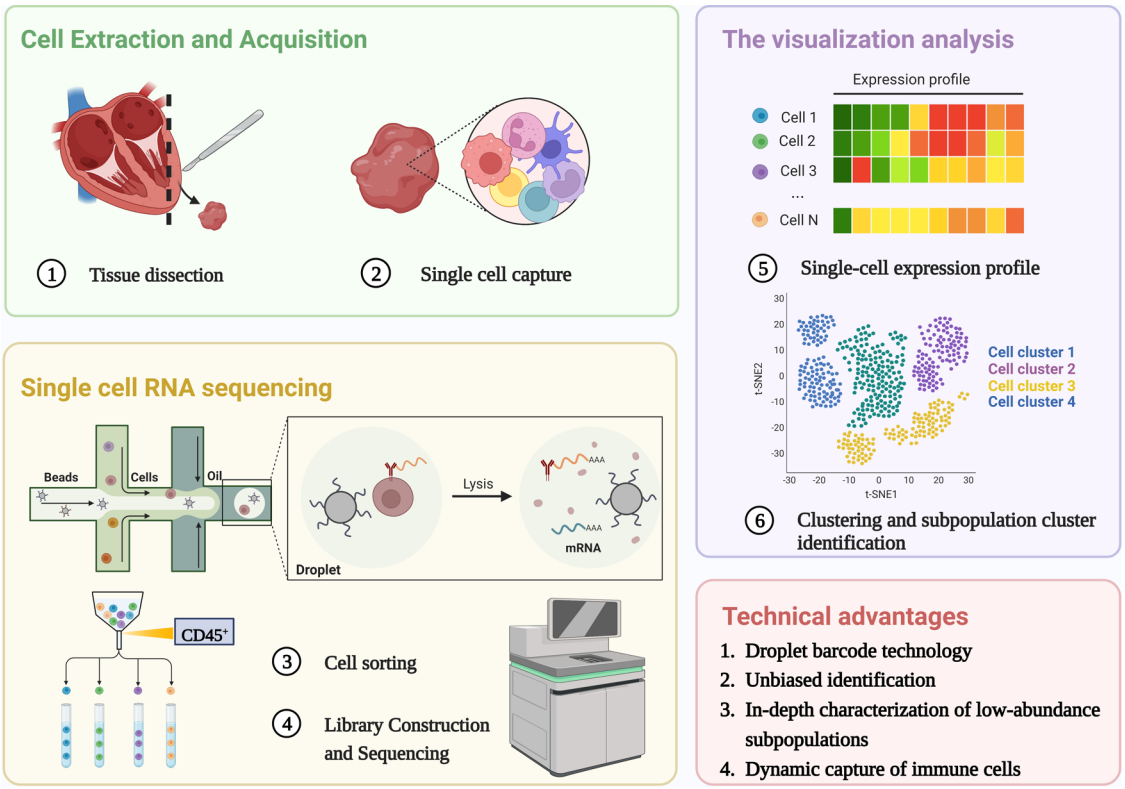


FIGURE 1
Single cell RNA sequencing workflow. The process shared by scRNA-Seq: single-cell isolation and capture, cell lysis, reverse transcription and amplification, library preparation, sequencing, and analysis. The unique advantages of scRNA seq in the study of immune heterogeneity: Application of barcode technology, unbiased identification and in-depth characterization of low abundance subpopulations, Dynamic capture of immune cells without ignoring transitional cell events. Created with BioRender.com.

enrichment plays a crucial role in accurately revealing the functional bias of these cell populations. Other exploratory analyses include pseudo-time analysis (32), cell-cell communication analysis (33), cell cycle prediction (34) and spatial transcriptome (35). It is noteworthy that studies on the heterogeneity of the immune system are best described using continuous models, and can only be approximated by grouping strategies that disperse cell phenotypes (13, 36). Traditional gating strategies to divide populations along any particular axis have the potential to mask the transitional cellular events that connect the complex environments. Given the asynchronous nature of the immune response, scRNA-seq generates static snapshots of the entire process which can be modeled as a coherence of transitional cellular states via capturing immune cells during a dynamic process. This kind of trajectory inference have provided a unique opportunity to track immune cells during differentiation and delineate lineage hierarchies (37). Ni et al. (38) investigated the heterogeneity of cardiac macrophages in mice with transverse aortic constriction by single-cell sequencing and identified macrophage subpopulations associated with cardiac injury. The pseudo-time trajectory confirmed that Rel gene was a key transcription factor driving CD 72^{high} macrophage differentiation which exerted a pro-inflammatory effect to induce cardiac injury. Furthermore, with the improved statistical power for large-scale single-cell datasets, it becomes possible to construct gene regulatory networks in dynamic processes by combining inferred trajectories and co-expression analyses (39). Zhuang et al. (40) analyzed the global characteristics and dynamics of single immune cells after myocardial infarction (MI), demonstrating not only dynamic inward flow of immune cells under ischemic conditions, but also identifying Fos/activator protein 1regulation as a key pro-inflammatory regulator. Multi-omics analysis further links transcriptional data with epigenetic characteristics, combines regulatory networks with clustering and trajectory inference, and constructs a regulatory landscape of the immune system (41–43). Chaffin et al. (44) formed the transcriptional landscape of the heart by analyzing left ventricular samples from dilated cardiomyopathy and hypertrophic cardiomyopathy at single-cell resolution. And They identified a reduction in the number of proliferating resident cardiac macrophages and transforming growth factor-β as the ultimate common transcriptional pathway in patients with cardiomyopathy. One such limitation is that cells of a single data set will show discrete clusters rather than continuous trajectories when dynamic trajectory analysis is performed—For example, based on the assumption that the underlying dynamic processes are stationary and ergodic, scRNA-seq captures a snapshot of cells at a particular point in time, so that scRNA-seq data with a large enough sample size can cover all possible intermediate states of a given population of

cells. However, the results may not meet expectations in the case of concomitant developmental time changes or external stimuli. scRNA-seq of immune cells may aggregate according to stress-responsive genes, but is mistaken for a different cell type. Highly important but lowly expressed genes were missed due to relatively low sequencing depth in single-cell studies. In addition, bicellular cells containing 2 different cell types may be misidentified as transitional cell types.

3. Immune cell populations in cardiac homeostasis

Studies on the immune system have typically focused on conditions of infection/injury, but it is becoming increasingly clear that leukocytes and secreted cytokines play a role in the growth and development of tissues as well as in healthy conditions (45–47). It is well known that there are a contingent of resident immune cells in the heart, which means that immune cells are not visitors (48–50), but rather early components of heart development. Transcriptional expression of long-term tissue-resident immune cell subsets correlate with tissue site. By residing in the tissues, these immune cells are involved in tissue homeostasis and repair. We must first understand the differences in cellular composition between healthy and diseased states. The immune system plays a vital role in regulating heart health and homeostasis. Elucidating the role of each immune cell subtype and their interactions is important for understanding the mechanism of "normal-damage-repair" in the heart.

The heart is composed of a small population of resident immune cells from the myeloid and lymphoid lineages. The heterogeneity of those cells has proven to be overly complex to characterize, and these cell populations from multiple lineages display a wide functional diversity and dynamically express molecular markers over time. Monocytes/macrophages are representative of cardiac immune cells. Macrophages, lymphocytes, dendritic cells and a small number of neutrophils cells constitute the mainstay of cardiac immune homeostasis (51, 52). Immune cell populations sense cues from the microenvironment and respond using specific gene expression profiles. Under homeostatic/pathogenic conditions, macrophages are activated and make decisions about immune regulation in conjunction with other immune cells. In this section, we will discuss the most relevant major immune cell types, as shown in **Figure 2**.

3.1. Macrophages

Macrophage phenotype is routinely defined as $CD45^+$ $CD11b^+$, and murine-derived $F4/80^+$ (53), but with further research, there are differences in the expression of major histocompatibility complex class II (MHC-II), CC-chemokine receptor 2 (CCR2), lymphatic vessel endothelial receptor 1 (LYVE1), T cell immunoglobulin and mucin domain containing 4 (TIMD4) and lymphocyte antigen 6C (Ly6C) in myocardium. Epelman et al. (54) found three macrophage subsets: The primary cardiac macrophage populations were $Ly6C^-MHC\text{-}II^{high}$, $Ly6C^-MHC\text{-}II^{low}$ and $Ly6C^+$ which represented ~2% of total macrophages in the adult heart. Then researchers investigated the mechanism of myocardial macrophage population recruitment in a mouse model and confirmed that these macrophages were renewed at steady state by *in situ* proliferation and were virtually unaffected by the input of circulating monocytes. Litviňuková et al. (45) used scRNA-seq to analyze the immune cell populations in adult healthy hearts and defined $LYVE1^+$ $TIMD4^-$-resident macrophages, speculating that these cells may be resident macrophages. It has been confirmed that the adult heart contains different resident macrophage subpopulations (55), Bajpai et al. (56) identified macrophages derived from circulating monocytes based on the expression of CCR2. The macrophages were divided into 3 distinct subsets: $CCR2^+MHC\text{-}II^{low}$, $CCR2^+MHC\text{-}II^{high}$, and $CCR2^-MHC\text{-}II^{high}$ cells. Meanwhile, the researchers found that in injured myocardial tissue, $CCR2^+MHC\text{-}II^{low}$ macrophages were only found adjacent to vessels located within areas of dense fibrosis and $CCR2^-$macrophages were mostly distributed in the myocardial area absence of scar tissue. Dick et al. (57) used TIMD4 as a persistent lineage marker for resident cardiac macrophage subpopulations and observed that four macrophage subpopulations in the adult steady-state myocardium, one of which was independent of circulating monocytes ($TIMD4^+LYVE1^+MHC\text{-}II^{low}CCR2^-$), one subpopulation was partially replaced by circulating monocytes ($TIMD4^-LYVE1^-MHC\text{-}II^{high}CCR2^-$), and other two $CCR2^+MHC\text{-}II^{high}$ subpopulations were completely replaced by monocytes. This study illustrates the contribution of recruited monocytes to different subpopulations of resident myocardial macrophages focusing on a hierarchical discussion. Follow-up single-cell sequencing (58) further confirmed that four genes (TIMD4, LYVE1, CCR2, MHC-II) defined the macrophage subpopulations present in the heart and provided new insights into the common starting point for understanding macrophage heterogeneity. In addition to resident macrophages expressing TIMD4 and/or LYVE1, there are also $MHC\text{-}II^{high}$ macrophages and macrophages expressing $CCR2^+$. At steady state, only a modest monocyte contribution is maintained until the monocyte-derived macrophages reach a defined upper limit, which does not exceed the preexisting resident macrophages.

In addition to the typical macrophage functions, including phagocytosis and stimulation of cytokine secretion, cardiac macrophages coordinate the following functions of the heart. Primitive embryonic-derived $CCR2^-$macrophages regulate coronary development through stimulation of coronary angiogenesis and remodeling of the coronary plexus (7). Driven by the cardiac autophagy mechanism, macrophages take up dysfunctional mitochondria to maintain the energy-intense unction of the heart (59). And the network of macrophages is involved in electrical conduction and lymphatic vessel growth. Hulsmans et al. (60) demonstrated that macrophages facilitated electrical conductance through the distal atrioventricular node, and that conducting cells are densely populated with elongated macrophages expressing connexin 43. In contrast, conditional deletion of connexin 43 in macrophages

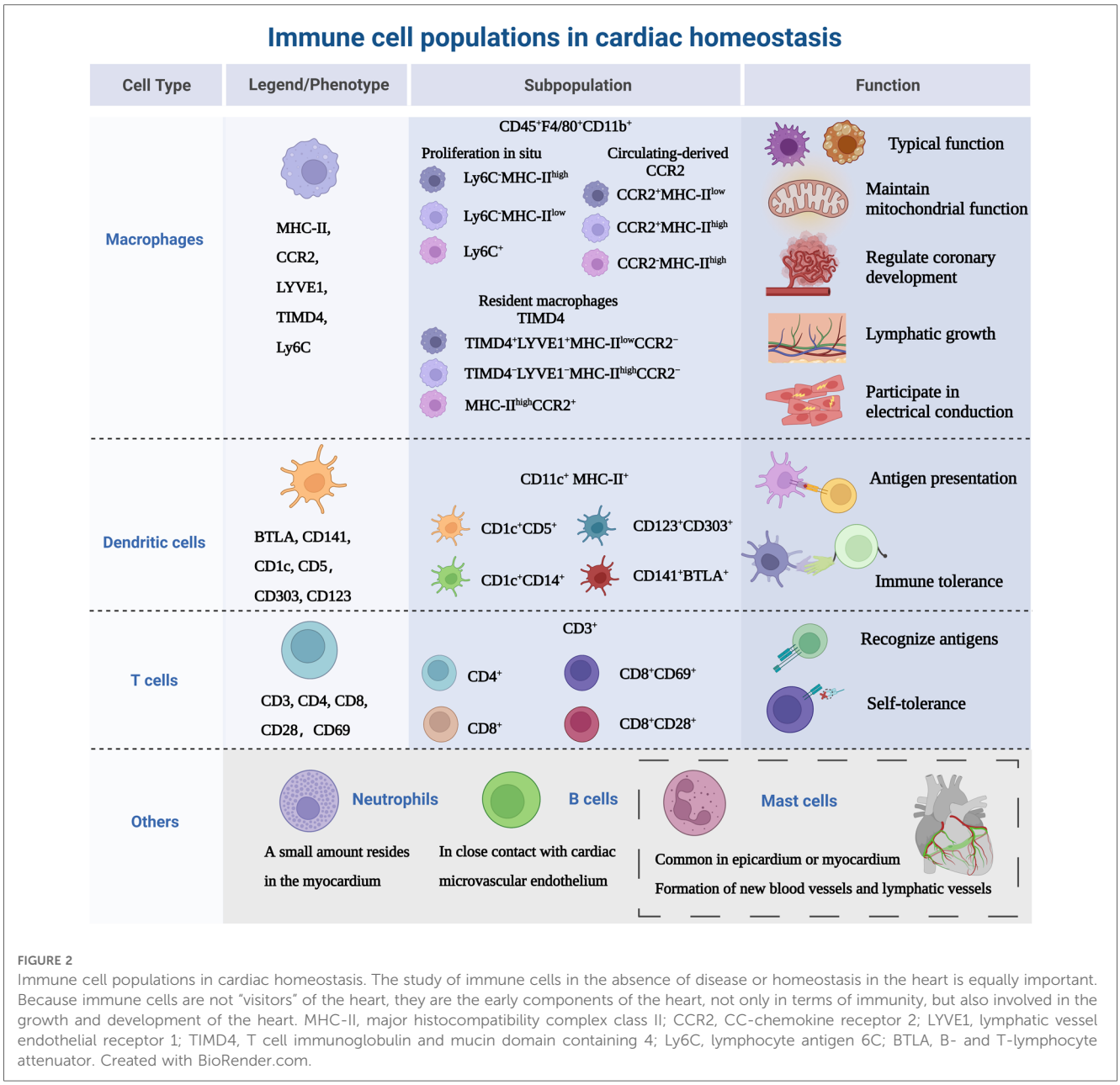


FIGURE 2
Immune cell populations in cardiac homeostasis. The study of immune cells in the absence of disease or homeostasis in the heart is equally important. Because immune cells are not "visitors" of the heart, they are the early components of the heart, not only in terms of immunity, but also involved in the growth and development of the heart. MHC-II, major histocompatibility complex class II; CCR2, CC-chemokine receptor 2; LYVE1, lymphatic vessel endothelial receptor 1; TIMD4, T cell immunoglobulin and mucin domain containing 4; Ly6C, lymphocyte antigen 6C; BTLA, B- and T-lymphocyte attenuator. Created with BioRender.com.

and lack of macrophages delayed atrioventricular conduction. Cahill et al. (61) observed that the distribution of resident macrophages in the subepicardial compartment of the developing mouse heart occurred concurrently the appearance of new lymphatics, and that macrophages interact intimately with the nascent lymphatic capillaries.

3.2. Dendritic cells

DCs distributed in the blood and tissues exposed to the environment, are professional antigen-presenting cells that act as a channel between the innate and adaptive immune system (62). Based on lineage, DCs are classified into three main populations: plasmacytoid (pDC) or conventional/myeloid DCs (cDC1 and cDC2) (63). There is no specific cell surface molecular marker for DCs, CD11c and MHC-II are regarded as the main markers, and they can also be expressed by macrophage subsets (64). Therefore, it is necessary to use different markers simultaneously to identify DCs. The introduction of scRNA-seq technology has greatly expanded the surface markers for identifying DCs, such as B- and T-lymphocyte attenuator (BTLA), CC-chemokine receptor 7(CCR7), CD1c, CD5 (65). DC subsets are sparsely distributed in cardiac tissue, especially abundant in the heart valves and aortic sinus, acting as "sentinels" to detect antigens (6, 66). At homeostasis, DCs share the ability of guiding immunity towards a state of tolerance (67–69). Borght et al. (70) observed in a healthy mouse model that cDC1 presented cardiac self-antigen (α-myosin heavy chain) to $CD4^+$ T cells only in heart-draining mediastinal lymph nodes, resulting in Treg expansion without cardiac injury and accumulation of multiple lymphoid

organs. The α-myosin heavy chain is a pathogenic autoantigen in a mouse model of spontaneous myocarditis, allowing $CD4^+$ T cells to specifically evade thymic negative selection and spread to the periphery (71). In addition, proliferation and recruitment of circulating precursors regulate the turnover of cardiac cDCs (72).

3.3. T cells

T cells are involved in adaptive immunity by recognizing diverse antigen, maintaining immunological memory and self-tolerance (73). The main surface marker of T cells is CD3 and are divided into two subsets: $CD4^+$ T cells can be divided into helper T cells (Th1, Th2 and Th17) and regulatory T cells (T_{reg}). $CD8^+$ T cells are called cytotoxic T cells (74). Skelly et al. (75) analyzed the transcriptional profile of non-cardiomyocytes in normal mice, and found T cells were also highly expressed Lat, Skap1, IL-17R. Unfortunately, the study did not further delineate the subtypes. It has been reported that CD69 is expressed in healthy human hearts, which may be a group of effector resident memory $CD8^+$ T cells (76, 77). CD69 is expressed on activated T cells and affects Th/T_{reg} balance and suppressive activity of T_{reg}, suggesting that cardiac resident T cells are associated with immune tolerance (78). To maintain immune dormancy at homeostasis, in addition to their low numbers in healthy myocardium, cardiac T cells are also regulated by self-antigen-specific tolerance (79). This process is regulated by both central and peripheral immune tolerance. Central immune tolerance refers to the clearance of immature T cells that recognize self-antigens. Self-reactive T cells that partially escape selection are inactivated by peripheral tolerance mechanisms, such as an active repression of T cell receptor signaling and a second signal, resulting in a long-term hyporesponsive state of T cells (80). Programmed cell death protein 1 (PD-1) is a T cell inhibitory molecule homologous to the co-stimulatory receptor CD28, and its role is to block activation signals from the T cell receptor and CD28 (81–83).

3.4. Other immune cell types involved in cardiac homeostasis

There is a gap in the knowledge of the heterogeneity of other immune cells in cardiac homeostasis. We will show the role of immune cells in the context of disease (mentioned in the next chapter). The presence of monocyte subtypes in the myocardium has been less studied with regard to monocyte/macrophage co-expressing cell subpopulations in the heart at steady state has been reported (84, 85). These subpopulations of cells are highly expressed Cd68 and Lyz2. Although extensive infiltration of neutrophils often occurs after cardiac injury (86), Martini et al. (87) used single-cell sequencing to show that small numbers of neutrophils were also present in mouse healthy myocardium. However the exact contribution of neutrophils to the heart is unknown. One study (88) confirmed that in the mouse heart, mast cells were more commonly found in the epicardium or myocardium, and scattered in the endocardium. Within the myocardium, 31% of mast cells were located around blood vessels. This is similar to the distribution of mast cells in the human heart (89). Cardiac mast cells may be involved in the formation of new blood vessels and lymphatic vessels, which are critical for cardiac development, myocardial healing, and maintenance of homeostasis (90, 91). When Adamo et al. (92) tracked myocardial B cell trafficking, they found that myocardial B cells represent a subset of circulating origin that in close contact with the cardiac microvascular endothelium, with less than 5% entered the myocardium. Analysis of B cell-deficient animals showed that cardiac B cells are involved in the growth and contractility of the myocardium (92).

4. The heterogeneous immune landscape in cardiovascular diseases

4.1. Atherosclerosis

Atherosclerosis is the underlying pathology of coronary artery disease and is regarded as a chronic inflammatory disease (93). The retained low-density lipoprotein in the arterial intima is recognized and phagocytosed by scavenger receptors on the surface of macrophages (94). This pathological process results in the accumulation of inflammatory cells and the deposition of lipid particles, which gradually form plaques in macroscopic view. Atherosclerosis is also accompanied by certain low-grade chronic inflammatory responses that attract cells of the innate and adaptive immune systems into atherosclerotic plaques.

4.1.1. Inflammatory responses in atherosclerosis

He et al. (95) confirmed that the aortic arch in the aorta is prone to vascular injury. And compared with healthy mice, the immune cells of high-fat mice were more heterogeneous and pro-inflammatory cells were found: Th17 cells, $CD8^+$ T cells and chemokine-enriched macrophages. The team's follow-up research found (85) that ascending aortic walls of mice fed a high-fat diet were invaded by a subset of macrophages that also amplified local inflammatory responses via scRNA-seq analysis. Four major monocyte/macrophage subpopulations were defined, including two resident clusters (cluster 1 and cluster 2), a blood-derived cluster and a B-cell-like cluster. In healthy mice, resident cluster 1 accounted for a larger proportion (>80%) in the subpopulation and significantly expressed CXC class chemokine genes. Resident cluster 2 contained proliferative properties, not only strongly expressed cell chemotactic genes and also expressed nuclear division-related genes (Top2a and Stmn1). In the presence of high cholesterol intake, the proportion of resident cluster 2 was upregulated, and cells continued to proliferate and express chemokines, increasing the likelihood that this subpopulation was programmed to exacerbate inflammation. The remaining two subpopulation had lower proportions, which might be involved in the regulation of immune effector processes. Atherosclerotic lesions are dominated by macrophages, monocytes and T cells.

The central inflammatory cells in plaques are macrophages. The classic theory of atherosclerosis holds that the classically-

activated macrophages (M1)/alternatively-activated macrophages (M2) balance in plaques is dynamic, with M1 predominant in progression and M2 predominant in regression (96). Interferons induce polarization of the M1 state, and at the other end of the polarization spectrum, M2 macrophages can be induced by IL-4. Lin et al. (97) analyzed plaque cells derived from fractalkine receptor (CX3CR1)-positive precursors in mice with atherosclerosis progression or regression, revealing the transformation and activation characteristics of monocytes/macrophages. During progression of atherosclerosis, there is a more prominent heterogeneity in the activation status of macrophages, including macrophages induced by interferon and IL-4, respectively. However, there was a significant IL-4-activated macrophage population in progressive plaques compared to regressive plaques, indicating that macrophage subpopulations were not simply dichotomous. As for the specific subpopulations in regression, only a small fraction of the total cells, including a subpopulation with high expression of B cell-related genes and some subpopulations with anti-inflammatory functions. They also found a cluster of monocytes with stem-like characteristics, suggesting that circulating monocytes did not differentiate immediately after entering the myocardium, but might continue to proliferate and self-renew at sites of myocardial inflammation. These circulating mononuclear cells may play a unique role in self-repair after myocardial injury. Although macrophages in myocardial tissue are partly derived from resident macrophages, the majority of plaque macrophages are most likely derived from circulating monocytes recruited during disease progression. Cochain et al. (98) clarified the presence of $CX3CR1^{+}$ monocyte-derived macrophage subpopulations in both the progression and regression of atherosclerosis, which represented a general inflammatory feature of atherosclerosis. These subpopulations of cells were almost exclusively detected in atherosclerotic aorta. The team described a subset of macrophages, enriched for triggered receptor expressed on myeloid cells-2 (TREM2). Such cells appear to have specialized functions in lipid metabolism and catabolism, and display gene expression signatures similar to osteoclasts, suggesting a role in pathological calcification. In addition, a class of monocyte-derived DCs with functions similar to inflammatory macrophages was defined, showing strong expression of IL-1β but low expression of chemokines. Li et al. (99) designed a program (AtheroSpectrum) to process data results from single-cell sequencing sources, revealing specific gene expression profiles associated with inflammatory macrophage foam cells, contributing to the development of therapeutic and prognostic strategies. For example, their study supported that the $TREM2^{+}$ macrophage subpopulation was consistent with a low inflammatory profile.

Importantly, the concepts of cardiac immune heterogeneity are equally corroborated in the human heart. Fernandez et al. (100) uncovered distinct features of leukocytes in carotid artery plaques of patients. Immune cells specifically enriched in plaques included macrophages and T cells. The defined immune cells largely correspond to known immune cell populations. Plaque macrophages comprised two macrophage subsets ($CD64^{+}HLADR^{+}CD206^{high}$ and $CD64^{+}HLADR^{+}CD206^{low}$) based on the varied expression of the M2 marker CD206. $CD8^{+}$ T cells and $CD4^{+}$ T cells were enriched in the plaques. Plaque T cells display transcriptional signatures associated with cellular activation, cytotoxicity, and cellular exhaustion. $CD4^{+}$ T cell sub-analysis showed that unlike circulating cells at rest, in plaques they displayed activated Th1 pro-inflammatory function and chemotaxis. $CD8^{+}$ T cells in plaques not only activated the pro-inflammatory interferon-γ pathway but also activated the programmed cell death-1 signaling pathway which was a marker of T cell exhaustion (101). They found a cluster of plaque-specific T cells expressing T cell activation gene and the only cluster associated with activation of pro-inflammatory Th1 and Th17 signaling pathways. The overlapping phenomena of T cell subsets activation and depletion may be due to the progressive loss of T cell function in chronic, persistent inflammatory responses induced by unresolved plaques. T cell expression in human atherosclerotic lesions perhaps is more active. Depuydt et al. (102) applied scRNA-seq to all live cells in advanced human atherosclerotic plaques and found T cell subsets were better characterized by an activated state. $CD4^{+}$ T cell subsets were broadly divided into cytotoxicity-signature cell cluster, regulatory T-cell cluster and central-memory cell cluster. The $CD4^{+}$ T with cytotoxic features was shown to be characterized by the expression of PRF1 and granzyme B. In addition, a large number of T cells did not express well-defined transcription factor signals, which can be missed when analyzing human data using the classic T-helper subsets of transcription factors.$CD8^{+}$ cell subsets were classified into cytotoxicity-signature cell cluster, effector-memory cell cluster and central-memory cell cluster as similar with $CD4^{+}$ T cells. However, unlike the data from Fernandez et al., the researchers did not detect a clear depleted phenotype in $CD8^{+}$ T cells.

4.1.2. Chemokines and immune cells in atherosclerosis

Trajectory analysis showed that not only macrophages, but also T cells differentiated into proliferative populations when the differentiation process was activated, and the other branch was Th17 cells (103). Furthermore, macrophages influenced T cell activation by modulating MHC-II and possibly fibroblasts through the transforming growth factor β pathway. In the cellular communication analysis, T cells, fibroblasts and macrophages had significant interactions, and the most active chemotactic communication among the three was CXC chemokine ligand 12 (CXCL12)-CXC chemokine receptor 4 (CXCR4) and CC chemokine ligand 5 (CCL5)-CCR5 (103). Chemokines are a class of small molecule cytokines that induce directed chemotaxis by activating G protein-coupled receptors. Chemokine subsets are determined according to the arrangement of amino acid (N-terminal) cysteines. Most chemokines are associated with cardiovascular disease and have pro-inflammatory and leukocyte recruitment effects (104, 105). Winkels et al. (106) constructed atlas of the immune cells in mouse atherosclerosis. They found enhanced expression of inflammatory genes CXCL2, CCL2, CCL3, Ly6C2, suggesting $Ly6C^{+}$ inflammatory monocytes. T cells clusters mainly included

Th2, Th17, $CD8^+$ cytotoxic T cells, memory T cells and the mix of $CD4^+/CD8^+$ T cells. Three principal B-cell subsets were detected by the markers CD43 and B220 ($CD43^{high}B220^{neg}$, $CD43^{neg}B220^{high}$, $CD43^{low}B220^{high}$). The $CD43^{high}B220^{neg}$ B cells tended to be pro-atherogenic chemokine CCL5, while $CD43^{neg}B220^{high}$ showed a predominant of the pro-inflammatory cytokine interferon γ. Further studying immune cells at different lesion sites, they found that in advanced stages of atherosclerosis, leukocyte signaling was relatively high in extravascular tissue rich in tertiary lymphoid organs, with a predominant distribution of B cells. This lesion site may be one of the major immune cell repositories in advanced atherosclerosis. CCR2, CCR5 and CX3CR1 are all involved in the recruitment of monocytes in the pathological process of atherosclerosis (107). Burger et al. (108) attempted to define macrophage subpopulations involved in atherosclerosis progression. They found that $LYVE1^+$ macrophages, which express high levels of CCL24, expand under hypercholesterolemia in $Apoe^{-/-}$ mice and promote the conversion of vascular smooth muscle cells to osteoblasts/chondrocytes in a CCL24-dependent manner. This suggests that $LYVE1^+$ resident macrophages may play an important role in vascular calcification and atherosclerotic plaque instability. The process by which monocytes, guided by chemokines, enter the plaque, engulf oxidized cholesterol, convert it into foam cells, and then exit the plaque is called atherosclerosis regression. One of the features of regression is the overall reduction of plaque macrophages. Rahman et al. (109) compared plaque regression in the aortic transplantation model with normo-lipidemic recipients and those deficient in chemokine receptors, explaining the role of the above chemokines in regression. They found that recruitment of CCR2-dependent and CX3CR1-dependent monocytes was critical for the regression of plaque macrophage content. Inadequate CCR2 or CX3CR1 in aortic transplant models prevented plaque regression. Furthermore, although CCR5 promoted the progression of atherosclerosis by recruiting monocytes, it was not required for the resolution of atherosclerosis. In addition, Park et al. (110) found that C-type lectin receptor CLEC4A2 induced monocytes to join the resident macrophage pool via scRNA-seq, which had the atheroprotective property. CLEC4A is the only classical C-type lectin receptor that possesses an intracellular immunoreceptor tyrosine-based inhibitory motif, which possibly transduces negative signals (111). Through gene deletion and competitive bone marrow chimerism experiments, they determined that CLEC4A2 played a protective role in atherosclerosis by maintaining myeloid steady-state. CLEC4A2 deficiency caused the loss of homeostatic properties of resident macrophages during atherogenesis, resulting in dysfunctional cholesterol metabolism, enhanced toll-like receptor triggering and aggravating the disease.

4.2. Myocardial infarction

MI usually occurs when a new thrombus or ruptured plaque blocks a coronary artery. MI is a common cause of acute aseptic heart injury (112). Because the heart requires sufficient energy and oxygen to maintain continuous contraction, disruption of coronary blood supply can irreversibly kill myocardial tissue in the affected area, especially when the heart is under high load. After MI, the distribution of immune cells in myocardium has changed significantly. The plasticity of subpopulation phenotypes enhances the need to analyze the characteristics of the cellular environment by single-cell techniques. Acutely dead cardiomyocytes recruit immune cells that express different phenotypes and trigger different alarms (113).

4.2.1. Intracardiac myeloid cells subpopulations

Neutrophils are short-term first responders to MI. In response to ischemic injury, circulating neutrophils and monocytes infiltrate the heart in a form that approximates a stacked wave. Calcagno et al. (114) explained the possibility of remote activation of myeloid cells from the peripheral circulation to infiltrating the infarcted heart via scRNA-seq. Similar monocyte subpopulations with high expression of interferon stimulated genes (ISGs) were found in the peripheral blood of both humans and mice suffering from MI. They also focused on myeloid cells in the infarcted heart and defined five intracardiac neutrophil subpopulations which had distinct temporal patterns between day 1 and day 4 post-MI. The first is a subpopulation of parental neutrophil that highly express Retnlg throughout the event. Another subset continued to express ISGs for 1–4 days after MI, and such expression was previously only found in monocyte-derived macrophages. The other three subpopulations highly expressed nuclear factor kappa B activation-related genes, hypoxia inducible factor 1 activation-related genes, and Siglecf, respectively. As for these three subpopulations, the first two subpopulations decreased proportionally over time and gradually were replaced by the third group. Macrophages in infarcted hearts also expressed ISGs. The team found that multiple myeloid cells at the infarct site expressed ISG and that ISG was detectable in early circulating neutrophils, which was enhanced as neutrophils matured. They also identified aberrant type I interferon innate immune activation in monocytes and neutrophils in the bone marrow of mice at steady-state and after MI. ISG became an important component in driving immune cell aggregation, suggesting that innate immunity was involved in the long-range activation associated with interferons. Macrophages have been shown to trigger a dual cascade of inflammation and repair in cardiac remodeling after MI. Farbehi et al. (115) and Jin et al. (116) successively confirmed that macrophage subpopulations have specific dominant subpopulations and dynamic changes at different stages after MI. $MHC\text{-}II^{high}$ subsets indicate a monocytic origin and robust inflammatory potential. The researchers observed that $MHC\text{-}II^{high}$ $CCR2^{high}$ subsets had expanded by day 3 after MI, whereas the $MHC\text{-}II^{high}$ $CCR2^{negative}$ subset which shared many inflammatory features (CCL5, CXCL9, Gpb2, etc) was more noticeable at a later stage of MI (116). The subsets enriched for features related to leukocyte migration and chemotaxis was significantly increased in the acute phase (day 3) and were typical post-infarct inflammatory clusters. The next four subsets showed low levels of CCR2 and MHC-II expression with marked

expansion from the restoration stage (day 7). These subsets showed the ability to resist cardiomyocyte injury, enrichment of chemokines and receptor-mediated endocytosis, and the ability to promote tissue fibrosis (angiogenesis and collagen-related gene expression), respectively (115). The profibrotic cluster manifestly appeared at day 14 after MI and involved in improving cardiac regeneration (116). Diffusion map analyses were used to model possible temporal changes in the main macrophages/monocytes population, revealing a cascade of states that break down from early infiltration and cell migration, inflammatory over-characterization, to late peak inflammatory resolution and repair.

4.2.2. Lymphocyte response to myocardial infarct

In addition to the acute myeloid response, MI triggers a lymphocyte response that affects recovery. Heinrichs et al. (117) conducted in mice a phenotyping of B-cell responses in infarcted hearts to dissect the mechanisms controlling B-cell mobilization and activity *in situ*. MI triggered synchronized B-cell responses in the infarct site and the mediastinal lymph nodes. A unique B cell cluster was also identified in the infarcted heart, showing obvious chemokine receptor characteristics, with high expression of CCR7 and CXCR5. These cells were enriched in the infarct site and peaked at Day 7 after MI. Quantitative analysis found that CXCR5 ligand (CXCL13) were expressed at higher levels than CCR7 ligands (CCL19 and CCL21), and then decreased myocardial B cell infiltration after MI was observed in CXCR5-deficient mice. It is likely that healing after MI is dependent on the CXCL13- CXCR5 axis, but not CCR7, to mobilize B cells to infiltrate the myocardium.

4.3. Heart failure

Heart failure (HF) is primarily a clinical diagnosis that occurs secondary to either left ventricular systolic and diastolic dysfunction. Ultimately the heart is not able to transport a sufficient amount of blood under normal filling pressures to meet our body's needs. Cardiac diseases that cause HF include, but are not limited to, MI, hypertension, heart valve disease, and myocarditis (118). Myocardial injury is bound to cause aseptic myocardial inflammation; whereas termination of inflammatory infiltration requires activation of anti-inflammatory stimulus pathways, which in turn initiate profibrotic signaling (119, 120).

4.3.1. Cardiac immune composition in heart failure

Under stress, immune cells even express a fibrotic phenotype (121). The relative balance between pathological inflammatory pathways and tissue repair processes determines the developmental trajectory of HF (122). Martini et al. (87) mapped the cardiac immune composition in the murine nonischemic, pressure-overload heart failure model. They demonstrated a major reorganization of immune cell abundance at late stages of HF (4 weeks after transverse aortic constriction), with macrophages and T cells being the most abundant. Activation of macrophage-dominated immune subpopulations in response to cardiac stress induced expansion, and these subpopulations clustered together macroscopically as cardiac inflammation. HF induced an increase in the abundance of resident $CCR2^{-}$ antigen-presenting macrophages, $CCR2^{-}$ phagocytic monocytes/macrophages, and proinflammatory, $CCR2^{+}Osm^{+}IL\text{-}1b^{+}$ recruited cells to varying degrees. Oncostatin M (OSM), a member of the IL-6 family, chronic activation of which promotes the development of heart failure, functions as a major mediator of cardiomyocyte remodeling under pathological conditions (123). B cells expressed observably higher levels of CXCR5 and CCR7 after the intervention. These chemokines targeted B cells homing and directed them to the T-cell compartment of the lymph node, which increased the likelihood that B and T cells were organized into tertiary lymph nodes, as they were in atherosclerotic aortic adventitia. PD-1 was found to be present in T_{reg} cells in the early stages of the disease and promoted T_{reg}-induced immunosuppressive function. T_{reg} cells were actively expressed and stable in number in the early stages of HF, and only in the later stages did they show a significant increase in number. In addition, they observed two neutrophil subpopulations, a relatively more numerous subpopulation of pro-inflammatory cells (CCR2) and a smaller subpopulation of pro-reparative cells (CCR1 and CXCR2), whose differing chemokine receptor expression might mean different microenvironment positioning modes. Koenig et al. (124) observed that HF had the greatest effect on the differential gene expression of myeloid cells and attempted to reveal their multiple subgroups. Patients with dilated cardiomyopathy displayed reduced numbers of resident macrophages and a greater number of intermediate monocytes, and other additional monocyte-derived macrophage subsets. The additional monocyte-derived macrophage subsets exhibited high chemotaxis and high inflammation. Anti-monocyte recruitment may reduce cardiac pathological inflammation and myocardial fibrosis.

4.3.2. Immune cells and myocardial fibrosis in heart failure

Revelo et al. (10) analyzed the role of cardiac immune cells, especially macrophages, in the decompensated hypertrophic response of the myocardium. Macrophages played an important role in the early stages of the hypertrophic remodeling process, as evidenced by the positive correlation between the number of macrophages and the level of cardiac hypertrophy. Resident macrophages were found to be depleted early, but recruited macrophages were rapidly replenished from circulating monocytes. The macrophage-derived factors transforming growth factor-β1 and IL-10 were increased in the heart after transverse aortic constriction, and these factors had been implicated in the progression of cardiac fibrosis. The researchers found that preferential depletion of resident macrophages reduced angiogenesis and worsened fibrosis, in addition to the fact that recruited macrophages were promoters of fibrosis and stimulated the aforementioned derived factors. Resident macrophages prevent deterioration of cardiac function and ameliorate fibrosis progression after pressure overload during the early stages of cardiac remodeling. Rao's team (125) solved the gene expression disturbance between leukocytes and myocardial fibrosis. The

results of the study showed that in the fibrotic myocardium, a large number of leukocytes infiltrated, especially $CD8^+$ T cells and $CD4^+$ T cells. In addition, tissue-resident $CXCL8^{high}CCR2^+HLADR^{high}$ macrophages were found especially in areas of severe fibrosis. $CXCL8^{high}CCR2^+HLADR^{high}$ cells highly express CXCL8, while its receptor duffy antigen/receptor for chemokines is expressed in activated endothelial cells. It interacted with activated endothelial cells via duffy antigen/receptor for chemokines and might promote leukocyte recruitment and infiltration in human heart failure. Several clinical studies have measured circulating subsets of monocytes or lymphocytes to assess the degree of recovery of left ventricular function or prognostic risk (126–128).

4.3.3. Circulating immune cells and heart failure

Abplanalp et al. (129) analyzed the impact of circulating immune cells on HF and gave new insights into the characteristics of immune cells at the single-cell level. Although the number of monocyte subpopulations was not affected by HF, the gene expression had significantly altered. The researchers classified monocyte subpopulations according to the expression of well-established markers CD14/CD16, which were roughly divided into classical monocytes (CD14/CD16 ratio >2.5), intermediate monocytes (CD14/CD16 ratio of 0.45–2.5), and non-classical monocytes (CD14/CD16 ratio <2.5). They found specific changes in gene expression patterns in the classical monocyte subpopulations, associated with cell differentiation, regulation of cell migration and stress response. Fatty acid binding protein 5 upregulation was enriched in the classical monocytes, which contributed to the marked induction of inflammatory signatures in patients with heart failure. Intermediate monocytes were more prone to post-stress adaptive signaling, and expressions were associated with epithelial-mesenchymal transition and cell proliferation. β-catenin was deeply enriched in intermediate monocytes involved in the regulation of multiple functions, the most important of which is the "epithelial-to-mesenchymal transition", indicating that it played an important role in HF progression and fibrosis. This suggested that modulating monocyte heterogeneity could be a new therapeutic target for rescue of HF.

5. Challenges and prospects

Single-cell technology uses unsupervised cluster analysis to define up to dozens of cell types, thereby assessing the importance of each cell cluster and completing the identification of key immune cell subsets. Opportunities to improve CVDs may arise following consensus on cell type and subtype identity characteristics and their functional implications. Theoretically, the selective suppression of detrimental immune cell subsets, or conversion to subsets with more beneficial functions, would lead to more effective prevention and treatment of malignant clinical events. Future increases in adaptive regulation of immune cell subpopulations, and even individualized interventions, are expected to have a profound impact on increasing life expectancy, improving quality of life, and clarifying diagnoses. Single-cell technology is now being used in clinical research. Fan et al. applied single-cell techniques to analyze the composition and phenotype of CD45+ cells in human peripheral blood samples. They revealed peripheral immune signatures associated with disease stages in patients with coronary artery disease and atherosclerotic cardiovascular disease and identified specific peripheral immune cell subpopulations that are strongly associated with coronary artery disease severity. These revealed stage-specific peripheral immune signatures become promising minimally invasive liquid biomarkers to potentially diagnose and monitor the progression of cardiovascular disease in humans (130). However, the development of such subset-specific inhibitors or reprogramming agents still needs to overcome numerous hurdles. The practical problem is that the study on immune heterogeneity cannot be separated from samples from healthy controls, and requires collecting as many immune cells as possible in the location and purifying the samples as much as possible to ensure that results are not biased towards specific cell types. Once a rare subset of immune cells has been identified using single-cell genomics, researchers need to further purify and specifically study them. The complexity of immune ecology in CVDs needs to be taken into account, and the inherent understanding of the phenotype of immune cells based on single-cell technology will be reconsidered, perhaps with the classification of gene expression profiles as a worthwhile attempt (131, 132).

As sequencing costs continue to decrease, it is economically promising to integrate multiple layers of bioinformatics and technical replication to improve the ability to detect changes in immunoheterogeneous expression. In addition, recent developments in spatial transcriptomics and multimodal omics will contribute to subpopulation-specific expression patterns and more comprehensive gene expression networks from a spatiotemporal perspective (133, 134).

6. Conclusion

Single-cell technology is now emerging as a powerful tool for comprehensive and unbiased analysis of the cellular composition of normal/diseased tissues. Experimental single-cell studies have revealed new heterogeneity in immune regions, demonstrating more clearly defined immune differences, compared to the relatively conserved normal situation. Single-cell techniques have also demonstrated significant immune cell cellular adaptability and plasticity in the pathological setting. In-depth single-cell studies on a limited number of precious human specimens have improved understanding of specific immune cell alterations, such as macrophages. In addition to revealing significant differences in gene expression, several variables potentially affecting immune cell behavior were analyzed in combination with in-depth analysis, considering genetic background, spatial location or intercellular interactions, showing changes in the composition of the driving cell population. The scRNA-seq analysis has remarkable resolution, even when dealing with small numbers of cells. It nicely demonstrates the enormous complexity of decision regulation in immune cells, which is a good example.

TABLE 1 Renewed phenotype marker genes of immune cells via single-cell RNA sequencing.

Disease	Model	Source of Sample	Sample Size	Methodology	Marker genes	Ref.
Under high fat condition	C57BL/6J mice	Aortic tissue	24,001	t-SNE, CellChat	Mo/MΦ cluster 1: C1q, PF4	(85)
					Mo/MΦ cluster 2: Top2a, Stmn1, Coro1a	
					Mo/MΦ cluster 3: COL1A1, Sparc, Gsn, Dcn	
					Mo/MΦ cluster 4: CD79a, CD20, Ly6d	
Under high fat/salt/glucose condition	C57BL/6J mice	Aortic tissue	216,612	t-SNE, CellPhone DB	Th17 cells: CD4, CCR6, IL-17	(95)
					$CD8^+$ T cells: CD8, IFNG,	
					Macrophages: CXCL10, CXCL2, CCL3	
					B cells: CD24, IL-10	
Atherosclerosis	$Ldlr^{-/-}$ mice	Aortic tissue	5,355	t-SNE	stem-like macrophages: CX3CR1, CD14, Adgre1, Csf1, CD68	(97)
under high fat condition	$Ldlr^{-/-}$ mice	Aortic tissue	1,226	t-SNE, SCDE	Macrophages: TREM2	(98)
					Monocyte-derived DC/DCs: Ifi30, Napsa, CD209a	
Atherosclerosis	Human	Aortic tissue, PBMCs	9,490	CyTOF, CITE-seq, aggregated scRNA-seq	T cells: NFATC2, FYN, ZAP70, GzmA, GzmK, EOMES, PDCD1, CD223	(100)
Atherosclerosis	Human	Aortic tissue	3,282	CCA, t-SNE	$CD4^+$ T cells:	(102)
					Cytotoxicity-signature cluster: GzmA, GzmK, PRF1	
					Regulatory T-cell cluster: FOXP3, CD25, CTLA4	
					Central-memory cell cluster: IL-7R, LEF1, SELL	
					$CD8^+$ T cells:	
					Cytotoxicity-signature cluster: GzmB, TBX21, NKG7	
					Effector-memory cell cluster: GzmK, GzmA, CD69	
					Central-memory cell cluster: LEF1, SELL, IL-7R	
Atherosclerosis	$Apoe^{-/-}$, $Ldlr^{-/-}$ mice	Aortic tissue	555	PCA, t-SNE, CyTOF	T cells: $CD43^{high}B220^{neg}$, $CD43^{neg}B220^{high}$, $CD43^{low}B220^{high}$	(106)
MI	Human, C57BL/6J mice	PBMCs, heart tissue	3,102/20,206	UMAP	Neutrophils cluster 1: Retnlg	(114)
					Neutrophils cluster 2: Isg15, Ifit3, Rsad2, Ifit1, Irf7	
					Neutrophils cluster 3: Nfkb1, Icam1, IL-1a, Sod2, Tnip1	
					Neutrophils cluster 4: Egln3, Hilpda, Vegfa	
					Neutrophils cluster 5: Siglecf	
MI	C57BL/6J mice	Heart tissue	17,932	t-SNE, UMAP	Mo/MΦ cluster 1: CXCL9, Igkc, Ighm	(116)
					Mo/MΦ cluster 2: CD81, C1qa, Fcrls	
					Mo/MΦ cluster 3: Ifitm1, CD74, Napsa	
					Mo/MΦ cluster 4: CCL5, CXCL9, Gpb2	
					Mo/MΦ cluster 5: S100a9, Rsad2, Retnlg	
					Mo/MΦ cluster 6: Fabp5, Hmox1, Gpnmb	
					Mo/MΦ cluster 7: CXCL3, Sod2, Prdx1	
					Mo/MΦ cluster 8: Gas6, Cbr2, Selenop	
					Mo/MΦ cluster 9: Fn1, Slc7a2, Sdc3	
					Mo/MΦ cluster 10: Dcn, Col3a1, Bgn	
					Mo/MΦ cluster 11: Stmn1, Birc5, Top2a	
MI	C57BL/6J mice	Heart tissue	6,588	UMAP	B cells: CCR7, CXCR5, CD69	(117)
HF	C57BL/6J mice	Heart tissue	18,115	CCA, t-SNE	Mo/MΦ cluster 1: CD1, CD163, Mrc1	(87)
					Mo/MΦ cluster 2: CCR2, IL-1b, Chil3	
					Mo/MΦ cluster 3: Ear2, Eno3, CD36	
					Mo/MΦ cluster 4: CCR2, Lyz2, Ifitm6	
					B cell cluster 1: CCR7, Pax5, Ly6d	
					B cell cluster 2: CCR7, CXCR5, Mef2c	
					B cell cluster 3: CXCR5, CCR7, CD19	
					B cell cluster 4: CCR7, CD19, CX3CR1	
					$CD8^+$ cell cluster 1: CD8, CD3, Thy1	
					$CD8^+$ cell cluster 2: CD8, IL-17, Lat	
					$CD4^+$ cell cluster 1: CD3, Lat, Lck	
					$CD4^+$ cell cluster 2: Ly6c1, CD28, Tcf7	

(*continued*)

TABLE 1 Continued

Disease	Model	Source of Sample	Sample Size	Methodology	Marker genes	Ref.
					T_{reg} cell cluster: Foxp3, Tnfrsf18, CTL4	
					Nonexpanding T cell-like cluster: Thy1, Rora, Arg1	
					Neutrophil cluster 1: CCR2, CD69	
					Neutrophil cluster 2: CCR1, CXCR2, Arg2	
HF	Human	Heart tissue	49,723	UMAP	Macrophage cluster 1: TREM2, SPP1, LGALS3	(124)
					Macrophage cluster 2: FOLR2, LYVE1	
					Macrophage cluster 3: LYVE1, HSPH1, HSPA1A	
					Macrophage cluster 4: CCL3, CCL4, PHLDA1	
					Macrophage cluster 5: KLF2, KLF4, EGR1	
					Monocyte cluster 1: LST1, LILRA5, CD16	
					Monocyte cluster 2: FCN1, CD14, S100a8	
					Monocyte cluster 3: FCN1, OLR1, PLAUR1	

scRNA-Seq, single-cell RNA sequencing; PCA, principle component analysis; t-SNE, t-distributed stochastic neighbor embedding; Th17, T helper 17; IL, interleukin; CCR, CC-chemokine receptor; CXCL, CXC chemokine ligand; CCL, CC chemokine ligand; IFNG, interferon-gamma; Mo/MΦ, monocytes/macrophages; PF4, platelet factor 4; Top2a, topoisomerase 2alpha; STMN1, stathmin 1; Coro1a, coronin-1A; COL1A1 collagen 1A1; Sparc, secreted protein acidic and rich in cysteine; Gsn, gelsolin; Dcn, decorin; Ly6d, lymphocyte antigen 6D; SCDE, single-cell differential expression analysis;TREM2, triggered receptor expressed on myeloid cells 2; PBMCs, peripheral blood mononuclear cells; Ifi30, gamma-interferon-inducible lysosomal thiol reductase; Napsa, napsin A aspartic peptidase; Adgre1, adhesion G protein-coupled receptor E1; Csf1, colony stimulating factor 1; NFATC2, nuclear factor of activated T cells 2; FYN, Src family tyrosine kinase; ZAP70, 70-kDa zeta-associated protein; Gzm, granzyme; EOMES, eomesodermin; CCA, Canonical Correlation Analysis; PDCD1, programmed death-1; PRF1, perforin 1; FOXP3, forkhead box protein P3; CTLA432; cytotoxic T-lymphocyte-associated protein 4; LEF1, lymphoid enhancer-binding factor 1; SELL, selectin L; UMAP, uniform manifold approximation and projection; TBX21, T box 21; NKG7, natural killer cell group 7; Igkc, immunoglobulin kappa constant; Ighm, immunoglobulin M heavy chain; Fcrls, Fc receptor-like proteins; Ifitm, interferon induced transmembrane protein; Gpb2, G-protein beta 2; S100a9, S100 calcium binding protein A9; Rsad2, radical S-adenosyl methionine domain containing 2; Retnlg, resistin like gamma; Fabp5, fatty acid binding protein 5; Hmox1, heme oxygenase 1; Gpnmb, glycoprotein NMB; Sod2, superoxide dismutase 2; Prdx1, peroxiredoxin 1; Gas6, growth arrest-specific 6; Cbr2, carbonyl reductase 2; Selenop, selenoprotein P; Fn1, fibronectin1; Slc7a2, solute carrier family 7 member 2; Sdc3, syndecan 3; Col3a1, collagen type III alpha 1 chain-1; Bgn, biglycan; MI, myocardial ischemia; CLCF1, cardiotrophin-like cytokine factor 1; snRNA-Seq, single-nucleus RNA sequencing; Birc5, baculoviral IAP repeat containing 5; Areg, amphiregulin; Isg15, interferon-stimulated gene 15; Ifit3, Interferon-induced protein with tetratricopeptide repeats 3; Irf7, interferon regulatory factor 7; Nfkb1, nuclear factor of kappa light polypeptide gene enhancer in B-cells 1; Icam1, intercellular cell adhesion molecule-1; Tnip1, TNFAIP3-interacting protein 1; Egln3, egl nine homolog 3; Hilpda, hypoxia inducible lipid droplet-associated; Vegfa, vascular endothelial growth factor A; Siglecf, sialic acid-recognizing lectin F; Mrc1, C1 mannose receptor; Chil3, chitinase-like protein 3; Ear2, nuclear receptor subfamily 2 group F member 6; Eno3, enolase 3; Lyz2, lysozyme 2; Pax5, paired-box 5; Mef2c, myocyte enhancer factor 2C; Thy1, Thy-1 cell surface antigen; Lat, linker for activation of T cells; Lck, lymphocyte-specific cytoplasmic protein-tyrosine kinase p56lck; Tcf7, transcription factor 7; Tnfrsf18, tumor necrosis factor receptor superfamily member 18; Rora, RAR-related orphan receptor A; Arg1, Arginase 1; SPP1, secreted phosphoprotein 1; LGALS3, lectin, galactoside-binding, soluble 3; FOLR2, folate receptor 2; LYVE1, lymphatic vessel endothelial hyaluronan receptor 1; HSPH1, Heat shock protein 105 kDa; HSPA1A, heat shock 70 kDa; PHLDA1, pleckstrin homology-like domain, family A, member 1; KLF2, kruppel-like factor 2; EGR1, early growth response gene 1; LST1, Leukocyte Specific Transcript 1; LILRA5, leukocyte immunoglobulin-like receptor, subfamily A, member 5; FCN1, ficolin 1; OLR1, oxidized low density lipoprotein (lectin-like) receptor 1; PLAUR1, plasminogen activator, urokinase receptor 1; Apoe, apolipoprotein E; Ldlr, low-density lipoprotein receptor.

In recent years, the extensive role of immune cells in cardiovascular pathophysiology has begun to receive attention. These range from heart development and repair, to the maintenance of myocardial homeostasis, to the participation of key mechanisms of CVDs. Macrophages and T cells have the most significant heterogeneity in CVDs, their numbers and proportions are more variable than other immune cells, and they interact closely with other non-cardiomyocytes. The influence of macrophages runs through all stages of atherosclerosis, from lesion initiation and enlargement, to necrosis, and then to regression. Circulation-derived macrophages are enriched in inflammation and chemotaxis, and exhibit richer phenotypic features compared to the classical polarization classification, which may be related to their environment and activated signaling pathways. Likewise, the potent role of resident macrophages cannot be ignored, and such cells exhibit proliferative properties in the pathological process of atherosclerosis. Acute MI triggers a large number of immune cells to infiltrate the injury site, and neutrophils and monocytes overlap and accumulate at the injury site in the early response. Depletion of monocytes/macrophages affects myocardial regeneration, but the source remains to be determined (resident vs. migrating monocytes/macrophages). In addition to myeloid cells, distinct B cell subsets are defined, are strongly chemotactic and have pro-reparative effects. Immune cell heterogeneity is closely related to the time points of HF; for example, early resident macrophages have been shown to prevent cardiac deterioration, but macrophages replenished from the circulation are at risk of aggravating pathological fibrosis. Immune cell subsets are involved in signaling mechanisms regulating myocardial fibrosis, which requires further exploration of causal relationships in the future. In addition, immune cell subsets may also affect lymphatic function and remodeling, and cardiac lymph angiogenesis is also one of the therapeutic targets for HF. The appearance and gene expressions of immune cell populations at the site of injury and in circulation correlate with disease progression. As shown in **Table 1**, this review summarizes the phenotypic marker genes of immune cells defined through scRNA-seq. At the same time, the study of high-parameter immunophenotyping has just started, and there are many important issues to be addressed in future work. The key is to understand which immune cell subpopulations exhibit

plasticity/instability and which are pathogenic clusters, and whether these can be targeted for drug therapy. Myocardial immune cell heterogeneity promotes disease progression under pathological conditions. Using healthy controls to study signature genes, combined with advances in clinical trial design, a bold approach to drug development will improve patient outcomes.

Funding

This work was supported by the National Natural Science Foundation of China (grant number 82174349), the CACMS Innovation Fund (grant number CI2021A00919), and the National Key R&D Program of China (grant numbers 2018YFC1704901 and 2018YFC1704900).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

1. Knuuti J, Wijns W, Saraste A, Capodanno D, Barbato E, Funck-Brentano C, et al. 2019 esc guidelines for the diagnosis and management of chronic coronary syndromes. *Eur Heart J*. (2020) 41(3):407–77. doi: 10.1093/eurheartj/ehz425

2. Flora GD, Nayak MK. A brief review of cardiovascular diseases, associated risk factors and current treatment regimes. *Curr Pharm Des*. (2019) 25(38):4063–84. doi: 10.2174/1381612825666190925163827

3. Mensah GA, Roth GA, Fuster V. The global burden of cardiovascular diseases and risk factors: 2020 and beyond. *J Am Coll Cardiol*. (2019) 74(20):2529–32. doi: 10.1016/j.jacc.2019.10.009

4. Roberts LB, Kapoor P, Howard JK, Shah AM, Lord GM. An update on the roles of immune system-derived micrornas in cardiovascular diseases. *Cardiovasc Res*. (2021) 117(12):2434–49. doi: 10.1093/cvr/cvab007

5. Strassheim D, Dempsey EC, Gerasimovskaya E, Stenmark K, Karoor V. Role of inflammatory cell subtypes in heart failure. *J Immunol Res*. (2019) 2019:2164017. doi: 10.1155/2019/2164017

6. Choi JH, Do Y, Cheong C, Koh H, Boscardin SB, Oh YS, et al. Identification of antigen-presenting dendritic cells in mouse aorta and cardiac valves. *J Exp Med*. (2009) 206(3):497–505. doi: 10.1084/jem.20082129

7. Leid J, Carrelha J, Boukarabila H, Epelman S, Jacobsen SE, Lavine KJ. Primitive embryonic macrophages are required for coronary development and maturation. *Circ Res*. (2016) 118(10):1498–511. doi: 10.1161/circresaha.115.308270

8. Swirski FK, Nahrendorf M. Cardioimmunology: the immune system in cardiac homeostasis and disease. *Nat Rev Immunol*. (2018) 18(12):733–44. doi: 10.1038/s41577-018-0065-8

9. Hulsmans M, Sager HB, Roh JD, Valero-Muñoz M, Houstis NE, Iwamoto Y, et al. Cardiac macrophages promote diastolic dysfunction. *J Exp Med*. (2018) 215 (2):423–40. doi: 10.1084/jem.20171274

10. Revelo XS, Parthiban P, Chen C, Barrow F, Fredrickson G, Wang H, et al. Cardiac resident macrophages prevent fibrosis and stimulate angiogenesis. *Circ Res*. (2021) 129(12):1086–101. doi: 10.1161/circresaha.121.319737

11. Nidorf SM, Fiolet ATL, Mosterd A, Eikelboom JW, Schut A, Opstal TSJ, et al. Colchicine in patients with chronic coronary disease. *N Engl J Med*. (2020) 383 (19):1838–47. doi: 10.1056/NEJMoa2021372

12. Hill CA, Fernandez DM, Giannarelli C. Single cell analyses to understand the immune continuum in atherosclerosis. *Atherosclerosis*. (2021) 330:85–94. doi: 10.1016/j.atherosclerosis.2021.04.003

13. Satija R, Shalek AK. Heterogeneity in immune responses: from populations to single cells. *Trends Immunol*. (2014) 35(5):219–29. doi: 10.1016/j.it.2014.03.004

14. Arendt D, Musser JM, Baker CVH, Bergman A, Cepko C, Erwin DH, et al. The origin and evolution of cell types. *Nat Rev Genet*. (2016) 17(12):744–57. doi: 10.1038/nrg.2016.127

15. Raj A, van Oudenaarden A. Nature, nurture, or chance: stochastic gene expression and its consequences. *Cell*. (2008) 135(2):216–26. doi: 10.1016/j.cell.2008.09.050

16. Ko MS, Nakauchi H, Takahashi N. The dose dependence of glucocorticoid-inducible gene expression results from changes in the number of transcriptionally active templates. *EMBO J*. (1990) 9(9):2835–42. doi: 10.1002/j.1460-2075.1990.tb07472.x

17. Jovic D, Liang X, Zeng H, Lin L, Xu F, Luo Y. Single-cell RNA sequencing technologies and applications: a brief overview. *Clin Transl Med*. (2022) 12(3):e694. doi: 10.1002/ctm2.694

18. Sinha N, Subedi N, Tel J. Integrating immunology and microfluidics for single immune cell analysis. *Front Immunol*. (2018) 9:2373. doi: 10.3389/fimmu.2018.02373

19. Liu T, Yamaguchi Y, Shirasaki Y, Shikada K, Yamagishi M, Hoshino K, et al. Single-cell imaging of caspase-1 dynamics reveals an all-or-none inflammasome signaling response. *Cell Rep*. (2014) 8(4):974–82. doi: 10.1016/j.celrep.2014.07.012

20. Newell EW, Sigal N, Nair N, Kidd BA, Greenberg HB, Davis MM. Combinatorial tetramer staining and mass cytometry analysis facilitate T-cell epitope mapping and characterization. *Nat Biotechnol*. (2013) 31(7):623–9. doi: 10.1038/nbt.2593

21. Shalek AK, Satija R, Adiconis X, Gertner RS, Gaublomme JT, Raychowdhury R, et al. Single-cell transcriptomics reveals bimodality in expression and splicing in immune cells. *Nature*. (2013) 498(7453):236–40. doi: 10.1038/nature12172

22. Oetjen KA, Lindblad KE, Goswami M, Gui G, Dagur PK, Lai C, et al. Human bone marrow assessment by single-cell RNA sequencing, mass cytometry, and flow cytometry. *JCI Insight*. (2018) 3(23):e124928. doi: 10.1172/jci.insight.124928

23. Wu H, Fu R, Zhang YH, Liu Z, Chen ZH, Xu J, et al. Single-cell RNA sequencing unravels upregulation of immune cell crosstalk in relapsed pediatric ependymoma. *Front Immunol*. (2022) 13:903246. doi: 10.3389/fimmu.2022.903246

24. Villani AC, Satija R, Reynolds G, Sarkizova S, Shekhar K, Fletcher J, et al. Single-cell RNA-seq reveals new types of human blood dendritic cells, monocytes, and progenitors. *Science*. (2017) 356(6335):eaah4573. doi: 10.1126/science.aah4573

25. Korenfeld D, Gorvel L, Munk A, Man J, Schaffer A, Tung T, et al. A type of human skin dendritic cell marked by Cd5 is associated with the development of inflammatory skin disease. *JCI Insight*. (2017) 2(18):e96101. doi: 10.1172/jci.insight.96101

26. Tang F, Barbacioru C, Wang Y, Nordman E, Lee C, Xu N, et al. MRNA-seq whole-transcriptome analysis of a single cell. *Nat Methods*. (2009) 6(5):377–82. doi: 10.1038/nmeth.1315

27. Ziegenhain C, Vieth B, Parekh S, Reinius B, Guillaumet-Adkins A, Smets M, et al. Comparative analysis of single-cell RNA sequencing methods. *Mol Cell*. (2017) 65(4):631–43.e4. doi: 10.1016/j.molcel.2017.01.023

28. Zhang X, Li T, Liu F, Chen Y, Yao J, Li Z, et al. Comparative analysis of droplet-based ultra-high-throughput single-cell RNA-seq systems. *Mol Cell*. (2019) 73 (1):130–42.e5. doi: 10.1016/j.molcel.2018.10.020

29. Picelli S. Single-cell RNA-sequencing: the future of genome biology is now. *RNA Biol*. (2017) 14(5):637–50. doi: 10.1080/15476286.2016.1201618

30. Chambers DC, Carew AM, Lukowski SW, Powell JE. Transcriptomics and single-cell RNA-sequencing. *Respirology*. (2019) 24(1):29–36. doi: 10.1111/resp.13412

31. Stoeckius M, Hafemeister C, Stephenson W, Houck-Loomis B, Chattopadhyay PK, Swerdlow H, et al. Simultaneous epitope and transcriptome measurement in single cells. *Nat Methods*. (2017) 14(9):865–8. doi: 10.1038/nmeth.4380

32. Saelens W, Cannoodt R, Todorov H, Saeys Y. A comparison of single-cell trajectory inference methods. *Nat Biotechnol*. (2019) 37(5):547–54. doi: 10.1038/s41587-019-0071-9

33. Jin S, Guerrero-Juarez CF, Zhang L, Chang I, Ramos R, Kuan CH, et al. Inference and analysis of cell-cell communication using cellchat. *Nat Commun*. (2021) 12(1):1088. doi: 10.1038/s41467-021-21246-9

34. Skinnider MA, Squair JW, Kathe C, Anderson MA, Gautier M, Matson KJE, et al. Cell type prioritization in single-cell data. *Nat Biotechnol*. (2021) 39(1):30–4. doi: 10.1038/s41587-020-0605-1

35. Moncada R, Barkley D, Wagner F, Chiodin M, Devlin JC, Baron M, et al. Integrating microarray-based spatial transcriptomics and single-cell RNA-seq reveals tissue architecture in pancreatic ductal adenocarcinomas. *Nat Biotechnol*. (2020) 38 (3):333–42. doi: 10.1038/s41587-019-0392-8

36. Provost G, Lavoie FB, Larbi A, Ng TP, Ying CTT, Chua M, et al. Novel approach to analysis of the immune system using an ungated model of immune surface marker abundance to predict health outcomes. *Immun Ageing*. (2022) 19(1):35. doi: 10.1186/s12979-022-00291-y

37. Zhang L, Yu X, Zheng L, Zhang Y, Li Y, Fang Q, et al. Lineage tracking reveals dynamic relationships of T cells in colorectal cancer. *Nature*. (2018) 564 (7735):268–72. doi: 10.1038/s41586-018-0694-x

38. Ni SH, Xu JD, Sun SN, Li Y, Zhou Z, Li H, et al. Single-cell transcriptomic analyses of cardiac immune cells reveal that rel-driven Cd72-positive macrophages induce cardiomyocyte injury. *Cardiovasc Res*. (2022) 118(5):1303–20. doi: 10.1093/cvr/cvab193

39. Matsumoto H, Kiryu H, Furusawa C, Ko MSH, Ko SBH, Gouda N, et al. Scode: an efficient regulatory network inference algorithm from single-cell RNA-seq during differentiation. *Bioinformatics*. (2017) 33(15):2314–21. doi: 10.1093/bioinformatics/btx194

40. Zhuang L, Wang Y, Chen Z, Li Z, Wang Z, Jia K, et al. Global characteristics and dynamics of single immune cells after myocardial infarction. *J Am Heart Assoc*. (2022) 11(24):e027228. doi: 10.1161/jaha.122.027228

41. Clark SJ, Argelaguet R, Kapourani CA, Stubbs TM, Lee HJ, Alda-Catalinas C, et al. Scnmt-seq enables joint profiling of chromatin accessibility DNA methylation and transcription in single cells. *Nat Commun*. (2018) 9(1):781. doi: 10.1038/s41467-018-03149-4

42. Song Q, Zhu X, Jin L, Chen M, Zhang W, Su J. Smgr: a joint statistical method for integrative analysis of single-cell multi-omics data. *NAR Genom Bioinform*. (2022) 4(3):lqac056. doi: 10.1093/nargab/lqac056

43. Li S, Wu B, Ling Y, Guo M, Qin B, Ren X, et al. Epigenetic landscapes of single-cell chromatin accessibility and transcriptomic immune profiles of T cells in COVID-19 patients. *Front Immunol*. (2021) 12:625881. doi: 10.3389/fimmu.2021.625881

44. Chaffin M, Papangeli I, Simonson B, Akkad AD, Hill MC, Arduini A, et al. Single-nucleus profiling of human dilated and hypertrophic cardiomyopathy. *Nature*. (2022) 608(7921):174–80. doi: 10.1038/s41586-022-04817-8

45. Litviňuková M, Talavera-López C, Maatz H, Reichart D, Worth CL, Lindberg EL, et al. Cells of the adult human heart. *Nature*. (2020) 588(7838):466–72. doi: 10.1038/s41586-020-2797-4

46. Sakai M, Troutman TD, Seidman JS, Ouyang Z, Spann NJ, Abe Y, et al. Liver-derived signals sequentially reprogram myeloid enhancers to initiate and maintain kupffer cell identity. *Immunity*. (2019) 51(4):655–70.e8. doi: 10.1016/j.immuni.2019.09.002

47. Ballesteros I, Rubio-Ponce A, Genua M, Lusito E, Kwok I, Fernández-Calvo G, et al. Co-option of neutrophil fates by tissue environments. *Cell*. (2020) 183 (5):1282–97.e18. doi: 10.1016/j.cell.2020.10.003

48. Wong NR, Mohan J, Kopecky BJ, Guo S, Du L, Leid J, et al. Resident cardiac macrophages mediate adaptive myocardial remodeling. *Immunity*. (2021) 54 (9):2072–88.e7. doi: 10.1016/j.immuni.2021.07.003

49. Zhang H, Xu A, Sun X, Yang Y, Zhang L, Bai H, et al. Self-maintenance of cardiac resident reparative macrophages attenuates doxorubicin-induced cardiomyopathy through the Sr-A1-C-myc axis. *Circ Res*. (2020) 127(5):610–27. doi: 10.1161/circresaha.119.316428

50. Sansonetti M, Waleczek FJG, Jung M, Thum T, Perbellini F. Resident cardiac macrophages: crucial modulators of cardiac (patho)physiology. *Basic Res Cardiol*. (2020) 115(6):77. doi: 10.1007/s00395-020-00836-6

51. Simões FC, Riley PR. Immune cells in cardiac repair and regeneration. *Development*. (2022) 149(8):dev199906. doi: 10.1242/dev.199906

52. Carrillo-Salinas FJ, Ngwenyama N, Anastasiou M, Kaur K, Alcaide P. Heart inflammation: immune cell roles and roads to the heart. *Am J Pathol*. (2019) 189 (8):1482–94. doi: 10.1016/j.ajpath.2019.04.009

53. Zaynagetdinov R, Sherrill TP, Kendall PL, Segal BH, Weller KP, Tighe RM, et al. Identification of myeloid cell subsets in murine lungs using flow cytometry. *Am J Respir Cell Mol Biol*. (2013) 49(2):180–9. doi: 10.1165/rcmb.2012-0366MA

54. Epelman S, Lavine KJ, Beaudin AE, Sojka DK, Carrero JA, Calderon B, et al. Embryonic and adult-derived resident cardiac macrophages are maintained through distinct mechanisms at steady state and during inflammation. *Immunity*. (2014) 40 (1):91–104. doi: 10.1016/j.immuni.2013.11.019

55. Lavine KJ, Epelman S, Uchida K, Weber KJ, Nichols CG, Schilling JD, et al. Distinct macrophage lineages contribute to disparate patterns of cardiac recovery and remodeling in the neonatal and adult heart. *Proc Natl Acad Sci U S A*. (2014) 111(45):16029–34. doi: 10.1073/pnas.1406508111

56. Bajpai G, Schneider C, Wong N, Bredemeyer A, Hulsmans M, Nahrendorf M, et al. The human heart contains distinct macrophage subsets with divergent origins and functions. *Nat Med*. (2018) 24(8):1234–45. doi: 10.1038/s41591-018-0059-x

57. Dick SA, Macklin JA, Nejat S, Momen A, Clemente-Casares X, Althagafi MG, et al. Self-renewing resident cardiac macrophages limit adverse remodeling following myocardial infarction. *Nat Immunol*. (2019) 20(1):29–39. doi: 10.1038/s41590-018-0272-2

58. Dick SA, Wong A, Hamidzada H, Nejat S, Nechanitzky R, Vohra S, et al. Three tissue resident macrophage subsets coexist across organs with conserved origins and life cycles. *Sci Immunol*. (2022) 7(67):eabf7777. doi: 10.1126/sciimmunol.abf7777

59. Nicolás-Ávila JA, Lechuga-Vieco AV, Esteban-Martínez L, Sánchez-Díaz M, Díaz-García E, Santiago DJ, et al. A network of macrophages supports mitochondrial homeostasis in the heart. *Cell*. (2020) 183(1):94–109.e23. doi: 10.1016/j.cell.2020.08.031

60. Hulsmans M, Clauss S, Xiao L, Aguirre AD, King KR, Hanley A, et al. Macrophages facilitate electrical conduction in the heart. *Cell*. (2017) 169 (3):510–22.e20. doi: 10.1016/j.cell.2017.03.050

61. Cahill TJ, Sun X, Ravaud C, Villa Del Campo C, Klaourakis K, Lupu IE, et al. Tissue-resident macrophages regulate lymphatic vessel growth and patterning in the developing heart. *Development*. (2021) 148(3):dev194563. doi: 10.1242/dev.194563

62. Joffre OP, Segura E, Savina A, Amigorena S. Cross-presentation by dendritic cells. *Nat Rev Immunol*. (2012) 12(8):557–69. doi: 10.1038/nri3254

63. Collin M, Bigley V. Human dendritic cell subsets: an update. *Immunology*. (2018) 154(1):3–20. doi: 10.1111/imm.12888

64. Shi H, Ge J, Fang W, Yao K, Sun A, Huang R, et al. Peripheral-blood dendritic cells in men with coronary heart disease. *Am J Cardiol*. (2007) 100(4):593–7. doi: 10.1016/j.amjcard.2007.03.067

65. Villar J, Segura E. Decoding the heterogeneity of human dendritic cell subsets. *Trends Immunol*. (2020) 41(12):1062–71. doi: 10.1016/j.it.2020.10.002

66. Forte E, Perkins B, Sintou A, Kalkat HS, Papanikolaou A, Jenkins C, et al. Cross-priming dendritic cells exacerbate immunopathology after ischemic tissue damage in the heart. *Circulation*. (2021) 143(8):821–36. doi: 10.1161/circulationaha.120.044581

67. Bigley V, Cytlak U, Collin M. Human dendritic cell immunodeficiencies. *Semin Cell Dev Biol*. (2019) 86:50–61. doi: 10.1016/j.semcdb.2018.02.020

68. Théry C, Amigorena S. The cell biology of antigen presentation in dendritic cells. *Curr Opin Immunol*. (2001) 13(1):45–51. doi: 10.1016/s0952-7915(00)00180-1

69. Waisman A, Lukas D, Clausen BE, Yogev N. Dendritic cells as gatekeepers of tolerance. *Semin Immunopathol*. (2017) 39(2):153–63. doi: 10.1007/s00281-016-0583-z

70. Van der Borght K, Scott CL, Nindl V, Bouché A, Martens L, Sichien D, et al. Myocardial infarction primes autoreactive T cells through activation of dendritic cells. *Cell Rep*. (2017) 18(12):3005–17. doi: 10.1016/j.celrep.2017.02.079

71. Lv H, Havari E, Pinto S, Gottumukkala RV, Cornivelli L, Raddassi K, et al. Impaired thymic tolerance to A-myosin directs autoimmunity to the heart in mice and humans. *J Clin Invest*. (2011) 121(4):1561–73. doi: 10.1172/jci44583

72. Clemente-Casares X, Hosseinzadeh S, Barbu I, Dick SA, Macklin JA, Wang Y, et al. A Cd103(+) conventional dendritic cell surveillance system prevents development of overt heart failure during subclinical viral myocarditis. *Immunity*. (2017) 47(5):974–89.e8. doi: 10.1016/j.immuni.2017.10.011

73. Kumar BV, Connors TJ, Farber DL. Human T cell development, localization, and function throughout life. *Immunity*. (2018) 48(2):202–13. doi: 10.1016/j.immuni.2018.01.007

74. Walsh AJ, Mueller KP, Tweed K, Jones I, Walsh CM, Piscopo NJ, et al. Classification of T-cell activation via autofluorescence lifetime imaging. *Nat Biomed Eng*. (2021) 5(1):77–88. doi: 10.1038/s41551-020-0592-z

75. Skelly DA, Squiers GT, McLellan MA, Bolisetty MT, Robson P, Rosenthal NA, et al. Single-cell transcriptional profiling reveals cellular diversity and intercommunication in the mouse heart. *Cell Rep*. (2018) 22(3):600–10. doi: 10.1016/j.celrep.2017.12.072

76. Tucker NR, Chaffin M, Fleming SJ, Hall AW, Parsons VA, Bedi KC Jr, et al. Transcriptional and cellular diversity of the human heart. *Circulation*. (2020) 142 (5):466–82. doi: 10.1161/circulationaha.119.045401

77. Casey KA, Fraser KA, Schenkel JM, Moran A, Abt MC, Beura LK, et al. Antigen-independent differentiation and maintenance of effector-like resident memory T cells in tissues. *J Immunol*. (2012) 188(10):4866–75. doi: 10.4049/jimmunol.1200402

78. Gorabi AM, Hajighasemi S, Kiaie N, Gheibi Hayat SM, Jamialahmadi T, Johnston TP, et al. The pivotal role of Cd69 in autoimmunity. *J Autoimmun*. (2020) 111:102453. doi: 10.1016/j.jaut.2020.102453

79. Grabie N, Lichtman AH, Padera R. T cell checkpoint regulators in the heart. *Cardiovasc Res*. (2019) 115(5):869–77. doi: 10.1093/cvr/cvz025

80. Xing Y, Hogquist KA. T-cell tolerance: central and peripheral. *Cold Spring Harb Perspect Biol*. (2012) 4(6):a006957. doi: 10.1101/cshperspect.a006957

81. Quigley M, Pereyra F, Nilsson B, Porichis F, Fonseca C, Eichbaum Q, et al. Transcriptional analysis of hiv-specific Cd8+ T cells shows that pd-1 inhibits T cell function by upregulating batf. *Nat Med*. (2010) 16(10):1147–51. doi: 10.1038/nm.2232

82. Hui E, Cheung J, Zhu J, Su X, Taylor MJ, Wallweber HA, et al. T cell costimulatory receptor Cd28 is a primary target for pd-1-mediated inhibition. *Science*. (2017) 355(6332):1428–33. doi: 10.1126/science.aaf1292

83. Li S, Tajiri K, Murakoshi N, Xu D, Yonebayashi S, Okabe Y, et al. Programmed death-ligand 2 deficiency exacerbates experimental autoimmune myocarditis in mice. *Int J Mol Sci*. (2021) 22(3):1426. doi: 10.3390/ijms22031426

84. Shang M, Hu Y, Cao H, Lin Q, Yi N, Zhang J, et al. Concordant and heterogeneity of single-cell transcriptome in cardiac development of human and mouse. *Front Genet*. (2022) 13:892766. doi: 10.3389/fgene.2022.892766

85. Kan H, Zhang K, Mao A, Geng L, Gao M, Feng L, et al. Single-cell transcriptome analysis reveals cellular heterogeneity in the ascending aortas of normal and high-fat diet-fed mice. *Exp Mol Med*. (2021) 53(9):1379–89. doi: 10.1038/s12276-021-00671-2

86. Calcagno DM, Zhang C, Toomu A, Huang K, Ninh VK, Miyamoto S, et al. Siglecf(Hi) marks late-stage neutrophils of the infarcted heart: a single-cell transcriptomic analysis of neutrophil diversification. *J Am Heart Assoc*. (2021) 10 (4):e019019. doi: 10.1161/jaha.120.019019

87. Martini E, Kunderfranco P, Peano C, Carullo P, Cremonesi M, Schorn T, et al. Single-cell sequencing of mouse heart immune infiltrate in pressure overload-driven heart failure reveals extent of immune activation. *Circulation*. (2019) 140 (25):2089–107. doi: 10.1161/circulationaha.119.041694

88. Ingason AB, Mechmet F, Atacho DAM, Steingrímsson E, Petersen PH. Distribution of mast cells within the mouse heart and its dependency on mitf. *Mol Immunol*. (2019) 105:9–15. doi: 10.1016/j.molimm.2018.11.009

89. Varricchi G, Marone G, Kovanen PT. Cardiac mast cells: underappreciated immune cells in cardiovascular homeostasis and disease. *Trends Immunol*. (2020) 41(8):734–46. doi: 10.1016/j.it.2020.06.006

90. Syväranta S, Helske S, Lappalainen J, Kupari M, Kovanen PT. Lymphangiogenesis in aortic valve stenosis–novel regulatory roles for valvular myofibroblasts and mast cells. *Atherosclerosis*. (2012) 221(2):366–74. doi: 10.1016/j.atherosclerosis.2011.12.034

91. Sun K, Li YY, Jin J. A double-edged sword of immuno-microenvironment in cardiac homeostasis and injury repair. *Signal Transduct Target Ther*. (2021) 6(1):79. doi: 10.1038/s41392-020-00455-6

92. Adamo L, Rocha-Resende C, Lin CY, Evans S, Williams J, Dun H, et al. Myocardial B cells are a subset of circulating lymphocytes with delayed transit through the heart. *JCI Insight*. (2020) 5(3):e134700. doi: 10.1172/jci.insight.134700

93. Chen X, He Y, Fu W, Sahebkar A, Tan Y, Xu S, et al. Histone deacetylases (hdacs) and atherosclerosis: a mechanistic and pharmacological review. *Front Cell Dev Biol*. (2020) 8:581015. doi: 10.3389/fcell.2020.581015

94. Moore KJ, Freeman MW. Scavenger receptors in atherosclerosis: beyond lipid uptake. *Arterioscler Thromb Vasc Biol*. (2006) 26(8):1702–11. doi: 10.1161/01.Atv.0000229218.97976.43

95. He D, Mao A, Zheng CB, Kan H, Zhang K, Zhang Z, et al. Aortic heterogeneity across segments and under high fat/salt/glucose conditions at the single-cell level. *Natl Sci Rev*. (2020) 7(5):881–96. doi: 10.1093/nsr/nwaa038

96. Peled M, Fisher EA. Dynamic aspects of macrophage polarization during atherosclerosis progression and regression. *Front Immunol*. (2014) 5:579. doi: 10.3389/fimmu.2014.00579

97. Lin JD, Nishi H, Poles J, Niu X, McCauley C, Rahman K, et al. Single-cell analysis of fate-mapped macrophages reveals heterogeneity, including stem-like properties, during atherosclerosis progression and regression. *JCI Insight*. (2019) 4 (4):e124574. doi: 10.1172/jci.insight.124574

98. Cochain C, Vafadarnejad E, Arampatzi P, Pelisek J, Winkels H, Ley K, et al. Single-cell RNA-seq reveals the transcriptional landscape and heterogeneity of aortic macrophages in murine atherosclerosis. *Circ Res*. (2018) 122(12):1661–74. doi: 10.1161/circresaha.117.312509

99. Li C, Qu L, Matz AJ, Murphy PA, Liu Y, Manichaikul AW, et al. Atherospectrum reveals novel macrophage foam cell gene signatures associated with atherosclerotic cardiovascular disease risk. *Circulation*. (2022) 145(3):206–18. doi: 10.1161/circulationaha.121.054285

100. Fernandez DM, Rahman AH, Fernandez NF, Chudnovskiy A, Amir ED, Amadori L, et al. Single-cell immune landscape of human atherosclerotic plaques. *Nat Med*. (2019) 25(10):1576–88. doi: 10.1038/s41591-019-0590-4

101. Kamphorst AO, Wieland A, Nasti T, Yang S, Zhang R, Barber DL, et al. Rescue of exhausted Cd8 T cells by pd-1-targeted therapies is Cd28-dependent. *Science*. (2017) 355(6332):1423–7. doi: 10.1126/science.aaf0683

102. Depuydt MAC, Prange KHM, Slenders L, Örd T, Elbersen D, Boltjes A, et al. Microanatomy of the human atherosclerotic plaque by single-cell transcriptomics. *Circ Res*. (2020) 127(11):1437–55. doi: 10.1161/circresaha.120.316770

103. Zhang C, Sun Y, Yang C, Wang Q, Lu Y. Single-cell sequencing of mouse heart cellular heterogeneity in hypercholesterolemia reveals the mechanism of myocardial damage. *Clin Transl Med*. (2022) 12(7):e951. doi: 10.1002/ctm2.951

104. Lu X, Wang Z, Ye D, Feng Y, Liu M, Xu Y, et al. The role of cxc chemokines in cardiovascular diseases. *Front Pharmacol*. (2021) 12:765768. doi: 10.3389/fphar.2021.765768

105. Hartmann P, Schober A, Weber C. Chemokines and microrna in atherosclerosis. *Cell Mol Life Sci*. (2015) 72(17):3253–66. doi: 10.1007/s00018-015-1925-z

106. Winkels H, Ehinger E, Vassallo M, Buscher K, Dinh HQ, Kobiyama K, et al. Atlas of the immune cell repertoire in mouse atherosclerosis defined by single-cell RNA-sequencing and mass cytometry. *Circ Res*. (2018) 122(12):1675–88. doi: 10.1161/circresaha.117.312513

107. Kang H, Li X, Xiong K, Song Z, Tian J, Wen Y, et al. The entry and egress of monocytes in atherosclerosis: a biochemical and biomechanical driven process. *Cardiovasc Ther*. (2021) 2021:6642927. doi: 10.1155/2021/6642927

108. Burger F, Baptista D, Roth A, Brandt KJ, da Silva RF, Montecucco F, et al. Single-cell RNA-seq reveals a crosstalk between hyaluronan receptor lyve-1-expressing macrophages and vascular smooth muscle cells. *Cells*. (2022) 11(3):411. doi: 10.3390/cells11030411

109. Rahman K, Vengrenyuk Y, Ramsey SA, Vila NR, Girgis NM, Liu J, et al. Inflammatory Ly6chi monocytes and their conversion to M2 macrophages drive atherosclerosis regression. *J Clin Invest*. (2017) 127(8):2904–15. doi: 10.1172/jci75005

110. Park I, Goddard ME, Cole JE, Zanin N, Lyytikäinen LP, Lehtimäki T, et al. C-type lectin receptor Clec4a2 promotes tissue adaptation of macrophages and protects against atherosclerosis. *Nat Commun*. (2022) 13(1):215. doi: 10.1038/s41467-021-27862-9

111. Nasu J, Uto T, Fukaya T, Takagi H, Fukui T, Miyanaga N, et al. Pivotal role of the carbohydrate recognition domain in self-interaction of Clec4a to elicit the itim-mediated inhibitory function in murine conventional dendritic cells in vitro. *Int Immunol*. (2020) 32(10):673–82. doi: 10.1093/intimm/dxaa034

112. Partida RA, Libby P, Crea F, Jang IK. Plaque erosion: a new in vivo diagnosis and a potential major shift in the management of patients with acute coronary syndromes. *Eur Heart J*. (2018) 39(22):2070–6. doi: 10.1093/eurheartj/ehx786

113. Swirski FK, Nahrendorf M. Leukocyte behavior in atherosclerosis, myocardial infarction, and heart failure. *Science*. (2013) 339(6116):161–6. doi: 10.1126/science.1230719

114. Calcagno DM, Ng RP Jr, Toomu A, Zhang C, Huang K, Aguirre AD, et al. The myeloid type I interferon response to myocardial infarction begins in bone marrow and is regulated by Nrf2-activated macrophages. *Sci Immunol*. (2020) 5(51): eaaz1974. doi: 10.1126/sciimmunol.aaz1974

115. Farbehi N, Patrick R, Dorison A, Xaymardan M, Janbandhu V, Wystub-Lis K, et al. Single-cell expression profiling reveals dynamic flux of cardiac stromal, vascular and immune cells in health and injury. *eLife*. (2019) 8:e43882. doi: 10.7554/eLife.43882

116. Jin K, Gao S, Yang P, Guo R, Li D, Zhang Y, et al. Single-cell RNA sequencing reveals the temporal diversity and dynamics of cardiac immunity after myocardial infarction. *Small Methods*. (2022) 6(3):e2100752. doi: 10.1002/smtd.202100752

117. Heinrichs M, Ashour D, Siegel J, Büchner L, Wedekind G, Heinze KG, et al. The healing myocardium mobilizes a distinct B-cell subset through a Cxcl13-Cxcr5-dependent mechanism. *Cardiovasc Res*. (2021) 117(13):2664–76. doi: 10.1093/cvr/cvab181

118. Bacmeister L, Schwarzl M, Warnke S, Stoffers B, Blankenberg S, Westermann D, et al. Inflammation and fibrosis in murine models of heart failure. *Basic Res Cardiol*. (2019) 114(3):19. doi: 10.1007/s00395-019-0722-5

119. Frangogiannis NG. The inflammatory response in myocardial injury, repair, and remodelling. *Nat Rev Cardiol*. (2014) 11(5):255–65. doi: 10.1038/nrcardio.2014.28

120. Clancy RM, Markham AJ, Jackson T, Rasmussen SE, Blumenberg M, Buyon JP. Cardiac fibroblast transcriptome analyses support a role for interferogenic, profibrotic, and inflammatory genes in anti-ssa/ro-associated congenital heart block. *Am J Physiol Heart Circ Physiol*. (2017) 313(3):H631–h40. doi: 10.1152/ajpheart.00256.2017

121. Frangogiannis NG. Cardiac fibrosis. *Cardiovasc Res*. (2021) 117(6):1450–88. doi: 10.1093/cvr/cvaa324

122. Dick SA, Epelman S. Chronic heart failure and inflammation: what do we really know? *Circ Res*. (2016) 119(1):159–76. doi: 10.1161/circresaha.116.308030

123. Kubin T, Gajawada P, Bramlage P, Hein S, Berge B, Cetinkaya A, et al. The role of oncostatin M and its receptor complexes in cardiomyocyte protection, regeneration, and failure. *Int J Mol Sci*. (2022) 23(3):1811. doi: 10.3390/ijms23031811

124. Koenig AL, Shchukina I, Amrute J, Andhey PS, Zaitsev K, Lai L, et al. Single-cell transcriptomics reveals cell-type-specific diversification in human heart failure. *Nat Cardiovasc Res*. (2022) 1(3):263–80. doi: 10.1038/s44161-022-00028-6

125. Rao M, Wang X, Guo G, Wang L, Chen S, Yin P, et al. Resolving the intertwining of inflammation and fibrosis in human heart failure at single-cell level. *Basic Res Cardiol*. (2021) 116(1):55. doi: 10.1007/s00395-021-00897-1

126. Silva N, Bettencourt P, Guimarães JT. The lymphocyte-to-monocyte ratio: an added value for death prediction in heart failure. *Nutr Metab Cardiovasc Dis*. (2015) 25(11):1033–40. doi: 10.1016/j.numecd.2015.07.004

127. Tsujioka H, Imanishi T, Ikejima H, Kuroi A, Takarada S, Tanimoto T, et al. Impact of heterogeneity of human peripheral blood monocyte subsets on

myocardial salvage in patients with primary acute myocardial infarction. *J Am Coll Cardiol*. (2009) 54(2):130–8. doi: 10.1016/j.jacc.2009.04.021

128. Charach G, Rogowski O, Karniel E, Charach L, Grosskopf I, Novikov I. Monocytes may be favorable biomarker and predictor of long-term outcome in patients with chronic heart failure: a cohort study. *Medicine (Baltimore)*. (2019) 98 (38):e17108. doi: 10.1097/md.0000000000017108

129. Abplanalp WT, John D, Cremer S, Assmus B, Dorsheimer L, Hoffmann J, et al. Single-cell RNA-sequencing reveals profound changes in circulating immune cells in patients with heart failure. *Cardiovasc Res*. (2021) 117(2):484–94. doi: 10.1093/cvr/cvaa101

130. Fan L, Liu J, Zhang Y, Zhang C, Shi B, Hu X, et al. High-dimensional single-cell analysis delineates peripheral immune signature of coronary atherosclerosis in human blood. *Theranostics*. (2022) 12(15):6809–25. doi: 10.7150/thno.73336

131. Giladi A, Amit I. Immunology, one cell at a time. *Nature*. (2017):547 (7661):27–9. doi: 10.1038/547027a

132. Saunders A, Macosko EZ, Wysoker A, Goldman M, Krienen FM, de Rivera H, et al. Molecular diversity and specializations among the cells of the adult mouse brain. *Cell*. (2018) 174(4):1015–30.e16. doi: 10.1016/j.cell.2018.07.028

133. Lin S, Liu Y, Zhang M, Xu X, Chen Y, Zhang H, et al. Microfluidic single-cell transcriptomics: moving towards multimodal and spatiotemporal omics. *Lab Chip*. (2021) 21(20):3829–49. doi: 10.1039/d1lc00607j

134. Gerbin KA, Grancharova T, Donovan-Maiye RM, Hendershott MC, Anderson HG, Brown JM, et al. Cell states beyond transcriptomics: integrating structural organization and gene expression in hipsc-derived cardiomyocytes. *Cell Syst*. (2021) 12(6):670–87.e10. doi: 10.1016/j.cels.2021.05.001

frontiers | Frontiers in Cardiovascular Medicine

TYPE Systematic Review
PUBLISHED 28 July 2023
DOI 10.3389/fcvm.2023.1227818

OPEN ACCESS

EDITED BY
Michel Puceat,
Institut National de la Santé et de la Recherche Médicale (INSERM), France

REVIEWED BY
Masanori Obana,
Osaka University, Japan
Yue Qiu,
Fudan University, China

*CORRESPONDENCE
Daniela Čiháková
✉ cihakova@jhmi.edu

RECEIVED 23 May 2023
ACCEPTED 03 July 2023
PUBLISHED 28 July 2023

CITATION
Daoud A, Lema DA, Won T and Čiháková D (2023) Integrative single-cell analysis of cardiac and pulmonary sarcoidosis using publicly available cardiac and bronchoalveolar lavage fluid sequencing datasets.
Front. Cardiovasc. Med. 10:1227818.
doi: 10.3389/fcvm.2023.1227818

Integrative single-cell analysis of cardiac and pulmonary sarcoidosis using publicly available cardiac and bronchoalveolar lavage fluid sequencing datasets

Abdel Daoud[1], Diego A. Lema[2], Taejoon Won[2] and Daniela Čiháková[1,2]*

[1]W. Harry Feinstone Department of Molecular Microbiology and Immunology, Johns Hopkins University Bloomberg School of Public Health, Baltimore, MD, United States, [2]Department of Pathology, Johns Hopkins University School of Medicine, Baltimore, MD, United States

Introduction: Cardiac presentation of autoimmune sarcoidosis, known as cardiac sarcoidosis (CS), is a poorly understood disease with high mortality and low diagnosis rate. While CS is an immunological syndrome, little is known about how cardiac parenchymal and stromal cells mediate its pathogenesis. Moreover, while most current sarcoidosis research is based on research in pulmonary sarcoidosis (PS), it remains unclear how much both presentations of sarcoidosis overlap. To tackle these concerns, we leveraged publicly available sarcoidosis transcriptomic datasets.

Methods: Two publicly available bronchoalveolar lavage single-cell RNA sequencing datasets were integrated to analyze PS relative to control. Additionally, two publicly available cardiac single-nucleus RNA sequencing datasets were integrated to analyze CS relative to control. Following integration, we ran cell-cell communication, transcription factor, and differential expression analyses on parenchymal, stromal, and immune subsets identified in our analysis.

Results: Our analysis revealed that there was an expansion of stromal and immune cells in PS and CS. We also observed upregulation of Th17.1 and attenuated activation transcriptional profiles in the immune cells of CS and PS relative to control. Additionally, we found upregulation of pro-inflammatory and pro-fibrotic transcriptional profiles in the cardiac stromal cells of CS relative to control. We also found that cardiomyocytes exhibited upregulated cardiac stress and proliferation transcriptional profiles in CS relative to control.

Conclusions: Our integrative transcriptomic analysis shows that despite tissue-specific differences, there are shared transcriptional trends between CS and PS. It also shows that stromal and parenchymal populations exhibit transcriptional trends that could explain their pathogenic role in CS.

KEYWORDS

cardiac sarcoidosis (CS), pulmonary sarcoidosis, single-cell RNA-seq (scRNA-seq), single-nucleus RNA-seq, meta-analysis, bioinformactics

Abbreviations

BAL, bronchoalveolar lavage fluid; CM, cardiomyocyte; CS, cardiac sarcoidosis; CXCL12, chemokine CXC motif ligand 12; DC, dendritic cell; ECM, extracellular matrix; FOLR2, folate receptor β; FOXP1, Forkhead box protein P1; GPNMB, glycoprotein Nmb; IFN, interferon; LEC, lymphatic endothelial cell; mTOR, mammalian target of rapamycin; MΦ, macrophage; NF-κB, nuclear factor kappa-light-chain-enhancer of activated B; PCA, principal component analysis; PDGF, platelet-derived growth factor; POSTN, periostin; PS, pulmonary sarcoidosis; scRNA-Seq, single-cell RNA sequencing; SMC, smooth muscle cell; snRNA-Seq single-nuclei RNA sequencing; SYTL3, synaptotagmin like 3; TF, transcription factor; TFPI, tissue factor pathway inhibitor; TGFβ, transforming growth factor β; UMAP, uniform manifold approximation and projection; WNK1, with no lysine/K lysine deficient protein kinase 1.

Introduction

Sarcoidosis is a multisystemic auto-inflammatory syndrome marked by the formation of non-necrotizing granulomas which are defined as small immune cellular aggregates constituted mainly of multinucleated macrophage giant cells as well as $CD4^+$ and $CD8^+$ T lymphocytes and surrounded by epithelial and fibrotic layers (1). While pulmonary sarcoidosis (PS), characterized by lung and/or intrathoracic lymph node involvement, constitutes at least 90% of sarcoidosis cases (2), only 5% of sarcoidosis patients present with clinically overt cardiac sarcoidosis (CS), defined as sarcoidosis with myocardial involvement (3, 4). It is estimated that at least a third of sarcoidosis cases have subclinical CS (4). This is further exacerbated by the fact that at least half CS cases present in an isolated manner, as opposed to a multisystem presentation, with sudden cardiac death constituting the first clinical sign of most of such cases (4). As sarcoidosis is an intricate and poorly understood immunological syndrome (1), many techniques have been utilized to better understand its pathobiology. Single cell resolution RNA sequencing techniques, such as single-cell RNA-sequencing (scRNA-Seq) and single-nucleus RNA-sequencing (snRNA-Seq), are examples of such techniques. Both scRNA-Seq and snRNA-Seq are particularly well-suited to understanding the transcriptional mechanism of multi-agent syndromes, such as sarcoidosis. While bulk RNA-sequencing has been utilized to study sarcoidosis, namely PS, in numerous past studies (5–7), there are very few scRNA-Seq studies that investigate sarcoidotic syndromes. Even PS, the most canonical of such syndromes, has only been investigated once so far via scRNA-Seq in a bronchoalveolar lavage fluid (BAL)-focused study (8). This is problematic because of the myeloid bias of BAL samples as well as the fact that this study failed to truly utilize the potential of the scRNA-Seq technology, including the ability to run cell-cell communication and protein-protein interaction analyses. Recently, a snRNA-Seq study performed transcriptional profiling of cardiac macrophages in CS (9). However, that study lacked a healthy control comparison and was in general more focused on myeloid populations involved in CS. In this study, we attempted to integrate publicly available scRNA-Seq and snRNA-Seq CS and PS datasets with appropriate controls utilizing novel transcriptomic algorithms, such as cell-cell communication and transcription factor analysis workflows. This approach allowed us to transcriptionally profile BAL and cardiac immune cells involved in sarcoidosis pathology, namely macrophages and T cells, as well as cardiac stromal cells implicated in CS, such as cardiac fibroblasts and endothelial cells. Using this approach, we show that, regardless of the tissue of presentation, sarcoidotic macrophages and T cells upregulate activation attenuation transcriptional profiles relative to control. We also show that CS cardiac fibroblasts and endothelial cells are enriched for pro-inflammatory and pro-fibrotic pathways relative to control. In addition, we show that viable cardiomyocytes in CS exhibit upregulated pro-inflammatory and cardiac stress transcriptional pathways relative to control.

Material and methods

Publicly available dataset and patient characterization

Two publicly available scRNA-Seq datasets were utilized to investigate PS, a control BAL dataset as well as a PS BAL dataset. The control scRNA-Seq dataset was retrieved using the accession code: GSE193782 (10). This dataset included 4 healthy control patients and 3 cystic fibrosis patients. Only the healthy control sequencing data were retained for the PS analysis. The PS scRNA-Seq dataset was retrieved using the accession code: GSE184735 (11). This dataset included 4 PS patients, 3 chronic beryllium disease patients and 2 beryllium-sensitized patients. Only the PS patients were retained for the PS analysis. The available clinical characteristics of the patients utilized for the PS analysis are shown in **Supplementary Table S1** (10, 11).

Two publicly available snRNA-Seq were utilized to investigate CS, a control cardiac dataset as well as a CS cardiac dataset. The control snRNA-Seq dataset was retrieved using the accession code: ERP123138 (12). This dataset included sequencing data of 14 healthy control patients collected from various regions of the heart. The CS snRNA-Seq dataset was retrieved using the accession code: GSE205734 (9). This dataset included sequencing data of 4 CS patients and 3 ischemic cardiomyopathy patients collected from the apex of the heart. Only the CS patients were retained for the CS analysis. The available clinical characteristics of the patients utilized for the CS analysis are shown in **Supplementary Table S2** (9, 12).

PS scRNA-Seq workflow

The Seurat scRNA-Seq SCT integration and analysis workflow was utilized (13). Separate Seurat objects were created for each of the PS samples ($n = 4$) using the Read10X() function after which they were merged into one Seurat object using the merge() function. For the healthy samples ($n = 4$), counts were read in using the Read10X() function after which they were merged into one Seurat object. To normalize for sample size, the control samples were downsampled to 8,000 cells. For the quality control of the healthy samples, cells with nFeature_RNA less than 1,850 or more than 7,450, and nCount_RNA more than 85,000 as well as mitochondrial gene content higher than 12% were discarded from further analysis. For the quality control of the PS samples, cells with nFeature_RNA more than 7,100 or less than 2,450, and nCount_RNA more than 87,750 as well as mitochondrial gene content higher than 15% were discarded from further analysis. The healthy and PS Seurat objects were separately normalized via the regularized negative binomial regression method by running the SCTransform() function to obtain SCT count assays, which were utilized for integration. Uniform Manifold Approximation and Projection (UMAP) reduction was run using the first 45 batch corrected principal component analysis (PCA) component. PCA batch correction was done using the Harmony package (14).

For subsequent gene expression analysis, the default assay was reverted to the RNA assay which was consequently log-normalized. Expression of canonical BAL cell type markers (**Supplementary Table S3**) was utilized to provide biological cluster annotations.

CS snRNA-Seq workflow

The Seurat snRNA-Seq SCT integration and analysis workflow was utilized (13). Separate Seurat objects were created for each of the CS samples ($n = 4$) using the Read10X() function after which they were merged into one Seurat object using the merge() function. For the healthy samples ($n = 14$), counts were read in and filtered for those in the apex region using the provided metadata and were merged into one Seurat object. The CS dataset yielded approximately 26,000 nuclei. To normalize for sample size, the control samples were downsampled to 26,000 nuclei. For the quality control of the healthy samples, nuclei with nFeature_RNA less than 450 or more than 3,400, and nCount_RNA more than 7,250 were discarded from further analysis. For the quality control of the CS samples, nuclei with nFeature_RNA less than 520 or more than 2,800, and nCount_RNA more than 7,600 were discarded from further analysis. As cardiomyocytes tend to have a high mitochondrial gene content, the use of mitochondrial gene content as a quality control metric was postponed till biological cluster annotation. The healthy and CS Seurat object were separately normalized via the regularized negative binomial regression method by running the SCTransform() function to obtain SCT count assays, which were utilized for integration. UMAP reduction was run using the first 50 batch corrected PCA component. PCA batch correction was done using the Harmony package (14).

For subsequent gene expression analysis, the default assay was reverted to the RNA assay which was consequently log-normalized. Expression of canonical cardiac cell type markers (**Supplementary Table S4**) was utilized to provide biological cluster annotations. Afterwards, the biological labels were exported, and the complete workflow was repeated with the added step that mitochondrial gene content was used as a quality control metric for non-cardiomyocyte cells. For the healthy sample, non-cardiomyocyte cells with a mitochondrial gene content greater than 5% were excluded. For the CS sample, non-cardiomyocyte cells with a mitochondrial gene content greater than 10% were excluded.

Subclustering workflow

Certain cell type clusters were subclustered to further inspect their transcriptional heterogeneity. For subclustering, only the clusters of the cell type of interest were retained and the SCT and integrated assays were cleared. Afterwards, the workflow starting from the SCT normalization was repeated. Biological subcluster annotations were based on annotation definitions developed by previous scRNA-Seq and snRNA-Seq datasets that inspected similar cell types (9–12).

Differential expression and gene set enrichment analysis

Genes differentially expressed in Sarcoidosis relative to control in cell types of interest were calculated by running the FindMarkers () function using the Wilcoxon Rank Sum test. Only genes with an adjusted p-value less than 0.05 were retained. Gene set enrichment analysis was conducted by obtaining gene sets associated with each differentially expressed gene via Metascape (15). Gene sets of interest were filtered by keeping gene sets with terms of interest using the grepl() text search command. Afterwards, gene sets of interest were retained by manually ensuring that they are of interest. As gene sets tend to poorly represent the direction of effect genes have on their respective gene sets, the sign of the $\log_2$[Average Fold Change] for gene set repressors was inverted. Since custom gene sets were developed for this analysis, a custom gene enrichment score had to be computed for the generated gene sets. This gene enrichment score (GES) was calculated by normalizing $\log_2$[Average Fold Change] of genes in a gene set to the maximum absolute $\log_2$[Average Fold Change] and summing all the adjusted $\log_2$[Average Fold Change] such that a positive score would indicate that a gene set was upregulated while a negative score would indicate otherwise.

$$\text{GES} = \sum_{i=1}^{i=G} \frac{\log_2(\text{avg Fold Change})_i}{\text{MAX}(|\log_2(\text{avg Fold Change})|)}; \quad \text{G: number of genes in a gene set}$$

Cell communication pathway analysis

The CellChat (16) package was used to run cell-cell communication analysis among subclustered macrophage populations, subclustered endothelial cell populations, and fibroblast populations. The normalized RNA assay was used for the ligand-receptor expression analysis with the labels set to the assigned cell type annotations. Significant pathways were defined as pathways with a communication probability higher than 0.2.

Transcription factor analysis

Differentially expressed transcription factors (TFs) were determined by separately running the FindMarkers() using the Wilcoxon Rank Sum test for the healthy and diseased samples. Afterwards, a human TF list, downloaded via the SCENIC package (17), was used to only retain TF genes that have an adjusted p-value less than 0.05. Subsequently, only TFs deemed to be differentially expressed in Sarcoidosis but not in control conditions were retained.

Software

All computational work was run via R v4.0.4 with the following packages additionally loaded: scales v1.1.1, lattice v0.20-41,

gridExtra v2.3, forcats v0.5.1, ggrepel v0.9.1, ggsignif v0.6.1, ggplot2 v3.3.3, multtest v2.46.0, Biobase v2.50.0, BiocGenerics v0.36.1, BiocManager, v1.30.12, patchwork v1.1.1, SeuratObject v4.0.0, Seurat v4.0.1, tidyr v1.1.3, dplyr v1.0.5, CellChat v1.5.0, igraph v1.3.4. Two-sample student *T* tests were run to execute all additional statistical testing at a 95% level of confidence.

Results

Compositional increase of immune and stromal cells in sarcoidosis in the heart and BAL at the cellular level

To examine similarities and differences in cardiac and lung sarcoidosis, we integrated publicly available CS and control cardiac snRNA-Seq datasets (**Supplementary Figure S1**), which yielded a final Seurat object constituted of 25,224 healthy control cardiac nuclei and 24,180 sarcoidotic cardiac nuclei (**Supplementary Figure S2A**). We also integrated publicly available PS and control BAL scRNA-Seq datasets (**Supplementary Figure S1**), which yielded a final Seurat object constituted of 7,423 control BAL cells and 6,296 sarcoidotic BAL cells (**Supplementary Figure S2B**). We observed immune cells, such as macrophages and T cells in both datasets (**Figures 1A**,**B**). Moreover, we identified cardiomyocytes and smooth muscle cells (SMCs) as well as stromal cells, such as endothelial cells, fibroblasts, lymphatic endothelial cells (LECs) and pericytes in the cardiac dataset that were not present in the BAL dataset. In order to confirm that the cell type designations utilized in our analysis represented transcriptionally and biologically distinct clusters, we ran unsupervised differential expression testing (**Figures 1C**,**D**, **Supplementary Figure S2C**, and **Supplementary Table S5**). Furthermore, to assess the proportional changes of different cell types in the cardiac and BAL Seurat objects, cells in each dataset were separated by disease status and the proportion of each cell type was compared across disease (**Figures 1E**,**F**). We found that, relative to control, there was a statistically significant increase in the proportion of macrophages, NK/T cells and endothelial cells in CS. We also noted a statistically significant increase in HLA-DR+ macrophages and epithelial cells in PS relative to control. Furthermore, we observed a statistically significant increase in the proportion of LECs as well as a statistically significant decrease in the proportion of cardiomyocytes and pericytes in CS relative to control (**Figures 1E**,**F**). While the *p*-value for the increase in the proportion of fibroblasts relative to control in the cardiac dataset was not statistically significant at a confidence level of 95%, it was less than 0.1 which suggests that there might be a trending increase in that cell type population in CS. Since the total number of cells for each dataset is approximately equal, the previous observations were applicable to cell type numbers of each dataset as well. Due to the multipotent nature of pericytes, we ran pseudotime trajectory analysis on pericytes, SMCs and fibroblasts. We noted that very few pericytes were contracted along the transitional path towards SMCs while most of the pericytes were diminished in sarcoidosis along the early portion of the transitional path towards fibroblasts in CS (**Supplementary Figure S2D**). Conversely, the proportion of the cells along the later portion of that path were shown to be expanded. Therefore, both sarcoidotic cardiac and BAL samples show a compositional shift towards increased immune cells, such as macrophages and T cells, as well as stromal cells, such as epithelial cells, endothelial cells and fibroblasts.

Sarcoidotic cardiac and BAL T cells exhibit immune attenuation and dysfunction

Interferon (IFN) γ-producing Th17 cells (Th17.1) are proposed to play a pathogenic role in PS (18). This subset has been functionally defined as $CD4^+$ T cells that produce both the Th17 cytokine, IL-17A, as well as the Th1 cytokine, IFNγ (19). To assess the contribution of this subtype in CS as well, we examined the proportion and number of Th17.1 cells, defined as cells/nuclei co-expressing *TBX21* and *RUNX1* in the heart and BAL. This definition has been previously verified as a valid transcriptional definition of Th17.1 (20). Qualitatively, we detected significantly more Th17.1 nuclei in sarcoidosis relative to controls in the heart (**Figure 2A**). While we saw more sarcoidotic Th17.1 cells in the BAL relative to control, there was not enough Th17.1-containing control donors to determine significance. The reason was that one control donor that yielded the only control BAL Th17.1 cell was determined to be an outlier by running Grubb's test for outliers and was accordingly removed from analysis. We also observed a trending increase in the proportion of Th17.1 cells in sarcoidotic cardiac T cells compared to control cardiac T cells (**Supplementary Figure S3A**). Both sarcoidotic cardiac and BAL Th17.1 constituted 6%–8% of all sarcoidotic cardiac and BAL T cells (**Supplementary Figure S3A**). Since none of the remaining control BAL T cells passed the Th17.1 transcriptional definition, we did not run any further comparison with the BAL dataset. Consequently, we ran differential expression testing between all cardiac Th17.1 cells and the other cardiac T cells (**Supplementary Table S6**). We discerned that only 2 genes were significantly upregulated in Th17.1 relative to all other cardiac T cells: *TBX21*, which encodes T-bet, and *AOAH*, a Th1 response transcription factor (**Supplementary Figure S3B**). We verified that these two genes were strong Th17.1 markers via ROC analysis (**Supplementary Figure S3C** and **Supplementary Table S6**). Hence, we show that both cardiac and BAL sarcoidotic T cells upregulated their composition of Th17.1 cells that are transcriptionally programmed to elicit both a Th1 and Th17 phenotype.

To determine how the transcriptional phenotype of sarcoidotic T cells differs in the heart and BAL, supervised gene-set enrichment was conducted (**Supplementary Table S7**). We noted that, relative to control, sarcoidotic cardiac and BAL T cells downregulated TCR signaling, immune response, IFN signaling and cytokine signaling transcriptional pathways (**Figure 2B**). In addition, relative to control, sarcoidotic cardiac and BAL T cells

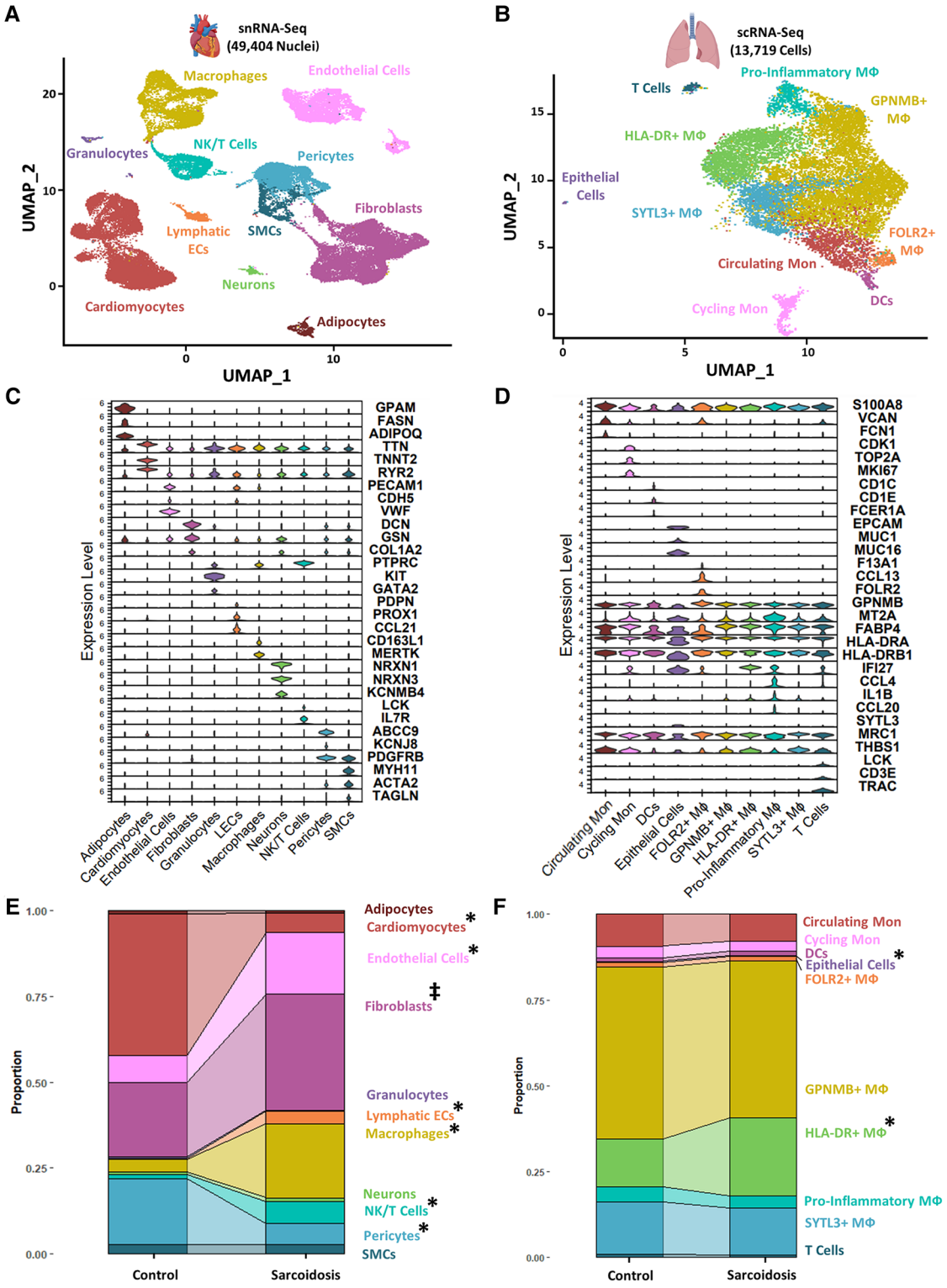


FIGURE 1
(A) UMAP clustering of integrated control and cardiac sarcoidosis (CS) snRNA-Seq datasets by cell type. The following biological cell populations were identified: adipocytes, cardiomyocytes, endothelial cells, fibroblasts, granulocytes, lymphatic ECs, macrophages (MΦ), neurons, NK/T cells, pericytes and smooth muscle cells (SMCs). (B) UMAP clustering of integrated control and pulmonary sarcoidosis (PS) scRNA-Seq datasets by cell type. The following biological cell populations were identified: circulating monocytes (Mon), cycling Mon, epithelial cells, FOLR2+ MΦ, GPNMB+ MΦ, HLA-DR+ MΦ, pro-inflammatory MΦ, SYTL3+ MΦ and T cells. (C) Violin plot showing the expression of canonical cardiac markers in each identified biological cluster. (D) Violin plot showing the expression of canonical BAL markers in each identified biological cluster. (E) Proportion analysis of each identified cardiac biological cluster across disease status. (F) Proportion analysis of each identified cardiac biological cluster across disease status. Statistically significant changes with a p-value <0.05 are indicated by a (*) while statistically trending changes with a p-value <0.1 are indicated by (‡).

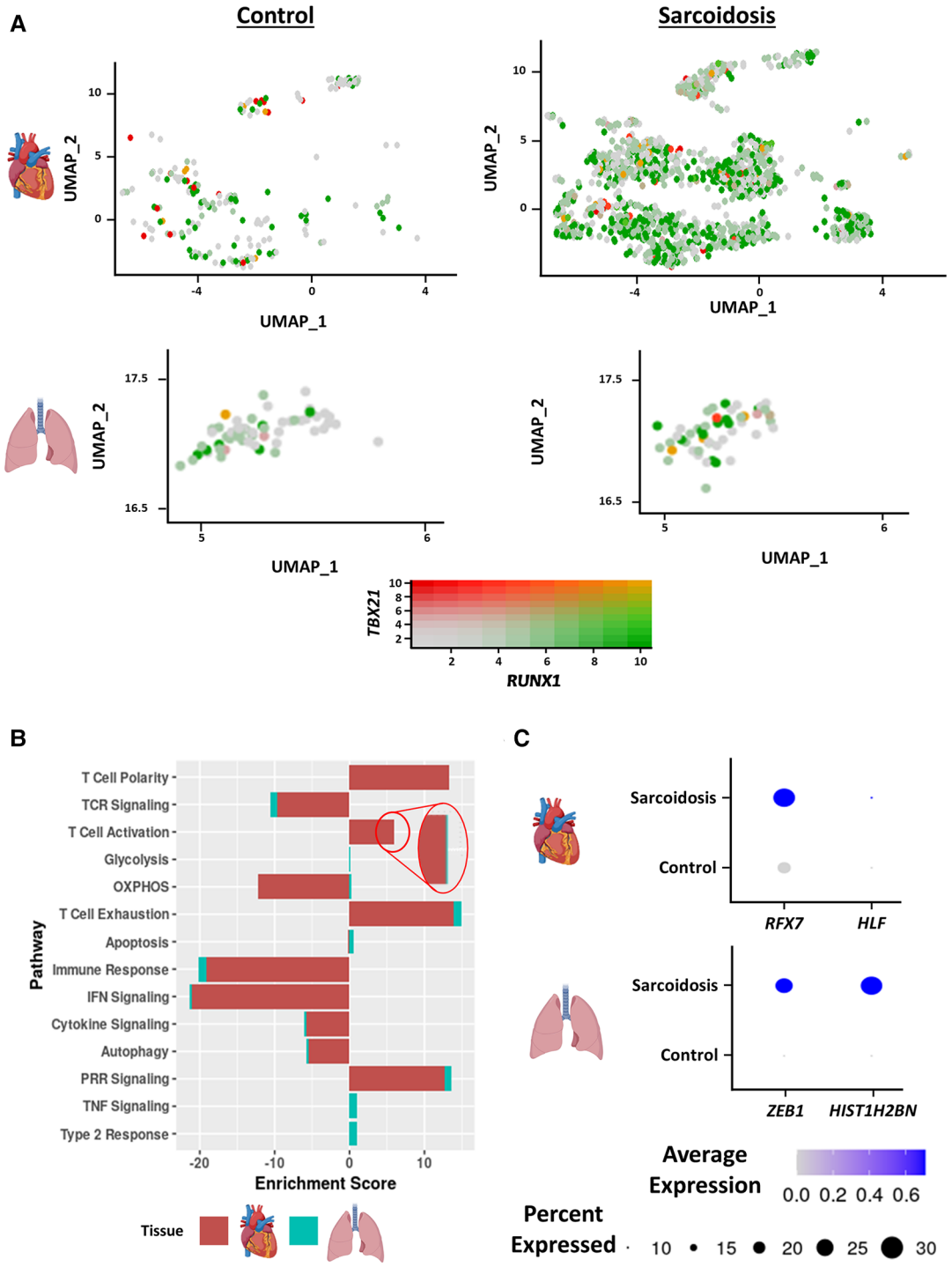


FIGURE 2
(A) Feature plot showing Th17.1 cells co-expressing *TBX21* and *RUNX1* in the cardiac (top) and BAL (bottom) datasets stratified by disease status (control shown on the left, sarcoidosis shown on the right). (B) Supervised gene-set enrichment results in cardiac and BAL T cells showing pathways shown to be significantly enriched in sarcoidosis relative to control. Positive enrichment scores indicate pathways shown to be significantly upregulated while negative enrichment scores indicate otherwise. (C) Dot plot showing the expression of sarcoidosis-specific transcription factors (TFs) in the cardiac and BAL datasets in control and sarcoidosis T cells. Gene expression was scaled from 0 to 0.6.

exhibited upregulation of immune dysfunction pathways, such as PRR signaling, autophagy attenuation, T cell activation and exhaustion. While apoptosis and oxidative phosphorylation pathways were upregulated in sarcoidotic BAL T cells, these processes were downregulated in sarcoidotic cardiac T cells (Figure 2B). Other transcriptional distinctions included that, relative to control, only BAL T cells showed upregulation of TNF signaling and type 2 response pathways while only cardiac T cells showed upregulation of the T cell structural polarity pathway (Figure 2B).

To inspect whether there was a shared transcriptional programming pathway for T cells in the heart and BAL, transcription factor (TF) analysis was conducted (**Supplementary Table S8**). We found that, unlike their control counterparts, both cardiac and BAL sarcoidosis T cells showed enrichment of memory T cell TF genes (*RFX7* and *ZEB1*), survival and proliferation TF genes (*HLF* and *HIST1H2BN*) (**Figure 2C**) and type 3 response TF genes (*ARNTL* and *ADARB1*) (**Supplementary Figure S3D**). Meanwhile, only sarcoidotic BAL T cells showed enrichment of *ZNF683*, a tissue residency TF, while only sarcoidotic cardiac T cells showed enrichment of *NR3C1*, a TF that has been associated with dysfunctional terminal T cell activation (**Supplementary Figure S3D**). Therefore, we demonstrate that both sarcoidotic cardiac and BAL T cells are transcriptionally programmed towards immune dysfunction processes, such as autophagy attenuation, exhaustion and Th17.1 response, as well as the attenuation of immune processes, such as IFN, cytokine and TCR signaling. However, while only sarcoidotic BAL T cells upregulated apoptosis and type 2 pathways, such as oxidative phosphorylation and TNF signaling, only sarcoidotic cardiac T cells exhibited activation priming profiles, indicated by downregulated apoptosis and oxidative phosphorylation as well as upregulated structural T cell polarity.

Sarcoidotic cardiac and BAL MΦ exhibit attenuated alternative activation

To inspect similarities and differences between macrophages (MΦ) in CS and PS, we subclustered cardiac MΦ into 6 subgroups: CD16 MΦ, Glycoprotein Nmb (GPNMB)+ MΦ, HLA-DR+ MΦ, resident MΦ, Synaptotagmin Like 3 (SYTL3)+ MΦ as well as CD1C+ dendritic cells (DCs) (**Figure 3A**). These annotations have been previously validated in cardiac MΦ by a previous transcriptomic study of CS (9). We further verified these MΦ subgroup annotations using canonical cell type markers as well as unsupervised differential expression testing (**Supplementary Figures S4A,B**, and **Supplementary Table S9**). These MΦ subpopulations were identifiable in the BAL dataset without further need for subclustering (**Figure 1A**). To inspect interactions between these subsets, we ran cell-cell communication analysis in cardiac and BAL MΦ (**Supplementary Figure S4C**, **Supplementary Table S10**). We observed that the expression of *CXCR4* [receptor for Chemokine CXC Motif Ligand 12 (CXCL12)], *SIGLEC1* (IFN-related factor), and *LGALS9* (immune checkpoint ligand) were upregulated in sarcoidotic BAL MΦ relative to control BAL MΦ (**Figure 3B**).

To further assess whether certain transcriptional pathways are conserved in sarcoidotic MΦ in the heart and BAL, we conducted supervised gene-set enrichment (**Supplementary Table S11**). We found that, relative to control, both cardiac and BAL sarcoidosis MΦ exhibited upregulated alternative activation and glycolysis pathways as well as downregulated cytokine signaling, phagocytosis/efferocytosis and autophagy pathways (**Figure 3C**). Gene-set enrichment showed that certain pathways were, however, tissue-specific in sarcoidotic MΦ. Relative to control, sarcoidotic cardiac MΦ downregulated apoptosis, but apoptosis was upregulated in sarcoidotic BAL MΦ (**Figure 3C**). In addition, relative to control, we noted that CXCL12 signaling was upregulated only in sarcoidotic cardiac MΦ while immune checkpoint signaling was upregulated only in sarcoidotic BAL MΦ. Additionally, only sarcoidotic BAL MΦ displayed downregulation of IFN signaling relative to control (**Figure 3C**, right).

To assess whether the transcriptional pathway results could be corroborated at the level of TF expression, we conducted TF analysis of sarcoidosis-specific TFs (**Supplementary Table S12**). We observed that, unlike their control counterparts, both sarcoidotic BAL and cardiac GPNMB+ MΦ enriched cytokine suppression TFs (*CREM* and *CUX1* as well as *NR1H3*) and alternative activation TFs (*JUND* and *ZNF331* in BAL MΦ as well as *MDM2* and *FOSL2* in cardiac MΦ) (**Figure 3D** and **Supplementary Figure S4D**). Other TFs enriched only in sarcoidotic BAL and cardiac MΦ included chronic immunity TFs (*ETS2* by BAL SYTL3+ MΦ and *ARID5B* by cardiac GPNMB+ MΦ) as well as mTOR TFs (*SREBF2* by BAL SYTL3+ MΦ and *RUNX2* by cardiac HLA-DR+ MΦ) (**Figures 3D**). Only sarcoidotic cardiac MΦ, however, revealed enrichment of *VDR*, a macrophage recruitment TF (**Supplementary Figure S4E**). Therefore, we show that, relative to control, both cardiac and BAL sarcoidotic MΦ exhibited attenuated alternative activation transcriptional phenotypes, as determined by the suppression of cytokine signaling, autophagy and phagocytosis/efferocytosis pathways, as well as by the simultaneous upregulation of glycolysis and type 2 response pathways. However, relative to control, only cardiac sarcoidotic MΦ presented with upregulation of the chemotaxis pathway and downregulation of the apoptosis pathway. Conversely, relative to control, only BAL MΦ displayed upregulation of immune checkpoint and apoptosis pathways as well as downregulation of the IFN signaling pathway.

Fibroblasts exhibit upregulated fibrotic, pro-inflammatory, and dysfunctional activation profiles in CS

To further investigate cardiac fibroblasts in CS, we subclustered cardiac fibroblasts into 6 biologically relevant and transcriptionally distinct fibroblast clusters: quiescent fibroblasts, ECM fibroblasts which were defined as fibroblasts highly expressing extracellular matrix (ECM)-related genes, myofibroblasts which were defined as fibroblasts highly expressing activation markers, inflammatory fibroblasts which were defined as fibroblasts highly expressing inflammatory markers, endothelial fibroblasts which were defined as fibroblasts highly expressing endothelial cell markers, and cardiomyocyte (CM)-like fibroblasts which were defined as fibroblasts highly expressing CM markers (**Figure 4A**, **Supplementary Figures S5A,B**, and **Supplementary Table S13**). We ran cell proportion analysis and showed that there was a statistically significant increase in the proportion of

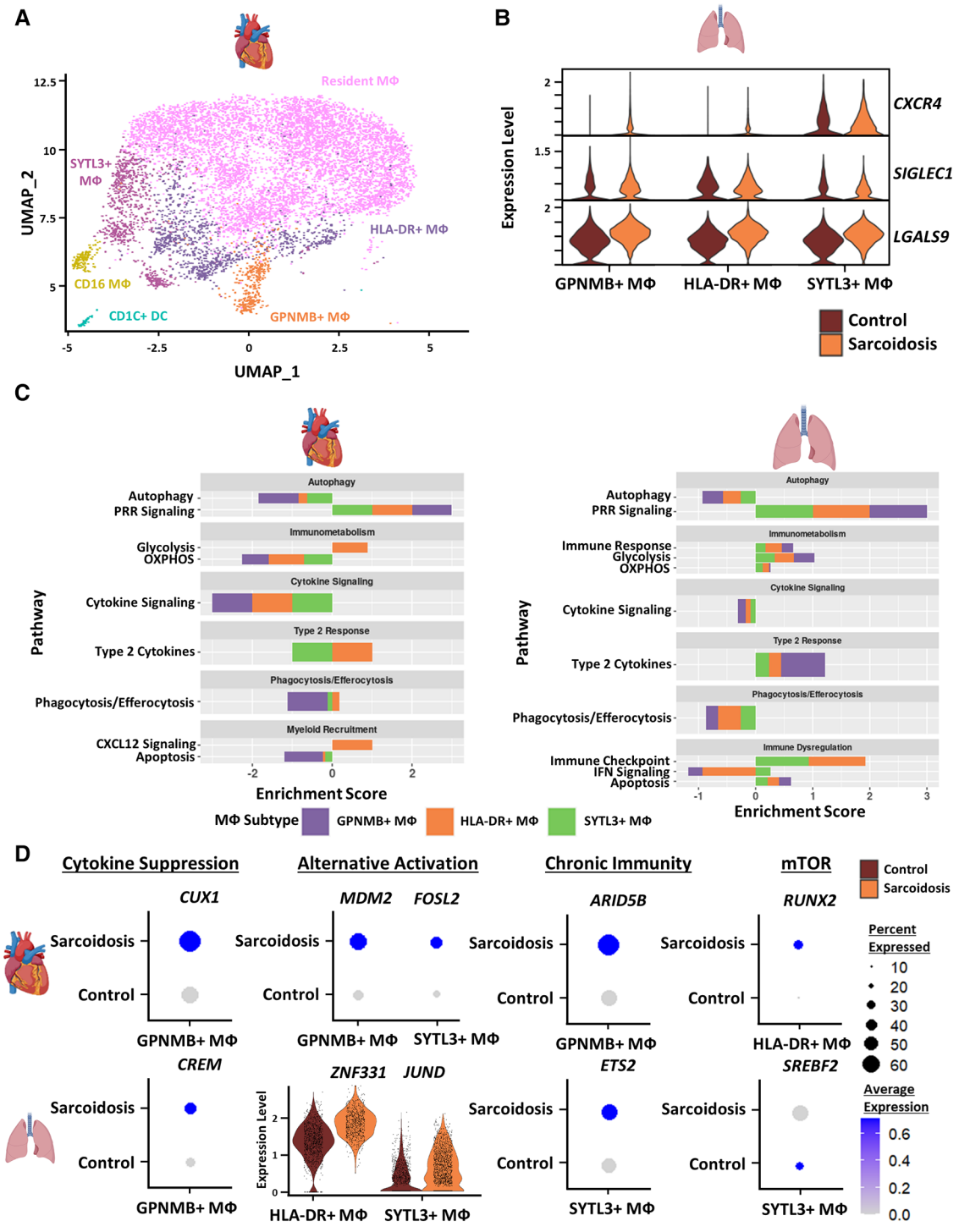


FIGURE 3
(A) UMAP clustering of subclustered cardiac myeloid subpopulations by cell type. The following myeloid subpopulations were identified: Resident MΦ, GPNMB+ MΦ, HLA-DR+ MΦ, CD16 MΦ, SYTL3+ MΦ and CD1C dendritic cells (DCs). (B) Violin plot showing gene expression of cell-cell communication ligands of interest differentially upregulated in sarcoidotic BAL MΦ subpopulations stratified by cell type and disease status. (C) Supervised gene-set enrichment results in cardiac and BAL MΦ subpopulations showing pathways shown to be significantly enriched in sarcoidosis relative to control. Positive enrichment scores indicate pathways shown to be significantly upregulated while negative enrichment scores indicate otherwise. (D) Dot plots showing the expression of sarcoidosis-specific transcription factors (TFs) in the cardiac and BAL datasets in control and sarcoidosis MΦ subpopulations. Average gene expression was scaled from 0 to 0.6. The expression of *ZNF331* and *JUND* was visualized using a violin plot stratified by cell type and disease status for aesthetic purposes. The classification group of each TF is shown colored above its respective gene expression plot.

myofibroblasts (Figure 4B). We also found a trending increase in the proportion of inflammatory and endothelial fibroblasts.

Subsequently, we conducted supervised gene-set enrichment to assess the transcriptional phenotypes of sarcoidotic cardiac fibroblasts (Supplementary Table S14). We observed that, relative to control, all sarcoidotic cardiac fibroblasts exhibited an increased pro-inflammatory transcriptional phenotype, as indicated by upregulated glycolysis, immune response and IL-1 signaling as well as downregulated HLA class I, susceptibility to T/NK mediated cytotoxicity, type 2 cytokines and oxidative phosphorylation (Figure 4C, left). Other transcriptional phenotypes differentially upregulated by sarcoidotic cardiac fibroblasts included tissue remodeling processes, such as ECM, Transforming Growth Factor β (TGFβ) and Platelet-Derived Growth Factor (PDGF) signaling and wound healing, as well as dysfunctional activation processes, such as autophagy, mammalian target of rapamycin (mTOR) signaling, Forkhead Box Protein P1 (FOXP1) signaling, apoptosis, migration and FOXP1 signaling. Only sarcoidotic inflammatory fibroblasts, however, upregulated HLA class II as well as IFN signaling pathways, relative to control (Figure 4C, right). Similarly, only ECM fibroblasts and myofibroblasts upregulated the TNF signaling pathway, relative to control. While only quiescent and inflammatory fibroblasts showed upregulation of the epithelial-mesenchymal transition pathway, relative to control (Figure 4C, right).

To assess how these enrichment analysis results corroborated with TF trends, we ran TF analysis (Supplementary Table S15). We observed that, unlike their control counterparts, only sarcoidotic inflammatory fibroblasts showed enrichment for immune response TFs, such as *PRDM1*, only sarcoidotic ECM fibroblasts showed enrichment of apoptosis TF, such as *MECOM*, and only sarcoidotic myofibroblasts and ECM fibroblasts showed enrichment of tissue remodeling TFs, such as *MEF2A* (Figure 4D). Unlike myofibroblasts and ECM fibroblasts which enriched fibroblast activation TFs, such as *GLIS1* and *TRPS1*, only in sarcoidosis, endothelial fibroblasts showed enrichment of *FLI1*, a fibroblast activation suppressor TF, and *GATA2*, an endothelial cell TF, strictly in sarcoidosis (Supplementary Figure S5C). Therefore, we show via gene-set enrichment and TF analysis that CS promotes transcriptional profiles characterized by upregulated tissue remodeling processes, such as ECM, TGFβ and PDGF signaling and wound healing, upregulated immunomodulatory processes, such as glycolysis, immune evasion and type 1 response, and dysfunctional activation, characterized by upregulated autophagy, apoptosis as well as migration.

Endothelial cells exhibit more immune modulation, angiogenesis and mTOR signaling in CS

To further study endothelial cells in CS, we subclustered cardiac endothelial cells into 7 transcriptionally distinct subtype clusters: capillary endothelial cells which highly express capillary markers, arterial endothelial cells which highly express arterial markers, Periostin (POSTN)+ endothelial cells, cardiomyocyte (CM)-like endothelial cells which highly express CM markers, immune endothelial cells which highly express immunomodulatory markers, venous endothelial cells which highly express venous markers, and fibroblast-like endothelial cells which highly express tissue remodeling markers (Figure 5A, Supplementary Figures S6A,B, Supplementary Table S16). While epithelial cells were identified in the BAL dataset, they were not detected in CS. We noted via cell proportion analysis that there was a statistically significant increase in fibroblast-like endothelial cells and a statistically significant decrease in CM-Like endothelial cells (Figure 5B). We also found a trending increase in the proportion of POSTN+ endothelial cells.

Differential expression and supervised gene-set enrichment testing was conducted in each sarcoidotic endothelial cell type relative to its control counterpart (Supplementary Table S17). As there was not enough control fibroblast-like endothelial cells to compare sarcoidotic Fib-like endothelial cells against, fibroblast-like endothelial cells were removed from enrichment analysis. In addition, CM-like endothelial cells were discarded from further analysis as there was an insufficient number of sarcoidotic CM-like endothelial cells. Regardless of subtype, all analyzed sarcoidotic endothelial cells showed increased immune transcriptional activation relative to control, as indicated by upregulation of glycolysis, immune response and immune synapse formation as well as by downregulation of HLA class I, susceptibility to T/NK-mediated cytotoxicity and oxidative phosphorylation (Figure 5C, left). Other transcriptional phenotypes included upregulation of angiogenesis pathways, such as coagulation and With No Lysine/K Lysine Deficient Protein Kinase 1 (WNK1) signaling, and downregulation of mTOR-dependent autophagy. While sarcoidotic arterial endothelial cells showed upregulation of TNF and IFN signaling pathways, they exhibited downregulation of endothelial cell activation pathways, as indicated by upregulated FOXP1 signaling (Figure 5C, right). Additionally, most endothelial cells, except for immune and POSTN+ endothelial cells, displayed upregulation of tissue remodeling pathways, such as TGFβ signaling and ECM. While sarcoidotic POSTN+ endothelial cells did not show clear upregulation of tissue remodeling pathways, they were the only subtype to upregulate pro-inflammatory pathways, such as IL-1, IL-2 and IL-23 signaling (Figure 5C, right).

To further assess transcriptional programs utilized by sarcoidotic endothelial cells, TF analysis was conducted (Supplementary Table S18). As venous and capillary endothelial cells mostly represent anatomical niches that are unlikely to be functionally unique in CS (Figure 5C), these subtypes were discarded from subsequent TF analysis. We found that, unlike their control counterparts, all analyzed sarcoidotic endothelial cells highly expressed unique angiogenic TFs, such as *VEGFA*, *PRDM1*, *CREB5* and *FOSB* (Figure 5D). We also found that, unlike their control counterparts, all analyzed sarcoidotic endothelial cells (except for arterial endothelial cells) were enriched for unique proliferation TFs, such as *PBX4*, *TCF4* and *RFX2*. Conversely, unlike their control counterparts, only sarcoidotic fibroblast-like endothelial cells displayed enrichment of *SPI1*, a tissue remodeling TF, and *IKZF1*, an

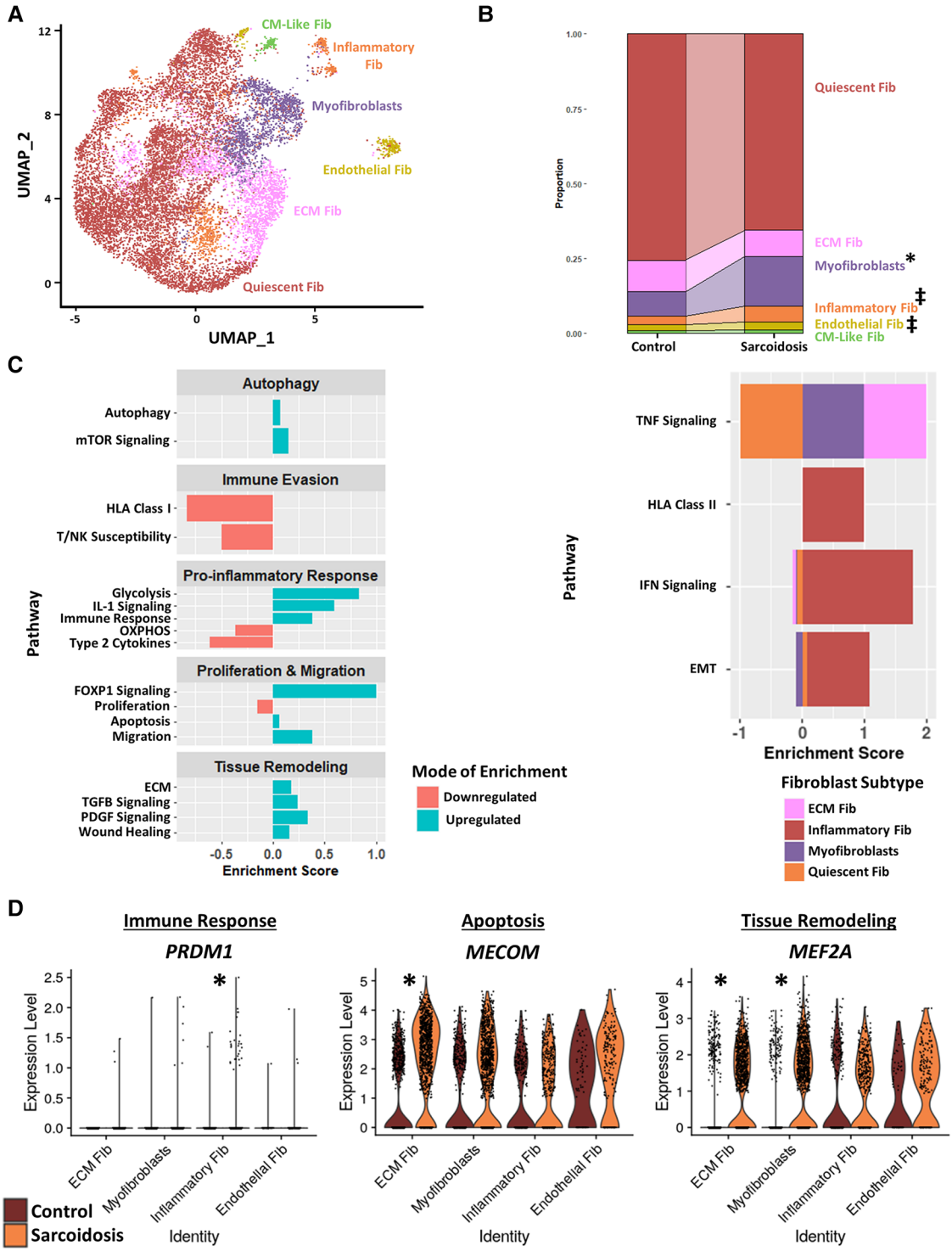


FIGURE 4
(A) UMAP clustering of subclustered cardiac fibroblast subpopulations by cell type. The following fibroblast subpopulations were identified: quiescent fibroblasts (Fib), ECM Fib, myofibroblasts, inflammatory Fib, endothelial Fib and CM-like Fib. (B) Proportion analysis of each identified cardiac fibroblast cluster across disease status. Statistically significant changes with a *p*-value <0.05 are indicated by a (*) while statistically trending changes with a *p*-value <0.1 are indicated by (‡). (C) Supervised gene-set enrichment results in cardiac fibroblast subpopulations showing pathways shown to be significantly enriched in sarcoidosis relative to control. Pathways shown to be significantly enriched in all fibroblast clusters are shown on the left while subpopulation-specific patterns are shown on the right. Positive enrichment scores indicate pathways shown to be significantly upregulated while negative enrichment scores indicate otherwise. (D) Violin plots stratified by cell type and disease status showing the expression of sarcoidosis-specific transcription factors (TFs) in the cardiac datasets in control and sarcoidosis fibroblast subpopulations. The classification group of each TF is shown colored above its respective gene expression plot.

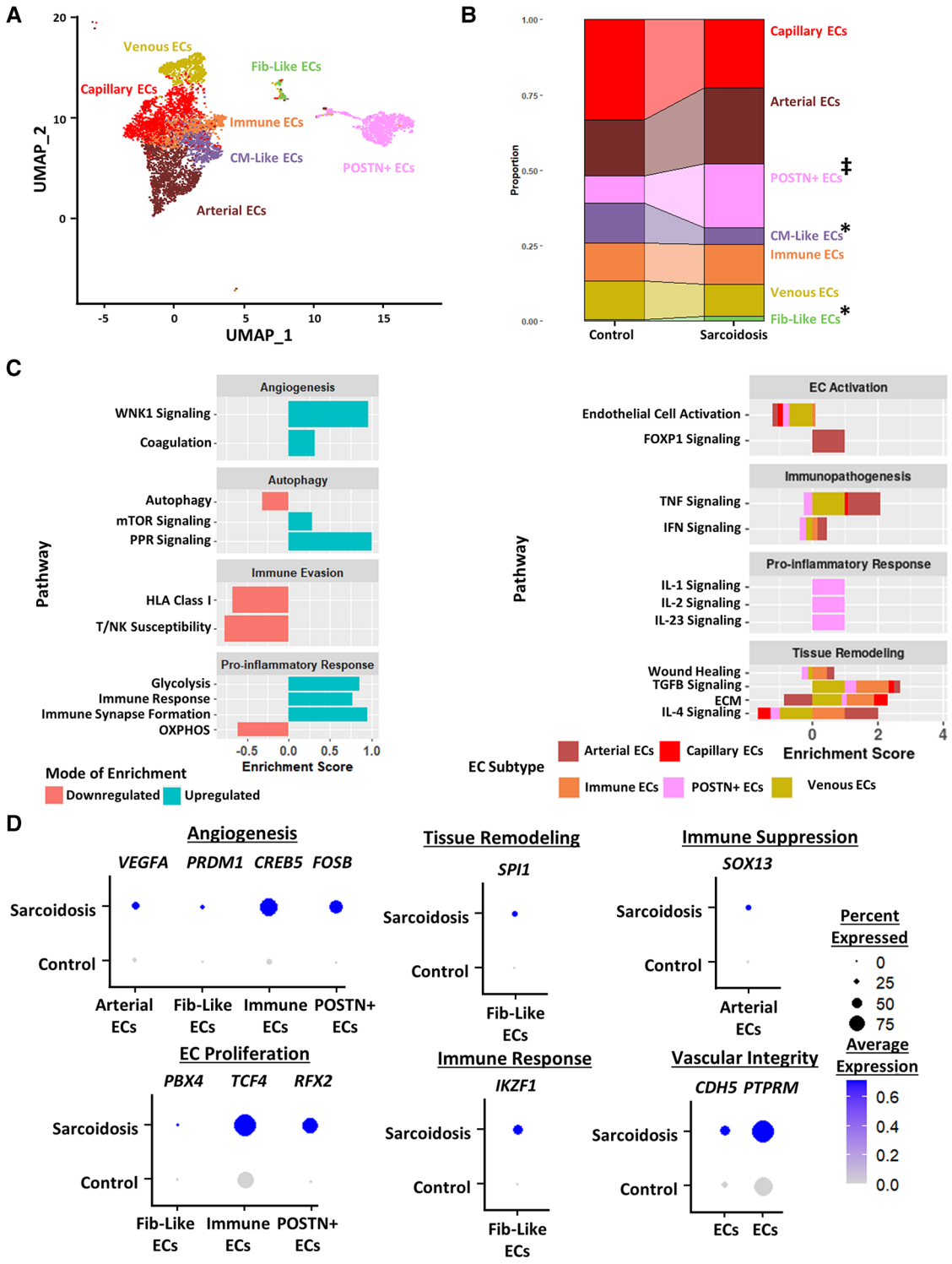


FIGURE 5
(A) UMAP clustering of subclustered endothelial cell subpopulations by cell type. The following fibroblast subpopulations were identified: capillary endothelial cells (ECs), arterial ECs, POSTN+ CM-like ECs, immune ECs, venous ECs and Fib-like ECs. (B) Proportion analysis of each identified cardiac endothelial cell cluster across disease status. Statistically significant changes with a *p*-value <0.05 are indicated by a (*) while statistically trending changes with a *p*-value <0.1 are indicated by (‡). (C) Supervised gene-set enrichment results in cardiac endothelial subpopulations showing pathways shown to be significantly enriched in sarcoidosis relative to control. Pathways shown to be significantly enriched in all endothelial cell clusters are shown on the left while subpopulation-specific patterns are shown on the right. Positive enrichment scores indicate pathways shown to be significantly upregulated while negative enrichment scores indicate otherwise. (D) Dot plots stratified showing the expression of sarcoidosis-specific transcription factors (TFs) in the cardiac datasets in control and sarcoidosis endothelial cell subpopulations. Average gene expression was scaled from 0 to 0.6. The classification group of each TF is shown colored above its respective gene expression plot.

immunomodulatory TF (Figure 5D). Meanwhile only sarcoidotic arterial endothelial cells were enriched for *SOX13*, an immunosuppressive TF. Additionally, all sarcoidotic endothelial cells showed upregulation of *CDH5* and its receptor *PTPRM* relative to control, as was indicated by cell-cell communication analysis (Figure 5D, Supplementary Figures S6C,D). Therefore, we demonstrate using gene-set enrichment and TF analysis that sarcoidotic endothelial cells transcriptionally upregulate mTOR signaling, immunomodulatory processes, such as glycolysis and immune response, as well as angiogenesis, characterized by CDH5-PTPRM communication. However, sarcoidotic endothelial cells exhibit heterogeneity in promoting tissue remodeling and cytokine signaling, as we showed that while the TNF signaling pathway was upregulated by arterial and immune endothelial cells, the IFN signaling pathway was uniquely upregulated by immune endothelial cells.

CS promotes pro-inflammatory, cardiotonic and proliferative profiles in cardiomyocytes

Due to the cardiac symptomology of CS, we next investigated the transcriptomic changes in sarcoidotic cardiomyocytes. Despite integration and harmony-based batch effect correction, most sarcoidotic cardiomyocytes identified initially in the heart (Figure 1A) clustered in a cluster termed "CM 3" while most control cardiomyocytes clustered in a cluster termed "CM 1" (Supplementary Figure S7A). Additionally, there was a third distinct cardiomyocyte cluster termed "CM 2", constituted mostly of sarcoidotic cardiomyocytes. To further analyze these three clusters, cardiomyocytes were isolated from the heart dataset for more granular integration and batch correction as well as for sub-clustering. Even after re-clustering, the three cardiomyocyte clusters were still spatially segregated by disease status (Figure 6A). Differential expression testing showed that the three clusters were transcriptionally distinct (Supplementary Figure S7B, Supplementary Table S19). Cell proportion analysis revealed that while there was no statistically significant difference in the proportion of cells in the CM 1 and CM 3 clusters across disease, there was a statistically significant increase in the proportion of cells in the CM 2 cluster in CS relative to control (Figure 6B).

To assess the functional heterogeneity of these cardiomyocyte clusters, we conducted supervised gene-set enrichment testing (Supplementary Table S20). We showed that CM 2 and CM 3 showed differential immune response activation, as indicated by upregulated WNK1 signaling, glycolysis, cytokine signaling, Nuclear Factor Kappa-light-chain-enhancer of Activated B (NF-κB) signaling, innate immunity and inflammasome activation (Figures 6C,D). CM 2 and CM 3 also displayed upregulated cardiac burden transcriptional processes, such as ECM and TGFβ signaling, and cardiac function, as indicated by upregulated calcium signaling, cardiac action potential, and cardiac myofibril assembly. In addition, CM 2 and CM 3 showed differential transcriptional upregulation of cardiomyocyte proliferation processes, as shown by downregulated apoptosis. While CM 3 showed upregulated IL-1 signaling, heart contraction and cardiac tissue regeneration, these pathways were downregulated in CM 2. Conversely, while mTOR and IFN signaling were downregulated in CM 3, these pathways were upregulated in CM 2. Furthermore, certain enrichment patterns, such as upregulation of TNF signaling and downregulation of FOXP1 signaling, a critical TF for cardiomyocyte proliferation, were unique to CM 2. These results were further corroborated by the observation that *TNNI3*, encoding for troponin I, was highly downregulated in CM (Supplementary Figure S7C). Therefore, we show that, despite transcriptional heterogeneity, sarcoidotic cardiomyocytes transcriptionally promote immune response processes, such as glycolysis, cytokine signaling, innate immunity and type 1 response, cardiac function processes, such as calcium signaling, heart contraction, cardiac action potential and cardiac myofibril assembly, and stress response processes, such as cardiac burden and cardiac tissue regeneration, as well as proliferation processes, via downregulating apoptosis.

Discussion

CS presents a significant public health burden, with high mortality risk (3, 4). However, the pathogenesis of this disease is largely understudied. In this transcriptomic study, we dissected the transcriptional complexity of major cardiac resident and immune cell types that might be involved in CS pathogenesis, comparing them to counterpart BAL cell types when possible. For this purpose, we separately integrated healthy and sarcoidosis cardiac snRNA-Seq as well as BAL scRNA-Seq datasets. We observed an expansion in immune (myeloid and lymphoid) as well as stromal populations, such as fibroblasts and endothelial cells, in both PS and CS, relative to control. This finding was in line with what has been shown about granuloma structure in different forms of sarcoidosis, such as renal and pancreatic sarcoidosis (21, 22). Interestingly, CS exhibited a statistically significant decrease in the proportion of pericytes relative to healthy heart control. As pericytes are known to differentiate into myofibroblasts (23), it could be argued that the observed reduction in pericytes might indicate that a portion of cardiac pericytes could be redirected into the fibroblast population in CS. This is corroborated by our findings that the fibroblast population, specifically myofibroblasts, were shown to be significantly expanded in CS relative to control. This hypothesis was further supported by our trajectory analysis of pericytes and fibroblasts that showed that the proportion of cells midway through the pericyte-fibroblast differentiation trajectory were higher in CS relative to control.

While it has been shown that T cells are a major constituent of sarcoidotic granulomas (1), studies investigating the pathogenic role of T cells in CS are lacking. Recently, there has been a paradigm shift from Th1 to Th17.1 (which are IFNγ+ IL-17A+) as the pathogenic T cell population in PS (18). This subset has also been implicated as potential mediators of sarcoidosis-like complications following CAR T cell treatment (24). Here, we show that Th17.1, which we defined based on previous *in vitro* work on IL-17A+ IFNγ+ T cells as *TBX21*+ *RUNX1*+ (20), expand in CS relative to control as well. While this expansion

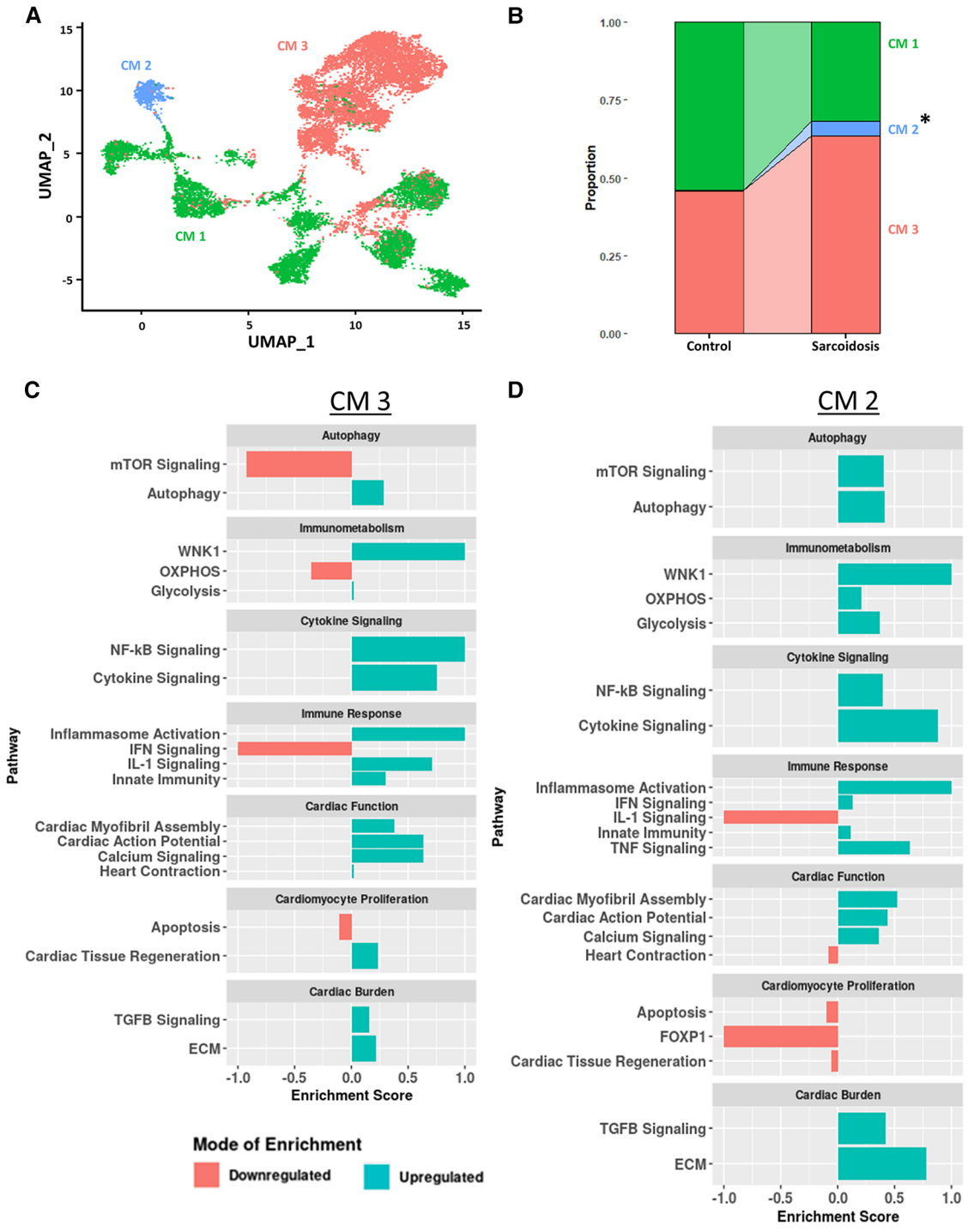


FIGURE 6
(A) UMAP clustering of subclustered cardiomyocyte subpopulations by cell type. Three distinct cardiomyocyte subclusters were identified termed: CM 1, CM 2 and CM 3. (B) Proportion analysis of each identified cardiomyocyte subcluster across disease status. Statistically significant changes with a *p*-value <0.05 are indicated by a (*). (C) Supervised gene-set enrichment results in cluster CM 3 showing pathways shown to be significantly enriched in CM 3 relative to CM 1 and CM 2. Positive enrichment scores indicate pathways shown to be significantly upregulated while negative enrichment scores indicate otherwise. (D) Supervised gene-set enrichment results in cluster CM 2 showing pathways shown to be significantly enriched in CM 2 relative to CM 1 and CM 3. Positive enrichment scores indicate pathways shown to be significantly upregulated while negative enrichment scores indicate otherwise.

was only trending in PS, that might be due to the myeloid bias of BAL samples. We also showed that, unlike their BAL counterparts, cardiac Th17.1 are best defined as *TBX21*+ *AOAH*+. This was reasonable as *TBX21* and *AOAH* mediate Th1 polarization, a known characteristic of Th17.1. Additionally, we showed via gene-set enrichment and transcription factor analysis that CS and PS exhibit downregulated immune response, cytokine and TCR signaling as well as upregulated chronic exhaustion pathways

relative to control. This was well-corroborated by other studies that reported that BAL-derived T cells exhibited low proliferative and cytokine signaling phenotypes in PS (25). While there was a study that reported increased Th1 cytokine signaling in CS (26), it was mostly focused on active disease and cytokine levels assessed via bulk measurements throughout large histological samples. While we observed that only BAL T cells upregulated apoptosis, TNF signaling and type 2 response processes, such as OXPHOS, relative to control, which has been previously reported in the PS literature (25, 27, 28), cardiac T cells were shown to downregulate OXPHOS with insufficient glycolysis compensational upregulation. Metabolic energy is critical for T cell activation and differentiation (29). Such metabolic insufficiency has been linked with quiescent and suppressed T cells (30). This transcriptional metabolic trend is well-corroborated with the fact that we observed downregulated activation profiles in CS T cells relative to their control counterparts. However, more protein and functional T cell studies in CS are needed to properly elucidate this transcriptional phenotype. Nevertheless, we show that both PS and CS T cells exhibit an attenuated activation profile as well as expansion of Th17.1 populations relative to control.

There has been a recent renewed interest in the pathogenic role of macrophages, a core constituent of sarcoidotic granulomas. For instance, a single cell transcriptomic analysis on cardiac macrophages in CS and ischemic cardiomyopathy has recently been reported (9), focusing mostly on the upregulation of mTOR signaling in CS. However, the pathogenic role of macrophages in sarcoidosis still remains unclear. As such, we aimed to elucidate other pathways involved in this pathogenic role. Gene-set enrichment and transcription factor analysis revealed that both CS and PS MΦ exhibit upregulated alternative activation yet attenuated effector function, as indicated by downregulated cytokine signaling and phagocytosis, relative to control. Similar observations have been previously reported in PS. For instance, previous studies have noted that, especially during the chronic stage of PS, BAL MΦ tend to exhibit alternative activation phenotypes (27) and attenuated phagocytic function (31). However, while it is still uncertain if such a phenotype is present in CS, there has been a report that CS granulomas tend to express Folate Receptor-β (FOLR2) which has been recently implicated in alternative activation polarization in cardiac MΦ (32, 33). Additionally, we observed downregulation of OXPHOS without sufficient glycolysis compensational upregulation only in CS MΦ. Such metabolic stress has been shown to drive macrophage alternative and activation (34) and to dysregulate their phagocytic function (35). This could explain our transcriptional observations that imply that CS MΦ might have impaired phagocytic function and be more polarized towards alternative activation. Intriguingly, only BAL MΦ exhibited upregulation of immune checkpoint signaling relative to control. This finding was well-corroborated by recent findings that T cells and NK cells upregulate the expression of immune checkpoint molecules, such as PD-1, CTLA-4 and TIGIT (36). In addition, this is consistent with several case reports that describe patients that developed immunotherapy-induced lung sarcoidosis with upregulated involvement of giant cell macrophages (37, 38). Conversely, only CS MΦ exhibited upregulated CXCL12 signaling but attenuated apoptosis relative to control. This could explain the study reporting high FOLR2 expression in CS MΦ as FOLR2 is also a marker of tissue resident MΦ (32, 33) which are known to exhibit tissue recruitment and self-renewal. Hence, we show that despite tissue-specific differences, both PS and CS MΦ exhibit attenuated alternative activation profiles relative to control.

In addition to inspecting immune cells, we investigated the transcriptomic signatures of stromal cells implicated in CS pathogenesis as well. One such cell subset is that of cardiac fibroblasts. Besides the observed increase in the fibroblast proportion, we noted a significant increase in the frequency of myofibroblasts as well as endothelial and inflammatory fibroblasts relative to control. Gene-set enrichment and transcription factor analysis revealed that all cardiac fibroblast subsets exhibited upregulation of pro-inflammatory and pro-fibrotic phenotypes relative to control. This is corroborated by studies that showed that CS patient cardiac histological samples exhibit high degrees of fibrosis as well as PET imaging studies that report increased fibroblast activation in CS (3, 39). Links between pro-inflammatory phenotypes and fibroblast activation in CS has previously been shown as it relates to monocyte-derived macrophages (40). This might be especially true since we have shown that only inflammatory fibroblasts exhibited upregulation of IFN signaling and HLA Class II signaling, a product of IFN signaling and a sign of fibroblast activation. This might be an important phenotype as IFN signaling has been implicated in the pathogenesis of PS (41) and could be implicated in CS as well. Intriguingly, we observed upregulation of apoptosis and autophagy by all cardiac fibroblast subsets. Autophagy has been shown to be downregulated by immune cells implicated in CS pathogenesis (9), such as cardiac MΦ. This finding implies that autophagy attenuation might be cell specific. Thence, we show that CS cardiac fibroblasts exhibit upregulated pro-fibrotic and pro-inflammatory profiles relative to control.

The second stromal cell population we examined was cardiac endothelial cells. In addition to the observed increase in the endothelial cell proportion, we observed a significant increase in POSTN+ and fibroblast-like endothelial cells as well as a significant decrease in cardiomyocyte-like endothelial cells relative to control. Fibrosis within blood vessel walls has been reported in CS patients, particularly in the aorta and coronary arteries (42, 43). It could be postulated that there could be a link between this histopathological feature of CS and the expansion of fibroblast-like endothelial cells. While angiogenesis has not explicitly been investigated before in CS, dysfunction in cardiac microvasculature, aortic elastic properties and coronary flow reserve have been reported in CS patients (43–45). Moreover, upregulated angiogenesis has been reported in other forms of sarcoidosis, such as neurosarcoidosis and PS (46, 47). Our analysis showed upregulation of angiogenic pathways in both our gene-set enrichment and transcription factor analysis relative to control. This was further supported by the expansion of POSTN + endothelial cells in CS relative to control which have been shown to be primary mediators of pathological angiogenesis in

various disease models (48). We also showed that cardiac endothelial cells upregulate immune response and evasion pathways in CS. Particularly, POSTN+ endothelial cells upregulated angiogenic cytokine signaling, such as IL-23 and IL-1 signaling. This is supported by studies that showed that cardiomyocytes in CS patients upregulate the expression of tissue factor pathway inhibitor (TFPI) which is implicated in IL-1 signaling as well as angiogenesis (49). In addition, there have been several case reports describing CS patients who develop vasculitis towards the end-stage of their disease (50, 51). This suggests that CS might be mediating this endothelial cell inflammation. We also detected an mTOR-dependent downregulation of autophagy in endothelial cells in CS relative to control. While autophagy defects have been long implicated in sarcoidosis pathology, the canonical thinking is in how this pathway is defected in immune cells, such as MΦ. In fact, mTOR signaling defects in cardiac MΦ have already been implicated in CS and PS pathology (9, 52). Here, we show that this defect might not be specific to immune cells. Another non-canonical observation was that all cardiac endothelial cells upregulated tissue remodeling pathways, such as TGFβ and ECM signaling, relative to control. This is certainly not a novel thought as many studies have showed that endothelial cells can have pro-fibrotic functions (53), but this suggests that histological fibrosis reported in CS might be the result of the concerted action of fibroblasts and endothelial cells. Thus, we show that cardiac endothelial cells exhibit upregulated angiogenic and pro-inflammatory phenotypes in CS relative to control.

The last cell type we inspected was cardiomyocytes. We observed that sarcoidotic cardiomyocytes were transcriptionally distinct from control cardiomyocytes. Importantly, sarcoidotic cardiomyocytes exhibited upregulated cardiac function processes, such as heart contraction and cardiac action potential, as well as pro-inflammatory and proliferative phenotypes relative to control. This was corroborated by studies that showed that CS cardiomyocytes upregulate the expression of TFPI which mediates IL-1 signaling (49). While previous reports have shown that CS is characterized by destruction of cardiomyocyte tissue (54), it is possible that the noted upregulation in proliferation and cardiac function profiles might reflect that the remaining viable cardiomyocytes promote such transcriptional pathways as a compensatory mechanism. This might be particularly true as we have shown that the proportion of cardiomyocytes are significantly reduced by as much as 7.5 folds in CS relative to control. This cardiomyocyte population contraction is corroborated by histological reports of cardiomyocyte degeneration in CS patients since 1,980 (55). The upregulation of cardiac processes by CS cardiomyocytes might provide potential mechanistic explanation of the clinical symptoms reported by CS patients, such as atrial arrhythmias, syncope, palpitations, and fatigue. Regardless, mechanistic studies investigating cardiomyocyte pathobiology in CS are lacking and these findings that CS cardiomyocytes appear to exhibit upregulated immune activation and cardiac stress profiles relative to control exemplify the need to dedicate more efforts to understanding this immunologically complex disease.

In conclusion, our transcriptomic analysis reveals that, despite tissue-specific differences, both sarcoidotic T cells and macrophages exhibit attenuated activation profiles relative to control. Intriguingly, we show that both CS and PS T cells exhibit expansion of Th17.1 populations relative to control. In addition to our findings that CS cardiac fibroblasts exhibit upregulated pro-fibrotic and pro-inflammatory phenotypes, previously reported in other forms of sarcoidosis, we report autophagy upregulation as well. Our findings also revealed that CS cardiac endothelial cells exhibit upregulated pro-angiogenic and pro-inflammatory pathways. In addition to these canonical observations, our findings revealed that CS cardiac endothelial cells exhibit upregulated tissue remodeling but downregulated autophagy phenotypes as well. Lastly, our findings revealed that CS cardiomyocytes exhibit upregulated pro-inflammatory and cardiac stress profiles. While these findings provide more insights into the intricate nature of CS pathology, BAL is a poor representative of pulmonary pathologies involving an array of immune and stromal populations due to its myeloid bias. Moreover, as the authors behind the study that made the CS snRNA-Seq dataset publicly available did not publish certain clinical details about the CS patients whose sequencing data was utilized for this analysis, such as their stage of disease progression and treatment regimen, it is challenging to ascertain how our conclusions are generalizable regardless of disease stage and treatment. As the cardiac samples utilized for the CS analysis were explanted transplant specimens, it is reasonable to assume that CS donors recruited for this dataset had reached the chronic stage of their disease by the time of sample collection. Therefore, work involving transcriptomic analysis of PS pulmonary specimens and a more representative CS cohort is needed to better profile potentially pathological transcriptional phenotypes involved in PS and CS pathology. Furthermore, as this analysis focuses on transcriptional profiling of CS and PS, it is critical to conduct protein level studies as well as biomarker and cardiac function testing using sarcoidosis CS and pulmonary sarcoidosis samples in order to better elucidate the causative processes involved in sarcoidosis pathology regardless of tissue presentation.

Data availability statement

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

Author contributions

Conceptualization, AD, DL, and DČ; formal methodology formulation and data analysis, AD; physiological contextualization and consultation, DL and TW; writing—original draft, AD and DČ; writing—reviewing and editing, AD, DL, TW, and DČ, visualization, AD and DL; supervision, DČ; fund acquisition, DČ. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the American Heart Association (AHA) 19TPA34910007 (PI: DČ), 20TPA35490421 (PI: DČ), 23EIA1040103 (PI: DČ) and 23POST1029569 (PI: DL), the National Institutes of Health (NIH)/National Heart, Lung, and Blood Institute (NHLBI) R01HL118183 and R01HL136586 (PI: DČ), the Global Autoimmune Institute, and the Matthew Poyner MVP Memorial Myocarditis Research Fund.

Acknowledgments

Supplementary Figure S1A was created using Biorender.com.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcvm.2023.1227818/full#supplementary-material

SUPPLEMENTARY FIGURE S1
Overview of the sn/scRNA-Seq workflow. Publicly available healthy & sarcoidotic sn/scRNA-Seq datasets were integrated separately for the heart & BALF comparisons. Following cell type clustering, cell type proportion, cell-cell communication and transcription factor analyses were conducted. Figure generated using Biorender.com.

SUPPLEMENTARY FIGURE S2
(A) UMAP clustering of integrated control and CS snRNA-Seq datasets by disease status. Metrics utilized for quality control are visualized on the right using violin plots that are stratified by disease status. (B) UMAP clustering of integrated control and PS scRNA-Seq datasets by disease status. Metrics utilized for quality control are visualized on the right using violin plots that are stratified by disease status. (C) Heatmaps showing the scaled expression of the top 10 markers of each identified biological cluster. (D) Pseudotime trajectory analysis conducted on cardiac stromal cells. Different parts of trajectory are highlighted and stratified by disease status. Pseudotime score indicates differentiation status.

SUPPLEMENTARY FIGURE S3
(A) Table showing number of Th17.1 cells in the heart and BAL stratified by disease status (top) and percentage of Th17.1 cells out of total T cells in the heart and BAL stratified by disease status. P-values are displayed above each table to indicate the corresponding statistical significance of Th17.1 changes in the heart across disease status. (B) Dot plot showing the cardiac gene expression of *TBX21* and *AOAH* in Th17.1 cells relative to other T cells. Average gene expression was scaled from 0 to 0.6. (C) Plot showing the ROC analysis results conducted in cardiac Th17.1 relative to other cardiac T cells. The AUC score is shown on the y-axis while the average log2(fold change) is shown on the x-axis. The closer a gene's AUC score is to 1 the stronger its potential is to be a positive classifier while the closer it is to 0.5, the weaker its potential is to be a classifier. (D) Dot plots showing the expression of sarcoidosis-specific transcription factors (TFs) in the cardiac and BAL datasets in control and sarcoidosis MΦ subpopulations. Average gene expression was scaled from 0 to 0.6.

SUPPLEMENTARY FIGURE S4
(A) Heatmap showing the scaled expression of the top 10 markers of each identified macrophage subcluster. (B) Dot plot showing the expression of markers used to identify each macrophage subcluster in the heart. Average gene expression was scaled from 0 to 2. (C) Top cell-cell communication pathways shown to be differentially upregulated in sarcoidotic BAL macrophages relative to control. The column graph shows the relative contribution of the highlighted pathways while the heatmaps show the direction of communication between the BAL macrophage clusters for each highlighted communication pathway. (D) Violin plot showing the gene expression of *NR1H3* in cardiac and BAL GPNMB+ macrophages stratified by disease status. (E) Dot plots showing the expression of sarcoidosis-specific transcription factors (TFs) in the cardiac and BAL datasets in control and sarcoidosis MΦ subpopulations. Average gene expression was scaled from 0 to 0.6.

SUPPLEMENTARY FIGURE S5
(A) Heatmap showing the scaled expression of the top 10 markers of each identified fibroblast subcluster. (B) Dot plot showing the expression of markers used to identify each fibroblast subcluster in the heart. Average gene expression was scaled from 0 to 2. (C) Violin plots showing the expression of sarcoidosis-specific transcription factors (TFs) in control and sarcoidotic cardiac fibroblast subpopulations. The classification group of each TF is shown colored above its respective gene expression plot.

SUPPLEMENTARY FIGURE S6
(A) Heatmap showing the scaled expression of the top 10 markers of each identified endothelial cell subcluster. (B) Dot plot showing the expression of markers used to identify each endothelial cell subcluster in the heart. Average gene expression was scaled from 0 to 2. (C) Cell-cell communication tree of the CDH5 communication pathway between cardiac endothelial cell subpopulations. (D) Cell-cell communication tree of the PTPRM communication pathway between cardiac endothelial cell subpopulations.

SUPPLEMENTARY FIGURE S7
(A) UMAP clustering of cardiomyocyte subclusters by disease status. (B) Heatmap showing the scaled expression of the top 10 markers of each identified cardiomyocyte subcluster. (C) Feature plot showing the expression of *TNNI3* in the different cardiomyocyte subclusters.

References

1. Sreeja C, Priyadarshini A, Premika Nachiammai N. Sarcoidosis: a review article. *J Oral Maxillofac Pathol.* (2022) 26(2):242–53. doi: 10.4103/jomfp.jomfp_373_21

2. Belperio JA, Shaikh F, Abtin FG, Fishbein MC, Weigt SS, Saggar R, et al. Diagnosis and treatment of pulmonary sarcoidosis: a review. *J Am Med Assoc.* (2022) 327(9):856–67. doi: 10.1001/jama.2022.1570

3. Mathai SV, Patel S, Jorde UP, Rochlani Y. Epidemiology, pathogenesis, and diagnosis of cardiac sarcoidosis. *Methodist Debakey Cardiovasc J.* (2022) 18 (2):78–93. doi: 10.14797/mdcvj.1057

4. Tan JL, Tan BE, Cheung JW, Ortman M, Lee JZ. Update on cardiac sarcoidosis. *Trends Cardiovasc Med.* (2022): [m5G;20:44]. doi: 10.1016/j.tcm.2022.04.007

5. Yoshioka K, Sato H, Kawasaki T, Ishii D, Imamoto T, Abe M, et al. Transcriptome analysis of peripheral blood mononuclear cells in pulmonary sarcoidosis. *Front Med*. (2022) 9:822094. doi: 10.3389/fmed.2022.822094

6. Zhang C, Tian R, Dreifus EM, Hashemi Shahraki A, Holt G, Cai R, et al. Activity of the growth hormone-releasing hormone antagonist MIA602 and its underlying mechanisms of action in sarcoidosis-like granuloma. *Clin Transl Immunology*. (2021) 10(7):e1310. doi: 10.1002/cti2.1310

7. Rosenbaum JT, Harrington CA, Searles RP, Fei SS, Zaki A, Arepalli S, et al. Identifying RNA biomarkers and molecular pathways involved in multiple subtypes of uveitis. *Am J Ophthalmol*. (2021) 226:226–34. doi: 10.1016/j.ajo.2021.01.007

8. Bhargava M, Liao SY, Crouser ED, Maier LA, Leach SM, et al. The landscape of transcriptomic and proteomic studies in sarcoidosis. *ERJ Open Res*. (2022) 8(1):00621–2021. doi: 10.1183/23120541.00621-2021

9. Liu J, Ma P, Lai L, Villanueva A, Koenig A, Bean GR, et al. Transcriptional and immune landscape of cardiac sarcoidosis. *Circ Res*. (2022) 131(8):654–69. doi: 10.1161/CIRCRESAHA.121.320449

10. Li X, Kolling FW, Aridgides D, Mellinger D, Ashare A, Jakubzick CV. ScRNA-Seq expression of IFI27 and APOC2 identifies four alveolar macrophage superclusters in healthy BALF. *Life Sci Alliance*. (2022) 5(11):e202201458. doi: 10.26508/lsa.202201458

11. Liao SY, Atif SM, Mould K, Konigsberg IR, Fu R, Davidson E, et al. Single-cell RNA sequencing identifies macrophage transcriptional heterogeneities in granulomatous diseases. *Eur Respir J*. (2021) 57(6):2003794. doi: 10.1183/13993003.03794-2020

12. Litviňuková M, Talavera-López C, Maatz H, Reichart D, Worth CL, Lindberg EL, et al. Cells of the adult human heart. *Nature*. (2020) 588(7838):466–72. doi: 10.1038/s41586-020-2797-4

13. Hao Y, Hao S, Andersen-Nissen E, Mauck WM 3rd, Zheng S, Butler A, et al. Integrated analysis of multimodal single-cell data. *Cell*. (2021) 184(13):3573–3587.e29. doi: 10.1016/j.cell.2021.04.048

14. Korsunsky I, Millard N, Fan J, Slowikowski K, Zhang F, Wei K, et al. Fast, sensitive and accurate integration of single-cell data with harmony. *Nat Methods*. (2019) 16(12):1289–96. doi: 10.1038/s41592-019-0619-0

15. Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun*. (2019) 10(1):1523. doi: 10.1038/s41467-019-09234-6

16. Jin S, Guerrero-Juarez CF, Zhang L, Chang I, Ramos R, Kuan CH, et al. Inference and analysis of cell-cell communication using CellChat. *Nat Commun*. (2021) 12(1):1088. doi: 10.1038/s41467-021-21246-9

17. Aibar S, González-Blas CB, Moerman T, Huynh-Thu VA, Imrichova H, Hulselmans G, et al. SCENIC: single-cell regulatory network inference and clustering. *Nat Methods*. (2017) 14(11):1083–6. doi: 10.1038/nmeth.4463

18. Ramstein J, Broos CE, Simpson LJ, Ansel KM, Sun SA, Ho ME, et al. IFN-γ-producing T-helper 17.1 cells are increased in sarcoidosis and are more prevalent than T-helper type 1 cells. *Am J Respir Crit Care Med*. (2016) 193(11):1281–91. doi: 10.1164/rccm.201507-1499OC

19. Boniface K, Blumenschein WM, Brovont-Porth K, McGeachy MJ, Basham B, Desai B, et al. Human Th17 cells comprise heterogeneous subsets including IFN-gamma-producing cells with distinct properties from the Th1 lineage. *J Immunol*. (2010) 185(1):679–87. doi: 10.4049/jimmunol.1000366

20. Wang Y, Godec J, Ben-Aissa K, Cui K, Zhao K, Pucsek AB, et al. The transcription factors T-bet and runx are required for the ontogeny of pathogenic interferon-γ-producing T helper 17 cells. *Immunity*. (2014) 40(3):355–66. doi: 10.1016/j.immuni.2014.01.002

21. Chaun H, King DM, Gofton JP, Sutherland WH, Bogoch A. Sarcoidosis of the pancreas. *Am J Dig Dis*. (1972) 17(8):725–30. doi: 10.1007/BF02231644

22. Calatroni M, Moroni G, Reggiani F, Ponticelli C. Renal sarcoidosis. *J Nephrol*. (2023) 36(1):5–15. doi: 10.1007/s40620-022-01369-y

23. Akasaka Y. The role of mesenchymal stromal cells in tissue repair and fibrosis. *Adv Wound Care*. (2022) 11(11):561–74. doi: 10.1089/wound.2021.0037

24. Leipold AM, Werner RA, Düll J, Jung P, John M, Stanojkovska E, et al. Th17.1 cell driven sarcoidosis-like inflammation after anti-BCMA CAR T cells in multiple myeloma. *Leukemia*. (2023) 37(3):650–8. doi: 10.1038/s41375-023-01824-0

25. Hawkins C, Shaginurova G, Shelton DA, Herazo-Maya JD, Oswald-Richter KA, Rotsinger JE, et al. Local and systemic CD4+ T cell exhaustion reverses with clinical resolution of pulmonary sarcoidosis. *J Immunol Res*. (2017) 2017:3642832. doi: 10.1155/2017/3642832

26. Terasaki F, Ukimura A, Tsukada B, Fujita S, Katashima T, Otsuka K, et al. Enhanced expression of type 1 helper T-cell cytokines in the myocardium of active cardiac sarcoidosis. *Circ J*. (2008) 72(8):1303–7. doi: 10.1253/circj.72.1303

27. Partida-Zavala N, Ponce-Gallegos MA, Buendía-Roldán I, Falfán-Valencia R. Type 2 macrophages and Th2 CD4+ cells in interstitial lung diseases (ILDs): an overview. *Sarcoidosis Vasc Diffuse Lung Dis*. (2018) 35(2):98–108. doi: 10.36141/svdld.v35i2.6691

28. Zielonka TM, Demkow U, Puscinska E, Golian-Geremek A, Filewska M, Zycinska K, et al. TNFalpha and INFgamma inducing capacity of sera from patients with interstitial lung disease in relation to its angiogenesis activity. *J Physiol Pharmacol*. (2007) 58(Suppl 5(Pt 2)):767–80. PMID: 18204191.

29. Domblides C, Lartigue L, Faustin B. Metabolic stress in the immune function of T cells, macrophages and dendritic cells. *Cells*. (2018) 7(7):68. doi: 10.3390/cells7070068

30. Hu C, Xuan Y, Zhang X, Liu Y, Yang S, Yang K. Immune cell metabolism and metabolic reprogramming. *Mol Biol Rep*. (2022) 49(10):9783–95. doi: 10.1007/s11033-022-07474-2

31. Orosi P, Nugent K. Studies of phagocytic and killing activities of alveolar macrophages in patients with sarcoidosis. *Lung*. (1993) 171(4):225–33. doi: 10.1007/BF00203722

32. Jahandideh A, Uotila S, Ståhle M, Virta J, Li XG, Kytö V, et al. Folate receptor β-targeted PET imaging of macrophages in autoimmune myocarditis. *J Nucl Med*. (2020) 61(11):1643–9. doi: 10.2967/jnumed.119.241356

33. Dick SA, Wong A, Hamidzada H, Nejat S, Nechanitzky R, Vohra S, et al. Three tissue resident macrophage subsets coexist across organs with conserved origins and life cycles. *Sci Immunol*. (2022) 7(67):eabf7777. doi: 10.1126/sciimmunol.abf7777

34. Boutens L, Stienstra R. Adipose tissue macrophages: going off track during obesity. *Diabetologia*. (2016) 59(5):879–94. doi: 10.1007/s00125-016-3904-9

35. Kolliniati O, Ieronymaki E, Vergadi E, Tsatsanis C. Metabolic regulation of macrophage activation. *J Innate Immun*. (2022) 14(1):51–68. doi: 10.1159/000516780

36. d'Alessandro M, Conticini E, Bergantini L, Mezzasalma F, Cameli P, Baglioni S, et al. PD1, CTLA4 and TIGIT expression on T and NK cells in granulomatous diseases: sarcoidosis and ANCA-associated vasculitis. *Int J Mol Sci*. (2022) 24(1):256. doi: 10.3390/ijms24010256

37. Zhao X, Yue D, Qian J, Zhang L, Song J, Zhang B, et al. Case report: sarcoid-like reactions and tertiary lymphoid structures following dual checkpoint inhibition in a patient with early-stage lung adenocarcinoma. *Front Immunol*. (2022) 13:794217. doi: 10.3389/fimmu.2022.794217

38. O'Connor C, Finnegan P, Power DG, Bennett M, Bourke JF. Pembrolizumab-associated erythema nodosum in the treatment of metastatic melanoma. *Immunotherapy*. (2022) 14(13):1021–6. doi: 10.2217/imt-2021-0239

39. Siebermair J, Kessler L, Kupusovic J, Rassaf T, Rischpler C. Cardiac fibroblast activation detected by ^{68}Gallium-FAPI-46 positron emission tomography-magnetic resonance imaging as a sign of chronic activity in cardiac sarcoidosis. *Eur Heart J Case Rep*. (2022) 6(1):ytac005. doi: 10.1093/ehjcr/ytac005

40. Patterson KC, Hogarth K, Husain AN, Sperling AI, Niewold TB. The clinical and immunologic features of pulmonary fibrosis in sarcoidosis. *Transl Res*. (2012) 160(5):321–31. doi: 10.1016/j.trsl.2012.03.005

41. Robinson BW, McLemore TL, Crystal RG. Gamma interferon is spontaneously released by alveolar macrophages and lung T lymphocytes in patients with pulmonary sarcoidosis. *J Clin Invest*. (1985) 75(5):1488–95. doi: 10.1172/JCI111852

42. Butany J, Bahl NE, Morales K, Thangaroopan M, Ross H, Rao V, et al. The intricacies of cardiac sarcoidosis: a case report involving the coronary arteries and a review of the literature. *Cardiovasc Pathol*. (2006) 15(4):222–7. doi: 10.1016/j.carpath.2006.02.005

43. Moyssakis I, Gialafos E, Tentolouris N, Floudas CS, Papaioannou TG, Kostopoulos CH, et al. Impaired aortic elastic properties in patients with systemic sarcoidosis. *Eur J Clin Invest*. (2008) 38(2):82–9. doi: 10.1111/j.1365-2362.2007.01906.x

44. Rimoldi O, Maranta F. Microvascular dysfunction in infiltrative cardiomyopathies. *J Nucl Cardiol*. (2019) 26(1):200–7. doi: 10.1007/s12350-017-0991-z

45. Kul S, Kutlu GA, Guvenc TS, Kavas M, Demircioglu K, Yilmaz Y, et al. Coronary flow reserve is reduced in sarcoidosis. *Atherosclerosis*. (2017) 264:115–21. doi: 10.1016/j.atherosclerosis.2017.05.005

46. Byg KE, Illes Z, Sejbaek T, Lambertsen KL, Ellingsen T, Nielsen HH. Inflammatory profiles in plasma and cerebrospinal fluid of patients with neurosarcoidosis. *J Neuroimmunol*. (2022) 367:577849. doi: 10.1016/j.jneuroim.2022.577849

47. Ziora D, Jastrzębski D, Adamek M, Czuba Z, Kozielski JJ, Grzanka A, et al. Circulating concentration of markers of angiogenic activity in patients with sarcoidosis and idiopathic pulmonary fibrosis. *BMC Pulm Med*. (2015) 15:113. doi: 10.1186/s12890-015-0110-3

48. Wasik A, Ratajczak-Wielgomas K, Badzinski A, Dziegiel P, Podhorska-Okolow M. The role of periostin in angiogenesis and lymphangiogenesis in tumors. *Cancers*. (2022) 14(17):4225. doi: 10.3390/cancers14174225

49. Kereveur A, Enjyoji K, Masuda K, Yutani C, Kato H. Production of tissue factor pathway inhibitor in cardiomyocytes and its upregulation by interleukin-1. *Thromb Haemost*. (2001) 86(5):1314–9. doi: 10.1055/s-0037-1616070

50. Poudel J, Risal U, Sigdel KR, Paudyal BP, Adhikari S, Basnyat B. Case report: co-existence of sarcoidosis and takayasu arteritis. *Wellcome Open Res*. (2020) 5:73. doi: 10.12688/wellcomeopenres.15837.2

51. Mahmoudzadeh R, Gopal A, Soares R, Dunn JP. Unilateral retinal arteritis and macroaneurysm in sarcoidosis. *Ocul Immunol Inflamm*. (2022) 30(7-8):1901–5. doi: 10.1080/09273948.2021.1970780

52. Linke M, Pham HT, Katholnig K, Schnöller T, Miller A, Demel F, et al. Chronic signaling via the metabolic checkpoint kinase mTORC1 induces macrophage

granuloma formation and marks sarcoidosis progression. *Nat Immunol*. (2017) 18 (3):293–302. doi: 10.1038/ni.3655

53. Biener L, Kruse J, Tuleta I, Pizarro C, Kreuter M, Birring SS, et al. Association of proangiogenic and profibrotic serum markers with lung function and quality of life in sarcoidosis. *PLoS One*. (2021) 16(2):e0247197. doi: 10.1371/journal.pone.0247197

54. Stephen JD. Fatal myocardial sarcoidosis; a case of sudden death. *Circulation*. (1954) 9(6):886–9. doi: 10.1161/01.cir.9.6.886

55. Sekiguchi M, Numao Y, Imai M, Furuie T, Mikami R. Clinical and histopathological profile of sarcoidosis of the heart and acute idiopathic myocarditis. Concepts through a study employing endomyocardial biopsy. I. Sarcoidosis. *Jpn Circ J*. (1980) 44(4):249–63. doi: 10.1253/jcj.44.249

frontiers | Frontiers in Cardiovascular Medicine

TYPE Brief Research Report
PUBLISHED 07 September 2023
DOI 10.3389/fcvm.2023.1251141

OPEN ACCESS

EDITED BY
Liming Yu,
The University of Texas Health Science Center at San Antonio, United States

REVIEWED BY
Mabruka Alfaidi,
LSU Health Sciences Center—Shreveport, United States
Qian M. A,
Augusta University, United States

*CORRESPONDENCE
Luke P. Brewster
lbrewst@emory.edu

RECEIVED 30 June 2023
ACCEPTED 15 August 2023
PUBLISHED 07 September 2023

CITATION
Sánchez Marrero G, Villa-Roel N, Li F, Park C, Kang D-W, Hekman KE, Jo H and Brewster LP (2023) Single-Cell RNA sequencing investigation of female-male differences under PAD conditions.
Front. Cardiovasc. Med. 10:1251141.
doi: 10.3389/fcvm.2023.1251141

Single-Cell RNA sequencing investigation of female-male differences under PAD conditions

Gloriani Sánchez Marrero[1], Nicolas Villa-Roel[1], Feifei Li[2], Christian Park[1], Dong-Won Kang[1], Katherine E. Hekman[2,3], Hanjoong Jo[1] and Luke P. Brewster[1,2,3]*

[1]Wallace H. Coulter Department of Biomedical Engineering, Georgia Institute of Technology and Emory University, Atlanta, GA, United States, [2]Department of Surgery, Emory University School of Medicine, Atlanta, GA, United States, [3]Surgical and Research Services, Atlanta Veteran Affairs Medical Center, Decatur, GA, United States

Peripheral arterial disease (PAD) is an age-related medical condition affecting mostly muscular arteries of the limb. It is the 3rd leading cause of atherosclerotic morbidity. The mechanical environment of endothelial cells (ECs) in PAD is characterized by disturbed blood flow (d-flow) and stiff extracellular matrices. In PAD, the stiffness of arteries is due to decreased elastin function and increased collagen content. These flow and stiffness parameters are largely missing from current models of PAD. It has been previously proven that ECs exposed to d-flow or stiff substrates lead to proatherogenic pathways, but the effect of both, d-flow and stiffness, on EC phenotype has not been fully investigated. In this study, we sought to explore the effect of sex on proatherogenic pathways that could result from exposing endothelial cells to a d-flow and stiff environment. We utilized the scRNA-seq tool to analyze the gene expression of ECs exposed to the different mechanical conditions both *in vitro* and *in vivo*. We found that male ECs exposed to different mechanical stimuli presented higher expression of genes related to fibrosis and d-flow *in vitro*. We validated our findings *in vivo* by exposing murine carotid arteries to d-flow via partial carotid artery ligation. Since women have delayed onset of arterial stiffening and subsequent PAD, this work may provide a framework for some of the pathways in which biological sex interacts with sex-based differences in PAD.

KEYWORDS

single-cell RNA sequencing, peripheral arterial disease conditions, sex-specific differences, human aortic endothelial cells, shear stress, substrate stiffness

1. Introduction

Peripheral arterial disease (PAD) manifests as obstructions in blood flow into the lower limbs. PAD most commonly affects the older populations, with an incidence of 12%–20% in those older than 65 years (1–3). The onset of arterial stiffness and endothelial cell (EC) dysfunction typically begins in the 4th decade (30–40 s), but is delayed approximately 10 years in females compared to males (4). This delay in stiffening corresponds to the subsequent delayed onset of PAD in females (5, 6). PAD is characterized by atherosclerotic lesions in regions commonly affected by disturbed flow (d-flow). D-flow is a combination of low and oscillatory shear stress (OS) at the vessel wall. Furthermore, atherosclerotic lesions have been found in locations where arteries have become stiffer, like those of aged populations (7, 8). Previous work from our group found that d-flow induces a stiffening arteriopathy closely mimicking that of aged arteries (9). The endothelial cells lining PAD arteries can sense both the d-flow of the blood and the stiffness in the vasculature's matrix (10–12). These environmental signals induce EC

activation of a proatherogenic pathway often resulting in increased EC permeability, inflammation (EndIT), endothelial-to-mesenchymal transition (EndMT), arterial stiffening, and thrombosis (11). Our research group and others have identified in ECs an important role for fibro-inflammatory pathways in PAD pathology related to thrombospondin-1 (*THBS1*) and cellular communication network factor 2 (*CCN2*), previously better known as connective tissue growth factor (9, 13).

Several comprehensive *in vitro* and *in vivo* studies have been done to understand the molecular pathways that lead to atherosclerosis (9, 14, 15). These mainly involved the transcriptomic bulk analysis of ECs. Recently, single-cell RNA sequencing (scRNA-seq) has become a popular transcriptomics tool because it allows for RNA analysis at a single-cell resolution for a large number of cells. This enables the study of EC heterogeneity as well as the ability to differentiate between cell types included in the sample. Studies using scRNA-seq methods to investigate EC heterogeneity under PAD or atherosclerotic conditions are scarce. One study reported proatherosclerotic EC reprogramming in carotid arteries of wild type mice exposed to d-flow (11), another reported EndMT progression in ECs cultured in stiff matrices (16). Since the characteristic mechanical conditions of PAD involve both pathological blood flow (d-flow) and a stiff vascular matrix, it is logical that combining these conditions is important to the identification of translationally important pathways. Unfortunately, it is not currently known how these two environmental mechanical conditions modulate EC phenotype. Furthermore, the effect of biological sex on EC response to these PAD flow and mechanical environments has not been studied with single cell analyses.

Despite the NIH mandate to include sex as a biological variable, sex-differences have not been adequately investigated in this space (17, 18). The inclusion of biological sex as a factor in cardiovascular disease research, and in PAD research, is imperative because the onset and presentation of these diseases in humans vary with biological sex. With regards to stiffening and flow, there is an *in vitro* study that found female human umbilical vein ECs (HUVECs) were less impacted than male HUVECs when exposed to higher and lower laminar shear stress and increasing substrate stiffness since they presented no change in morphology and YAP1 nuclear translocation (19). In parallel, we have been integrating substrate stiffness (softer and stiffer reflecting stiffness of healthy and diseased arteries) with low and oscillatory wall shear stress (mimicking the d-flow environment around blockages in PAD) and using laminar wall shear stress (stable flow) to reflect non-obstructed flow. Here, we incorporate scRNA-seq methodologies to enable cell specific gene expression data, which cannot be harvested using bulk RNA sequencing. The objective of this study is to rigorously investigate the impact of stiffness and flow on sex differences in the *in vitro* response of human aortic ECs (HAECs) in the context of PAD conditions (stiff substrate and d-flow). Given the powerful independent effect of stiffness or flow on ECs, we hypothesized that scRNA-seq would reveal sex-based phenotypic differences under PAD biomechanical conditions that may contribute to sex-based differences in PAD.

2. Methods

2.1. HAECs culturing and shear stress exposure

HAECs from three age-appropiate male (ages 46, 54, and 63) and two female (ages 46 and 53) donors (Lot#: 2,895, 2,824, 1,622, 1,576, 1,851; Cell Applications, San Diego, CA) were expanded, cultured, and sheared in complete endothelial cell growth medium (Cat.# 211-500; Cell Applications, San Diego, CA). HAECs of passages 4–6 were seeded onto modified 6-well plates functionalized with collagen-coated hydrogels (Cat.# SW6-COL-25 and SW6-COL-100; Matrigen, Irvine, CA) of different stiffnesses (25 kPa and 100 kPa) that mimic the EC environment in physiologic (25 kPa) and PAD conditions (100 kPa) (20, 21). The HAECs were sheared utilizing a modified orbital shaker method (11, 22, 23). In brief, hydrogel-containing 6-well plates were modified by preventing HAEC growth in the outer (> 10.5 mm) or inner (< 10.5 mm) portion of the well by placing autoclaved vinyl obstructions prior to seeding the ECs. HAECs grown in the outer ring of the well mimicked laminar shear or stable flow (LS, timed-averaged shear stress of 7 dynes/cm^2) while HAECs grown in the inner circle mimicked disturbed flow (OS, timed-averaged shear stress of 3 dynes/cm^2). The cells were sheared for 96 h utilizing an orbital shaker at 100 rpm, changing media every two days. For comparison, HAECs cultured on tissue culture plastic (TCP) immediately prior to subculture conditions above were analyzed. ECs on TCP (2 GPa) and static conditions are known to mimic many of the pathways seen under s-flow conditions, including *THBS1* (9, 24).

2.2. Single-cell collection

Following exposure to shear stress, HAECs were recovered utilizing TrypLE (Cat.# 12563011; Gibco, Thermo Fisher Scientific, Waltham, MA) and suspended in 1X PBS (Cat.# MT21040CV) with 0.04% BSA. After ensuring the proper cell concentration was collected, the samples were delivered to the Emory Integrated Genomics Core for single-cell processing and barcoding with the 10X Genomics Chromium device. NovaSeq S4 was utilized to prepare the libraries and sequence the samples. Then, the CellRanger software was utilized to demultiplex and align the sequenced data.

2.3. scRNA-seq analysis and gene ontology

scRNA-seq data was processed and analyzed using the Seurat package in R (RRID:SCR_016341). To remove low-quality cells or cell multiplets, standard quality control filters were applied (11). Cells containing more than 8,000 or less than 200 unique feature counts were removed. In addition, cells containing more than 15% mitochondrial counts were removed from the analysis as dying cells often express mitochondrial contamination. Next,

the data was normalized and scaled. The data pertaining to the static culture group was merged across donors, subjected to unsupervised clustering, and visualized with UMAPs, dot plots, and heatmaps. The data pertaining to all shearing conditions and the static culture group were merged for each HAEC donor and the data was visualized using violin plots and feature plots. Finally, we utilized gene ontology (GO) analysis to identify pathways enhanced in the different shear, stiffness, and biological sex groups of HAECs. GO analysis was performed by obtaining the top 200 differentially expressed genes across the different groups. Then, we utilized PANTHER to identify the GO biological process terms enhanced in each of the groups. We plotted the top 10 upregulated GO terms per condition, detailing the fold enrichment.

2.4. In vivo validation of EC Sex-differences: PCL surgery and qPCR

Partial carotid artery ligation (PCL) surgery was performed on 9 to 15-week-old S129 and C57BL/6 wild type (WT) mice as previously published (7, 25). The left carotid artery (LCA) was subjected to d-flow by PCL while the right carotid artery (RCA) served as *in situ* control. D-flow has been shown to induce arterial stiffening similar to that seen in aged 80-week-old animal (9). 24 h following PCL, EC-enriched RNA was collected from the carotid arteries as previously described (7, 25). 2–3 mouse arteries were pooled for each sample (n) and analyzed. RNA was isolated from the cell lysate and One-Step SYBR green (Cat.# 1725150; Bio-Rad, Hercules, CA) was used for gene analysis utilizing quantitative real-time polymerase chain reaction (qPCR, RT-PCR). The primers used were *CCN2* and *THBS1* with *18S* as the housekeeping gene (Integrated DNA Technologies, USA). The gene analysis for the *in vivo* validation was presented as the fold change normalized against *18S*.

2.5. Statistical analysis

The validation data collected from quantitative PCR was analyzed utilizing GraphPad Prism. Each data point represents 2–3 LCAs or RCAs from mouse following 24 h of PCL. The fold change gene expression of the two WT strains utilized here (S129 and C57BL/6) were pooled for plotting. A two-sample Wilcoxon-Mann–Whitney test was done to compare between biological sex. Significance was set as a p-value of <0.05.

3. Results

3.1. Sex-based differences identified in HAECs in static culture

To identify sex-specific differences in static conditions, HAECs on TCP and under static culture were pooled across donors. This condition mimics how ECs are typically expanded. After pooling all donors, eight distinct clusters were identified following unsupervised clustering and are shown using a UMAP plot (**Figure 1A**). The clusters were defined to be donor-specific, sex-specific, or common to all donors. Clusters that were donor-specific are those only present in specific HAEC donors. For example, clusters 4 and 5 are only found in male donors 3 and 2, respectively (**Figure 1E**). **Figure 1B** shows the top three differentially expressed genes per cluster, indicating HAEC heterogeneity in static culture. As shown in **Figure 1C**, many clusters were, in part, organized by the donor's biological sex. The histogram in **Figure 1D** confirms that some clusters, as presented in **Figure 1A**, are sex specific. Then, we evaluated donor heterogeneity in static culture by viewing UMAPs separated by donor (**Figure 1E**) and found no considerable differences within donors of the same sex, except those already summarized in **Figure 1A**. These UMAPs show clear transcriptional sex-specific differences in HAECs in static culture. Next, we performed a differential gene (pseudo-bulk RNA) expression analysis to identify the sex-specific differences in our scRNA-seq data of HAECs under static culture. The top 30 differentially expressed genes ranked by normalized expression in male and female HAECs are presented in the heatmap plot (**Figure 1F**). As expected, sex-linked genes like *XIST* and *RPS4Y1* rank among the highest in female and male, respectively. However, other non-sex-linked genes were identified. Specifically, in male HAECs, genes like *VWF* and *BMP4* are differentially expressed, while in female HAECs, *PIEZO2* was increased.

3.2. Sex-specific differential response of HAECs exposed to shear stress and substrate stiffness

Motivated by the sex-specific differences shown in the static culture group, we next investigated the existence of sex-specific differences when HAECs were exposed to different magnitudes of shear stress and substrate stiffness. The relative gene expression across all shear stress and substrate conditions per donor was plotted in violin plots (**Figure 2A**). Of interest, we can note that the gene expression level of HAECs in response to mechanical stimuli was stronger in males than in females. This is particularly notable when evaluating profibrotic genes like *CCN1*, *CCN2*, *COL8A1* and d-flow associated genes like *THBS1* that present higher expression in male HAECs after mechanical stimuli.

3.3. In vivo validation

We validated two important fibro-inflammatory markers from our *in vitro* scRNA-seq findings with two commonly used WT murine strains (S129 and C57BL/6). Here, EC-enriched RNA was collected after 24 h of d-flow, induced by PCL in the LCA; the contralateral RCA serving as an *in situ* control (7, 25). *CCN2* and *THBS1* expression levels were quantified by

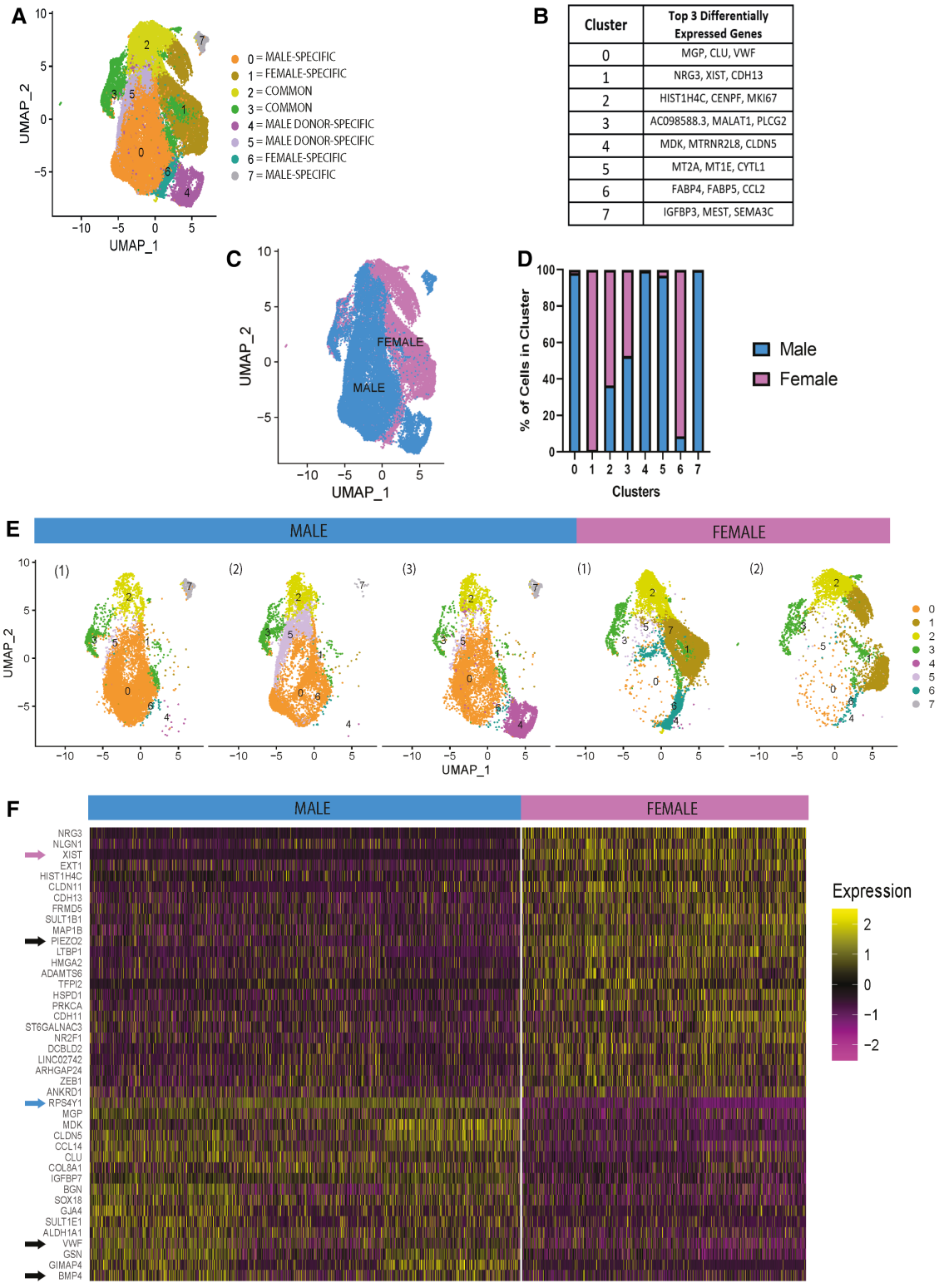


FIGURE 1
In static conditions, gene expression of HAECs is different across biological sex. (**A**–**F**) scRNA-seq data for HAECs cultured in static conditions. (**A**) UMAP presenting the 8 distinct clusters identified. (**B**) Table showing top three differentially expressed genes per cluster. (**C**) UMAP identifying female and male HAECs. (**D**) Histogram presenting the percentage of HAECS that are male or female per cluster. (**E**) UMAPs separated by donor, $n = 3$ male, $n = 2$ female. (**F**) Following differential gene expression analysis, a pseudo-bulk heat map was generated showing the top 30 differentially expressed genes for the male and female HAECs group.

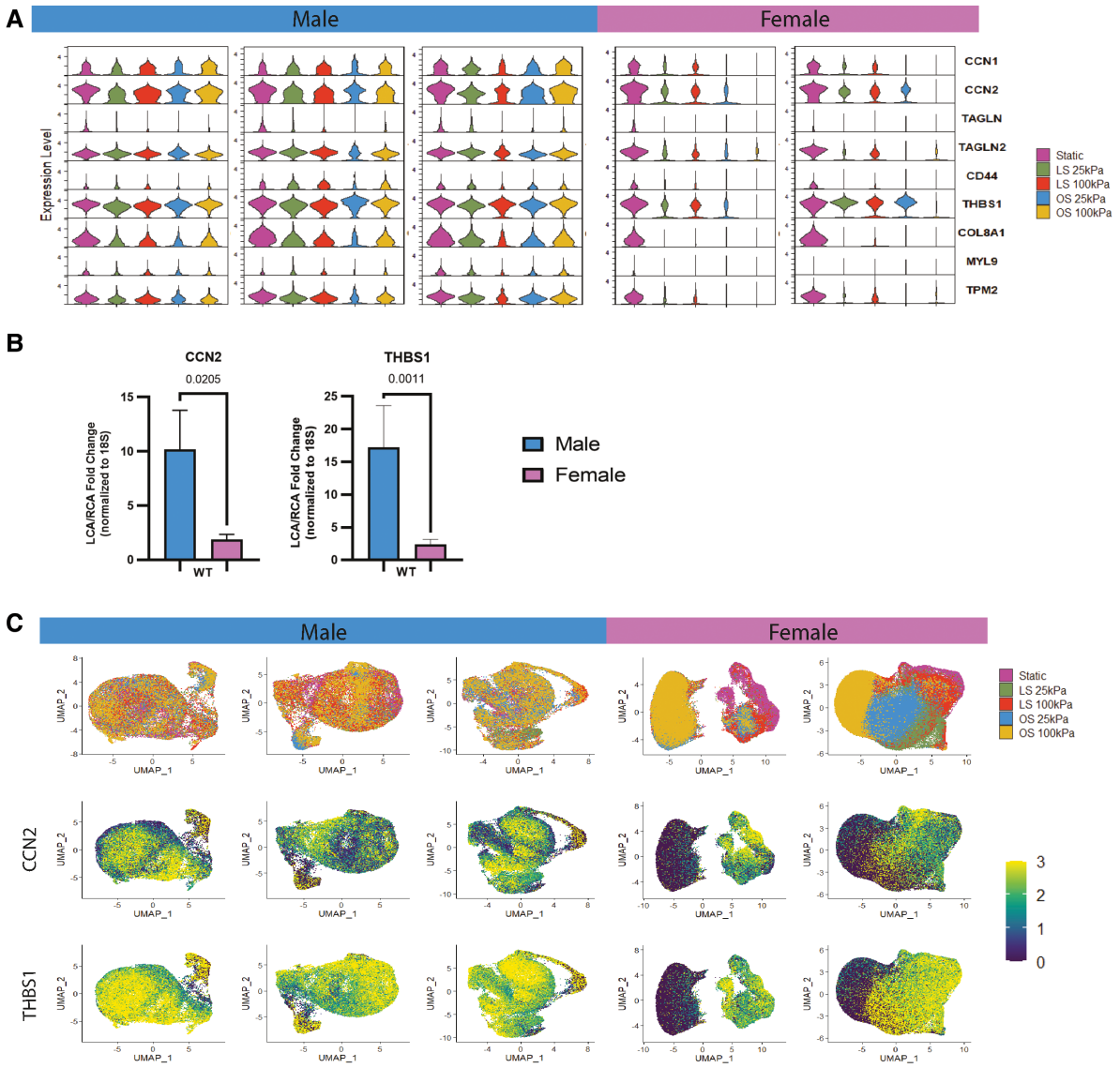


FIGURE 2
Male and female HAECs respond differently to shear and substrate stiffness. (A) Violin plots from scRNA-seq data showing the different gene expressions of female and male HAECs under laminar shear (LS), oscillatory shear (OS) or static culture and seeded on soft (25 kPa) and stiff (100 kPa) substrates. (B) *In vivo* validation of male and female scRNA-seq analysis showing fold change differences in *CCN2* and *THBS1* in EC-enriched RNA of the d-flow left carotid artery (LCA) normalized to the s-flow right carotid artery (RCA) of mice after 24 h. D-flow in LCA was imposed by partial carotid ligation. The data is presented as the mean ± SEM (n = 3-5). (C) The top row shows donor-specific UMAP plots merging all shear and substrate stiffness conditions. The bottom two rows are feature plots from scRNA-seq in each donor detailing *CCN2* and *THBS1* expression level merging all shear and substrate stiffness conditions.

qPCR (Figure 2B). Fold change gene expressions of both WT strains were merged. *CCN2* and *THBS1* expression was significantly higher in males than females. These trends were consistent with our scRNA-seq data (Figure 2C) which show higher gene expression levels in male HAECs when compared to female HAECs for *THBS1* and *CCN2*. The UMAP and feature plots shown in Figure 2C allow for the evaluation of which HAECs across the different shear and substrate conditions are expressing more *CCN2* and *THBS1*. Moreover, the degree of expression of these genes is consistent across scRNA-seq *in vitro* and bulk *in vivo* data, with *THBS1* being abundantly expressed followed by *CCN2*.

3.4. GO analysis

To test relevant signaling pathways, we evaluated enhanced GO terms in each condition across biological sex (Figure 3). We found that male HAECs exposed to a mechanical environment mimicking PAD exhibited many of the EndMT pathways associated with vasculature development while female HAECs showed processes related to cellular signaling and metabolism. Moreover, when we evaluated the GO terms shared between the static, OS 25 kPa, and OS 100 kPa groups for each biological sex (delineated with * in Figure 3), we found a large overlap of GO terms in males and less so in female HAECs. We also

investigated the GO terms unique to the LS 25 kPa and LS 100 kPa groups for each biological sex (marked with ^ in **Figure 3**) and discovered more unique LS GO terms in females compared to males.

4. Discussion

This study leveraged scRNA-seq analysis to investigate phenotypic differences by cell-specific gene expression between male and female HAECs under clinically relevant mechanical environments (softer/stiffer substrates; laminar/oscillatory shear stress). After finding distinct differences between male and female HAECs on TCP under static culture conditions, we considered the effect of mechanical stimuli on sex-specific differences. We found that male HAECs presented higher expression of genes related to fibrosis and d-flow. Lastly, we validated our scRNA-seq results with *in vivo* results from a murine PCL model of d-flow. The EC-enriched RNA taken from these arteries showed gene expressions consistent with those found in the *in vitro* model that was analyzed with scRNA-seq.

This concise study opens the door to several avenues of investigation. While our findings clearly showed male-female differences in gene expression for HAECs exposed to mechanical stimuli, gene expression differences across LS and OS groups are not as clear. Particularly, common EC d-flow markers such as *CCN2*, *TAGLN*, and *THBS1* (9, 11) that at this time point, did not show marked upregulation by OS. Since these d-flow genes have not been previously studied in HAECs cultured on substrates of different stiffnesses, future work will address temporal changes that may be important to this process. This is easily incorporated into this model system that innovatively incorporates stiffness/flow mimicking PAD conditions *in vitro*, and *in vivo* models can match such time points for translational testing. Understanding how the various mechanical stimuli differentially affect molecular pathways within ECs is of interest to our group, particularly how a PAD environment that includes a stiff ECM and d-flow leads to atherosclerosis onset and progression. In this regard, our *in vitro* scRNA-seq and *in vivo* data shows a clear sex-based difference in *THBS1* and *CCN2*, which are currently being investigated as part of fibroinflammatory pathways in PAD (9). In addition, further investigation of our pathway enrichment analysis is ongoing, and we expect to be able to discover gene groups and molecular pathways could play important roles in PAD.

This study included 5 HAEC donors (3 male and 2 female), which is similar in donor number to many publications (26–28). All donors were in the age range for PAD (46–63 years). Nonetheless, to investigate if genetic factors in donors significantly affect our conclusions more HAEC donors would be needed. To date, we are not sure of the importance of donor age on the findings presented here, and commercially available donors are quite limited. This is particularly so for aged donors. As such, we plan to use our access to patients and patient tissue in future studies that utilize scRNA-seq on PAD patient arteries under d-flow and s-flow conditions. Arteries will be further segregated by stiffness using mechanical testing as published (9, 29–31). This will permit us to validate key findings and

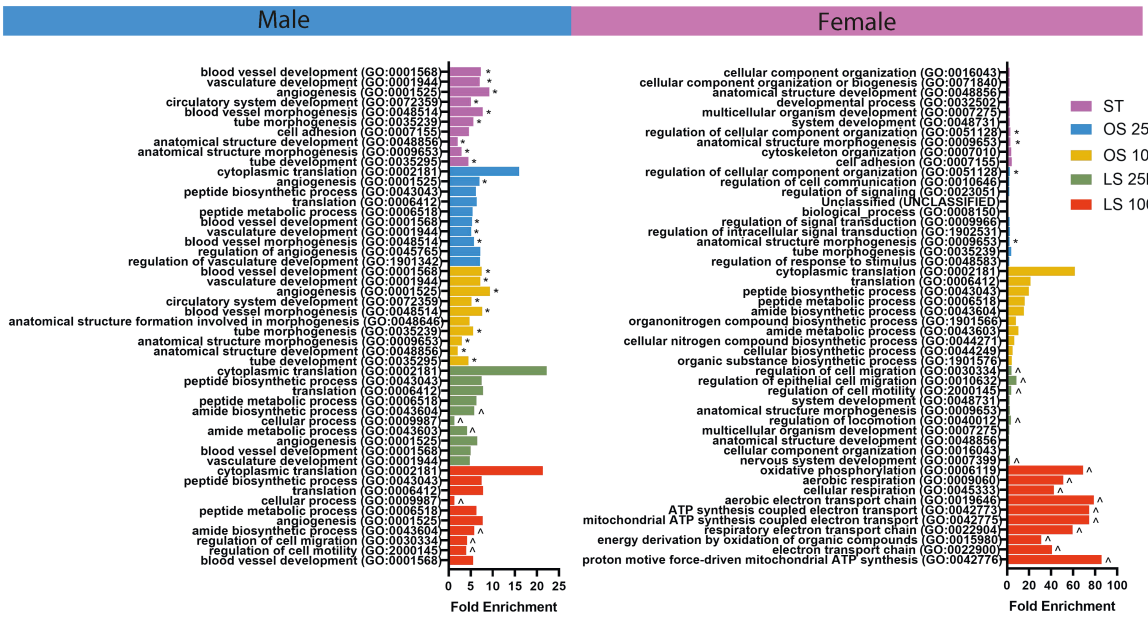


FIGURE 3
Male HAECs share more GO terms across static and OS conditions while female HAECs show more unique GO terms in LS groups. GO analysis results showing the top 10 most significant GO biological process terms across stiffness, shear, and biological sex conditions. GO was performed by obtaining the top 200 differentially expressed genes across stiffness, shear, and biological sex conditions. Asterisks (*) indicate the GO terms that are shared between the static culture, OS 25 kPa, and OS 100 kPa groups within biological sex. The circumflex symbols (^) denote the GO terms that were unique to the LS 25 kPa and LS 100 kPa groups within biological sex.

discover new pathways that may lead to translational therapies to better treat both male and female PAD patients.

Others have previously utilized scRNA-seq analysis tools to examine the effects of the mechanical environment of ECs on the transcriptome at a single-cell resolution as it relates to atherosclerosis, but biological sex was not taken into consideration (11, 16). A recent study by Zamani et al. used scRNA-seq analysis to show that ECs cultured on TCP presented high heterogeneity and higher mesenchymal transcriptional features suggesting EndMT progression (16). Similarly, our group of HAECs exposed to no flow and cultured on tissue culture plastic presented high heterogeneity (**Figure 1**). Moreover, our static culture group showed high gene expression of smooth muscle cell and mesenchymal cell markers like *TAGLN*, *TAGLN2*, *MYL9*, and *TPM2* in male and female HAECs (**Figure 2A**). This might be due to the high stiffness of the tissue culture plastic which, as proposed by Zamani et al., could lead to EndMT progression. It is well known that static culture on TCP has a genetic signature very similar to that of disturbed flow and in contrast to that of stable flow (9, 32). The idea of arterial stiffness being involved in atherosclerosis progression has been studied previously and with our experimental design we can begin investigating how biological sex could play a role in this (16, 33, 34). The work we have done here also takes into consideration both the pathological mechanical and hemodynamic microenvironment provided by a stiff ECM and disturbed flow conditions. This ensured that the pathological conditions presented in atherosclerotic lesions *in vivo* are more accurately represented in the *in vitro* model utilized. Furthermore, our study centers on the effect of biological sex on the HAEC response to shear stress and substrate stiffness, which appears to be a key component of EC response to flow and stiffness that has been largely overlooked in prior publications.

Biological sex is underreported in *in vitro* and *in vivo* research studies in the cardiovascular field (17, 18). In the case of PAD, the inclusion of biological sex as a study variable might help explain the delayed onset of PAD in females. Recent work by James and Allen shows the imperative need of including biological sex as an experimental variable in *in vitro* studies. After applying controlled shear stress (15 dynes/cm^2 or 5 dynes/cm^2) to HUVECs seeded on substrates of different stiffnesses (10 kPa or 100 kPa), they discovered that female HUVECs remained mostly invariant to the different combinations of mechanical stimuli suggesting that females might be more resistant to changes in the mechanical environment of ECs (19). Here, we exposed female HAECs to a complex mechanical environment mimicking PAD and found an attenuated response in comparison to males. This is consistent with the delayed onset of PAD in females. In contrast, transforming growth factor beta (*TGF-β*) has been proposed to be important to atherosclerotic remodeling, but the data presented was all from male animals (35). Since *THBS1* can activate *TGF-β*, this mechanism may be relatively more important in males compared to females. Biological sex-based differences are present in cardiovascular diseases such as pulmonary arterial hypertension, aortic valve stenosis, PAD, and more (6, 26, 36). Studies investigating sexual dimorphisms in cardiovascular disease have queried genes encoded in sex chromosomes since recent findings suggested that non-hormonal differences might affect the differential disease presentation observed (37, 38). For gene dosage balance, female cells undergo X chromosome inactivation (XCI). Failure to do so could lead to the unbalanced transcription of X-chromosome genes (27). The effect this could have on cardiovascular disease is being investigated, but no studies have examined XCI in PAD. This is a potential and exciting future direction for our work. These intrinsic biological differences warrant the specification of sex in studies utilizing cells. Furthermore, if sex-specific differences are heightened by the experimental conditions, in our case, mechanical stimulation, future research must mandate better inclusion of biological sex as a variable in experimental design.

This work highlights significant sex-based EC phenotypic differences induced by various clinically relevant culture conditions that may have profound impact on our understanding of sex-based clinical differences. In particular, this study found that male HAECs on TCP and cultured under static conditions and HAECs exposed to OS conditions had relatively similar gene expression in markers of EndMT and mesenchymal cells. In contrast, female HAECs showed higher EndMT and mesenchymal cell markers in HAECs cultured on static conditions when compared to other shear and substrate stiffness conditions. The *in vitro* model utilized in this work also enables rigorous delineation of sex-based EC differences in a setting that can exclude hormonal interactions of *in vivo* studies as well as comparisons between humans and mice that have distinct hormonal patterns across all ages. Still, this *in vitro* platform could be used to incorporate exogenous hormone levels to test *in vivo* conditions. Further, our *in vivo* validation supports acute sex-based EC expression differences under PAD flow conditions.

In conclusion, sex-based differences are seen in HAECs under PAD conditions. These EC phenotypes may contribute to sex-based clinical differences in PAD. The use of scRNA-seq is a powerful technique that will allow better understanding of ECs under complex biomechanical conditions. Future work may further aid in the discovery of EC heterogeneity, genetic differences of PAD patients, as well as sex-based donor differences under static culture and under clinically relevant conditions.

Data availability statement

The data presented in this study are deposited in the NCBI BioProject repository with accession number: PRJNA1008616. The data is available at: https://www.ncbi.nlm.nih.gov/bioproject/PRJNA100861.

Ethics statement

Ethical approval was not required for the studies on humans in accordance with the local legislation and institutional requirements

because only commercially available established cell lines were used. The animal study was approved by Emory University Institutional Animal Care and Use Committee. The study was conducted in accordance with the local legislation and institutional requirements.

Author contributions

GSM, NV-R, FL, HJ, and LB contributed to the conception and design of the study. FL, CP, and D-WK performed experiments. GSM and NV-R performed scRNA-seq analysis. GSM wrote the manuscript. GSM, NV-R, KH, and LB critically reviewed and edited the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This work was fully funded by the National Institutes of Health, grant HL143348.

Acknowledgments

This study was supported in part by the Emory Integrated Genomics Core (EIGC), which is subsidized by the Emory University School of Medicine and is one of the Emory Integrated Core Facilities.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

1. Belch JJ, Topol EJ, Agnelli G, Bertrand M, Califf RM, Clement DL, et al. Critical issues in peripheral arterial disease detection and management: a call to action. *Arch Intern Med*. (2003) 163(8):884–92. doi: 10.1001/archinte.163.8.884

2. Ware JE Jr. The status of health assessment 1994. *Annu Rev Public Health*. (1995) 16(1):327–54. doi: 10.1146/annurev.pu.16.050195.001551

3. Criqui MH, Aboyans V. Epidemiology of peripheral artery disease. *Circ Res*. (2015) 116(9):1509–26. doi: 10.1161/CIRCRESAHA.116.303849

4. Reneman RS, Van Merode T, Hick P, Muytjens AM, Hoeks AP. Age-related changes in carotid artery wall properties in men. *Ultrasound Med Biol*. (1986) 12 (6):465–71. doi: 10.1016/0301-5629(86)90218-8

5. DuPont JJ, Kenney RM, Patel AR, Jaffe IZ. Sex differences in mechanisms of arterial stiffness. *Br J Pharmacol*. (2019) 176(21):4208–25. doi: 10.1111/bph.14624

6. Pabon M, Cheng S, Altin SE, Sethi SS, Nelson MD, Moreau KL, et al. Sex differences in peripheral artery disease. *Circ Res*. (2022) 130(4):496–511. doi: 10.1161/CIRCRESAHA.121.320702

7. Nam D, Ni C-W, Rezvan A, Suo J, Budzyn K, Llanos A, et al. Partial carotid ligation is a model of acutely induced disturbed flow, leading to rapid endothelial dysfunction and atherosclerosis. *Am J Physiol-Heart Circ Physiol*. (2009) 297(4): H1535–43. doi: 10.1152/ajpheart.00510.2009

8. Labropoulos N, Leon L Jr, Brewster L, Pryor L, Tiongson J, Kang S, et al. Are your arteries older than your age? *Eur J Vasc Endovasc Surg*. (2005) 30(6):588–96. doi: 10.1016/j.ejvs.2005.06.011

9. Kim CW, Pokutta-Paskaleva A, Kumar S, Timmins LH, Morris AD, Kang D-W, et al. Disturbed flow promotes arterial stiffening through thrombospondin-1. *Circulation*. (2017) 136(13):1217–32. doi: 10.1161/CIRCULATIONAHA.116.026361

10. Wood JA, Liliensiek SJ, Russell P, Nealey PF, Murphy CJ. Biophysical cueing and vascular endothelial cell behavior. *Materials (Basel)*. (2010) 3(3):1620–39. doi: 10.3390/ma3031620

11. Andueza A, Kumar S, Kim J, Kang D-W, Mumme HL, Perez JI, et al. Endothelial reprogramming by disturbed flow revealed by single-cell RNA and chromatin accessibility study. *Cell Rep*. (2020) 33(11):108491. doi: 10.1016/j.celrep.2020.108491

12. Coutinho T, Rooke TW, Kullo IJ. Arterial dysfunction and functional performance in patients with peripheral artery disease: a review. *Vasc Med*. (2011) 16(3):203–11. doi: 10.1177/1358863X11400935

13. Smadja DM, d'Audigier C, Bieche I, Evrard S, Mauge L, Dias J-V, et al. Thrombospondin-1 is a plasmatic marker of peripheral arterial disease that modulates endothelial progenitor cell angiogenic properties. *Arterioscler Thromb Vasc Biol*. (2011) 31(3):551–9. doi: 10.1161/ATVBAHA.110.220624

14. Jiang Y-Z, Jiménez JM, Ou K, McCormick ME, Zhang L-D, Davies PF. Hemodynamic disturbed flow induces differential DNA methylation of endothelial kruppel-like factor 4 promoter in vitro and in vivo. *Circ Res*. (2014) 115(1):32–43. doi: 10.1161/CIRCRESAHA.115.303883

15. Dunn J, Qiu H, Kim S, Jjingo D, Hoffman R, Kim CW, et al. Flow-dependent epigenetic DNA methylation regulates endothelial gene expression and atherosclerosis. *J Clin Invest*. (2014) 124(7):3187–99. doi: 10.1172/JCI74792

16. Zamani M, Cheng YH, Charbonier F, Gupta VK, Mayer AT, Trevino AE, et al. Single-cell transcriptomic census of endothelial changes induced by matrix stiffness and the association with atherosclerosis. *Adv Funct Mater*. (2022) 32(47):2203069. doi: 10.1002/adfm.202203069

17. Taylor KE, Vallejo-Giraldo C, Schaible NS, Zakeri R, Miller VM. Reporting of sex as a variable in cardiovascular studies using cultured cells. *Biol Sex Differ*. (2011) 2:1–7. doi: 10.1186/2042-6410-2-11

18. Clayton JA, Gaugh MD. Sex as a biological variable in cardiovascular diseases: JACC focus seminar 1/7. *J Am Coll Cardiol*. (2022) 79(14):1388–97. doi: 10.1016/j.jacc.2021.10.050

19. James BD, Allen JB. Sex-specific response to combinations of shear stress and substrate stiffness by endothelial cells in vitro. *Adv Healthcare Mater*. (2021) 10 (18):2100735. doi: 10.1002/adhm.202100735

20. Chai C-K, Speelman L, Oomens CW, Baaijens FP. Compressive mechanical properties of atherosclerotic plaques—indentation test to characterise the local anisotropic behaviour. *J Biomech*. (2014) 47(4):784–92. doi: 10.1016/j.jbiomech.2014.01.018

21. Ebenstein DM, Coughlin D, Chapman J, Li C, Pruitt LA. Nanomechanical properties of calcification, fibrous tissue, and hematoma from atherosclerotic plaques. *J Biomed Mater Res Part A*. (2009) 91(4):1028–37. doi: 10.1002/jbm.a.32321

22. Ghim M, Pang KT, Arshad M, Wang X, Weinberg PD. A novel method for segmenting growth of cells in sheared endothelial culture reveals the secretion of an anti-inflammatory mediator. *J Biol Eng*. (2018) 12(1):1–13. doi: 10.1186/s13036-018-0107-6

23. Pang KT, Ghim M, Arshad M, Wang X, Weinberg PD. Segmenting growth of endothelial cells in 6-well plates on an orbital shaker for mechanobiological studies. *J Vis Exp*. (2021) 172. doi: 10.3791/61817

24. Simmons RD, Kumar S, Jo H. The role of endothelial mechanosensitive genes in atherosclerosis and omics approaches. *Arch Biochem Biophys*. (2016) 591:111–31. doi: 10.1016/j.abb.2015.11.005

25. Nam D, Ni C-W, Rezvan A, Suo J, Budzyn K, Llanos A, et al. A model of disturbed flow-induced atherosclerosis in mouse carotid artery by partial ligation and a simple method of RNA isolation from carotid endothelium. *J Vis Exp*. (2010) 40:e1861. doi: 10.3791/1861

26. Aguado BA, Walker CJ, Grim JC, Schroeder ME, Batan D, Vogt BJ, et al. Genes that escape X chromosome inactivation modulate sex differences in valve myofibroblasts. *Circulation*. (2022) 145(7):513–30. doi: 10.1161/CIRCULATIONAHA.121.054108

27. Tukiainen T, Villani A-C, Yen A, Rivas MA, Marshall JL, Satija R, et al. Landscape of X chromosome inactivation across human tissues. *Nature*. (2017) 550 (7675):244–8. doi: 10.1038/nature24265

28. Wirka RC, Wagh D, Paik DT, Pjanic M, Nguyen T, Miller CL, et al. Atheroprotective roles of smooth muscle cell phenotypic modulation and the TCF21 disease gene as revealed by single-cell analysis. *Nat Med*. (2019) 25 (8):1280–9. doi: 10.1038/s41591-019-0512-5

29. Dong H, Ferruzzi J, Liu M, Brewster LP, Leshnower BG, Gleason RL Jr. Effect of Aging, Sex, and Gene (Fbln5) on Arterial Stiffness of Mice: 20 Weeks Adult Fbln5-knockout Mice Have Older Arteries than 100 Weeks Wild-Type Mice. bioRxiv. (2023):2023.05. 30.542920.

30. Wang R, Raykin J, Li H, Gleason RL Jr, Brewster LP. Differential mechanical response and microstructural organization between non-human primate femoral and carotid arteries. *Biomech Model Mechanobiol*. (2014) 13(5):1041–51. doi: 10.1007/s10237-014-0553-0

31. Wang R, Raykin J, Brewster LP, Gleason RL Jr. A novel approach to assess the in situ versus ex vivo mechanical behaviors of the coronary artery. *J Biomech Eng*. (2017) 139(1):011010. doi: 10.1115/1.4035262

32. Qiao C, Meng F, Jang I, Jo H, Chen YE, Zhang J. Deep transcriptomic profiling reveals the similarity between endothelial cells cultured under static and oscillatory shear stress conditions. *Physiol Genomics*. (2016) 48(9):660–6. doi: 10.1152/physiolgenomics.00025.2016

33. Palombo C, Kozakova M. Arterial stiffness, atherosclerosis and cardiovascular risk: pathophysiologic mechanisms and emerging clinical indications. *Vasc Pharmacol*. (2016) 77:1–7. doi: 10.1016/j.vph.2015.11.083

34. Hansen L, Taylor WR. Is increased arterial stiffness a cause or consequence of atherosclerosis? *Atherosclerosis*. (2016) 249:226–7. doi: 10.1016/j.atherosclerosis.2016.04.014

35. Chen P-Y, Qin L, Li G, Wang Z, Dahlman JE, Malagon-Lopez J, et al. Endothelial TGF-β signalling drives vascular inflammation and atherosclerosis. *Nat Metabol*. (2019) 1(9):912–26. doi: 10.1038/s42255-019-0102-3

36. Predescu DN, Mokhlesi B, Predescu SA. The impact of sex chromosomes in the sexual dimorphism of pulmonary arterial hypertension. *Am J Pathol*. (2022) 192 (4):582–94. doi: 10.1016/j.ajpath.2022.01.005

37. Li J, Chen X, McClusky R, Ruiz-Sundstrom M, Itoh Y, Umar S, et al. The number of X chromosomes influences protection from cardiac ischaemia/reperfusion injury in mice: one X is better than two. *Cardiovasc Res*. (2014) 102(3):375–84. doi: 10.1093/cvr/cvu064

38. Ji H, Zheng W, Wu X, Liu J, Ecelbarger CM, Watkins R, et al. Sex chromosome effects unmasked in angiotensin II–induced hypertension. *Hypertension*. (2010) 55 (5):1275–82. doi: 10.1161/HYPERTENSIONAHA.109.144949

frontiers | Frontiers in Cell and Developmental Biology

TYPE Original Research
PUBLISHED 24 October 2023
DOI 10.3389/fcell.2023.1258959

OPEN ACCESS

EDITED BY
Hanjoong Jo,
Emory University, United States

REVIEWED BY
Triet M. Bui,
Dana–Farber Cancer Institute, United States
Feng Yuan,
The University of Iowa, United States

*CORRESPONDENCE
Long Zeng,
lozengo@163.com

†These authors share first authorship

RECEIVED 14 July 2023
ACCEPTED 09 October 2023
PUBLISHED 24 October 2023

CITATION
Li X, Xu C, Li Q, Shen Q and Zeng L (2023), Exploring key genes associated with neutrophil function and neutrophil extracellular traps in heart failure: a comprehensive analysis of single-cell and bulk sequencing data.
Front. Cell Dev. Biol. 11:1258959.
doi: 10.3389/fcell.2023.1258959

Exploring key genes associated with neutrophil function and neutrophil extracellular traps in heart failure: a comprehensive analysis of single-cell and bulk sequencing data

Xudong Li[1†], Changhao Xu[2†], Qiaoqiao Li[3†], Qingxiang Shen[4] and Long Zeng[5*]

[1]State Key Laboratory of Organ Failure Research, Department of Cardiology, Nanfang Hospital, Southern Medical University, Guangzhou, China, [2]Department of Cardiology, The First Affiliated Hospital of Nanjing Medical University, Nanjing, China, [3]Department of Cardiology, The Second Affiliated Hospital of Chongqing Medical University, Chongqing, China, [4]Department of Obstetrics and Gynecology, The Second Affiliated Hospital of the University of South China, University of South China, Hengyang, Hunan, China, [5]Department of Cardiology, Shangrao People's Hospital, Shangrao, Jiangxi, China

Background: Heart failure (HF) is a complex and heterogeneous manifestation of multiple cardiovascular diseases that usually occurs in the advanced stages of disease progression. The role of neutrophil extracellular traps (NETs) in the pathogenesis of HF remains to be explored.

Methods: Bioinformatics analysis was employed to investigate general and single-cell transcriptome sequencing data downloaded from the GEO datasets. Differentially expressed genes (DEGs) associated with NETs in HF patients and healthy controls were identified using transcriptome sequencing datasets and were subsequently subjected to functional enrichment analysis. To identify potential diagnostic biomarkers, the random forest algorithm (RF) and the least absolute shrinkage and selection operator (LASSO) were applied, followed by the construction of receiver operating characteristic (ROC) curves to assess accuracy. Additionally, single-cell transcriptome sequencing data analysis identified key immune cell subpopulations in TAC (transverse aortic constriction) mice potentially involved in NETs regulation. Cell-cell communication analysis and trajectory analysis was then performed on these key cell subpopulations.

Results: We identified thirteen differentially expressed genes (DEGs) associated with NET through differential analysis of transcriptome sequencing data from HF (heart failure) samples. Utilizing the Random Forest and Lasso algorithms, along with experimental validation, we successfully pinpointed four diagnostic markers (CXCR2, FCGR3B, VNN3, and FPR2) capable of predicting HF risk. Furthermore, our analysis of intercellular communication, leveraging single-cell sequencing data, highlighted macrophages and T cells as the immune cell subpopulations with the closest interactions with neutrophils. Pseudo-trajectory analysis sheds light on the differentiation states of distinct neutrophil subpopulations.

Conclusion: In this study, we conducted an in-depth investigation into the functions of neutrophil subpopulations that infiltrate cardiac tissue in TAC

mice. Additionally, we identified four biomarkers (CXCR2, FCGR3B, VNN3, and FPR2) associated with NETs in HF. Our findings enhance the understanding of immunology in HF.

KEYWORDS

bulk RNA sequencing, heart failure, neutrophil, single-cell RNA sequencing, neutrophil extracellular traps (NET)

Introduction

As the immune system's first line of defense against infections, neutrophils play a critical role. They regulate the immune response through three mechanisms: phagocytosis, degranulation, and the release of NETs (Papayannopoulos, 2018). NETs are a significant component of the innate immune response. They entrap and eliminate pathogenic microorganisms, including viruses, bacteria, fungi, and protozoa, to prevent their wider spread *in vivo* (Carmona-Rivera et al., 2019). NETs are assembled from cytolytic and granular proteins, which are arranged on a dense chromatin scaffold. The formation of these structures results in the release of high concentrations of toxic proteins, which are lethal to entrapped microorganisms. The primary mode of NET release from neutrophils is through a cell death process known as NETosis. The multifaceted function of NETs in the immune system highlights their importance as an effective strategy against infectious diseases.

Recent research has shown that NETs play a crucial role in human immune responses, and their involvement in pathologies such as systemic lupus erythematosus (Kraaij et al., 2018; Papayannopoulos, 2018; Frangou et al., 2019a; Frangou et al., 2019b), rheumatoid arthritis (Khandpur et al., 2013; O'Neil et al., 2023; Sakkas et al., 2014; de Bont et al., 2020), and cystic fibrosis (Skopelja et al., 2016; Gray et al., 2018; Guerra et al., 2020; Morán et al., 2022) has been extensively studied. However, a growing body of literature also highlights their contribution to HF. Studies have found that cardiac pressure overload triggers NETosis, which can lead to a decrease in left ventricular ejection fraction (LVEF) in wild-type (WT) mice (Martinod et al., 2017). In Seipin/Bslc2 knockout mice, an Asian lean diabetic model, the formation of interstitial fibrosis associated with NETs exacerbates left ventricular sclerosis and further contributes to HF during its progression wang (Wang et al., 2019). While current research predominantly targets cardiomyocytes, fibroblasts, and other immune cell subsets as therapeutic targets, the potential contribution of non-resident immune cell subpopulations, such as neutrophils in cardiac tissue, to the development of HF remains poorly understood. Therefore, further investigation and development of the role of neutrophils in HF progression are necessary to better understand their potential therapeutic value.

The rapid development of Bulk RNA sequencing technology and single-cell sequencing technology has facilitated the discovery of new diagnostic and prognostic markers for diseases. In this study, we downloaded HF-related datasets from the GEO database and employed bioinformatics to screen for NET-related diagnostic markers in HF. Our analysis of intercellular communication, based on single-cell sequencing data, revealed that macrophages and T cells are the immune cell subpopulations with the most prominent interactions with neutrophils. Additionally, through pseudo-trajectory analysis, we gained insights into the differentiation status of various neutrophil clusters. These findings provide new insights into the role of NETs in HF and have significant implications for the development of targeted treatments and prevention strategies.

Materials and methods

Dataset and preprocessing

The RNA-seq datasets analyzed in this study were retrieved from the GEO database (https://www.ncbi.nlm.nih.gov/geo/) and consisted of 131 samples from 2 separate datasets. Of these, the GSE145154 dataset comprised 67 samples (52 HF cases and 15 healthy controls), while the GSE116250 dataset had 64 samples (50 HF cases and 14 healthy controls). The GSE145154 dataset was utilized as the training dataset, while the GSE116250 dataset was used as the validation dataset. It is worth noting that the 69 initial biomarkers of NETs included in this study were obtained from prior research studies (Zhang et al., 2022) (Supplementary Table S1).

Identification of NETs-related differential genes

To identify differentially expressed genes (DEGs) associated with NETs, we conducted differential analysis on the training dataset samples utilizing the "Deseq2" package (version 1.38.2) (Anders and Huber, 2010). The established thresholds were set at p-value <0.05 and |logFC|>1. After intersecting DEGs with NETs genes, a final set of 13 NETs-related differential genes were obtained between HF and control samples.

Identification of NETs-related diagnostic biomarkers

To pinpoint key NETs-related biomarkers, we harnessed the Random Forest (RF) algorithm and the LASSO regression model.

LASSO is a widely-used regression method for selecting variables to improve prediction accuracy, implemented through the "glmnet" R package (version 4.1), we selected the optimal λ value and removed genes that displayed partial collinearity to reduce potential bias. In contrast, the RF algorithm, a supervised classification method relying on decision trees, was executed

using the "randomForest" R package (version 4.7). We evaluated error rates for tree counts ranging from 1 to 500 and determined the optimal number of trees by selecting the configuration with the lowest error rate. Furthermore, we gauged the feature importance scores for each gene, identifying candidate biomarkers as those with importance values exceeding 2 for subsequent analysis.

Enrichment analysis

In our study, we employed the "clusterprofiler" R package (version 4.6.0) to conduct Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses (Yu et al., 2012). GO analysis is considered the gold standard for large-scale functional enrichment studies, as it covers various biological processes, molecular functions, and cellular components (The Gene Ontology Resource, 2019). Additionally, we utilized the KEGG database, which provides comprehensive information on genomes, biological pathways, diseases, and drugs (Kanehisa et al., 2023). Significant enrichment was defined as a critical *p*-value threshold of <0.05, and the results of the functional enrichment analysis are visually represented using bar charts.

Single-cell data source and preprocessing

We processed the dataset GSE122930 using the "Seurat" R package (version 4.3.0) (Butler et al., 2018). Initial quality control involved the removal of cells based on the following criteria (Papayannopoulos, 2018): cells with fewer than 200 genes or more than 5,000 genes (UMI >0) (Carmona-Rivera et al., 2019), cells with more than 20,000 UMI, and (Frangou et al., 2019a) cells with over 12.5% mitochondrial UMI count. Subsequently, the data were log-normalized using default parameters. We selected the 894 most variable genes using the "FindVariableGenes" function and scaled the data using the "ScaleData" function to remove unnecessary sources of variation. Principal Component Analysis (PCA) was performed using the "RunPCA" function, and the number of principal components was determined visually using the "ElbowPlot" function. We constructed a shared nearest neighbor (SNN) plot for the first 15 principal components with the "FindNeighbors" function and clustered the cells using the "FindClusters" function, setting the "Resolution" parameter to 0.6. For visualization, we utilized the "RunUMAP" function to create UMAP plots. To identify marker genes in each cluster, we employed the "FindAllMarkers" function, setting the parameter "min.pct" to 0.2 and the "thresh.use" parameter to 0.2. Additionally, we used the "celltypist" to assist in cellular annotation (Domínguez Conde et al., 2022), followed by manual annotation for further refinement.

Trajectory analysis of single cells

We used the CytoTRACE R package (version 0.3.3) to help predict the direction of cell differentiation (Gulati et al., 2020). For our single-cell trajectory analysis, we employed the R package Monocle2 (version 2.16.0) (Qiu et al., 2017). We initially identified clusters corresponding to cancer stem cells and epithelial cells, and subsequently loaded these clusters into the R environment. To facilitate the analysis, we created an object using the "newCellDataSet" function. Within the trajectory analysis, we harnessed the "FindVariableGenes" gene set to perform pseudo-temporal sorting of all cells within the target cell subpopulation. Next, we reduced the dimensions of the dataset using the "reduceDimension ()" function, utilizing the parameters "reduction_method = "DDRTree" and "max_components = 2." For visualization purposes, we employed the "plot_cell_trajectory" function to generate a spanning tree of cells. Finally, we utilized the "differentialGeneTest" function to identify genes that exhibited significant changes over pseudotime [q-value <10(-5)], and we visualized the expression changes of the top 100 genes over pseudotime using the "plot_pseudotime_heatmap" function.

Cell-cell communication analysis

We utilized CellChat, a tool that quantitatively infers intercellular communication networks from scRNA-seq data (Jin et al., 2021). Based on a database of mouse ligand-receptor interactions and pattern recognition techniques, CellChat can detect intercellular communication at the pathway level and calculate the communication network of aggregated cells. Use default settings for all parameters.

Transverse aortic constriction (TAC)

Male 8-week-old adult wild-type (WT) C57BL/6J mice were procured from Charles River Laboratory (Charles River, China). Briefly, after randomizing the mice into groups, consisting of three mice in the TAC group and three mice in the SHAM group, the mice were anesthetized using isoflurane and underwent a transthoracic thoracotomy. Following the exposure of the aortic arch, a suture was passed through the aortic arch positioned between the right innominate artery and the left common carotid artery. Subsequently, the aortic arch was ligated to a 27-gauge needle, and the needle was carefully withdrawn upon the completion of ligation. Mice in the sham-operated group underwent identical procedures but were not subjected to ligation. Following the surgical intervention, the mice were placed on a heating pad for recovery and closely monitored. Four weeks after either the sham or TAC surgery, the mice were anesthetized using an overdose of pentobarbital (100 mg/kg, Sigma-Aldrich), and their hearts were extracted through an open-chest procedure for subsequent molecular analysis.

RNA isolation and real-time quantitative PCR (qRT-PCR)

Total RNA was extracted from cardiac tissue using Freezol reagent (Vazyme, R711), following the manufacturer's instructions. Subsequently, qRT-PCR analysis was conducted on the QuantStudio™ 5 Real-Time PCR Detection System using

10.3389/fcell.2023.1258959

Gene name	Forword primer (5′→3′)	Reverse primer (5′→3′)
Mpo	AATATGGCACGCCCAACAAC	TCTCCCACCAAAACCGATCAC
Vnn3	GCTGTGGGTTCAATGGACACT	CTGCCAGCTTGATTGCACTCT
Cxcr2	GTAATTCTGGCCCTGCCCAT	CAGGATACGCAGTACGACCC
Fcgr4	GAGGTCCATATGGGCTGGCTA	CTTGCCTTTGCCGTTCTGTAA
Fpr2	CATTTGGTTGGTTCATGTGCAA	AATACAGCGGTCCAGTGCAAT
Gapdh	AGGTCGGTGTGAACGGATTTG	TGTAGACCATGTAGTTGAGGTCA

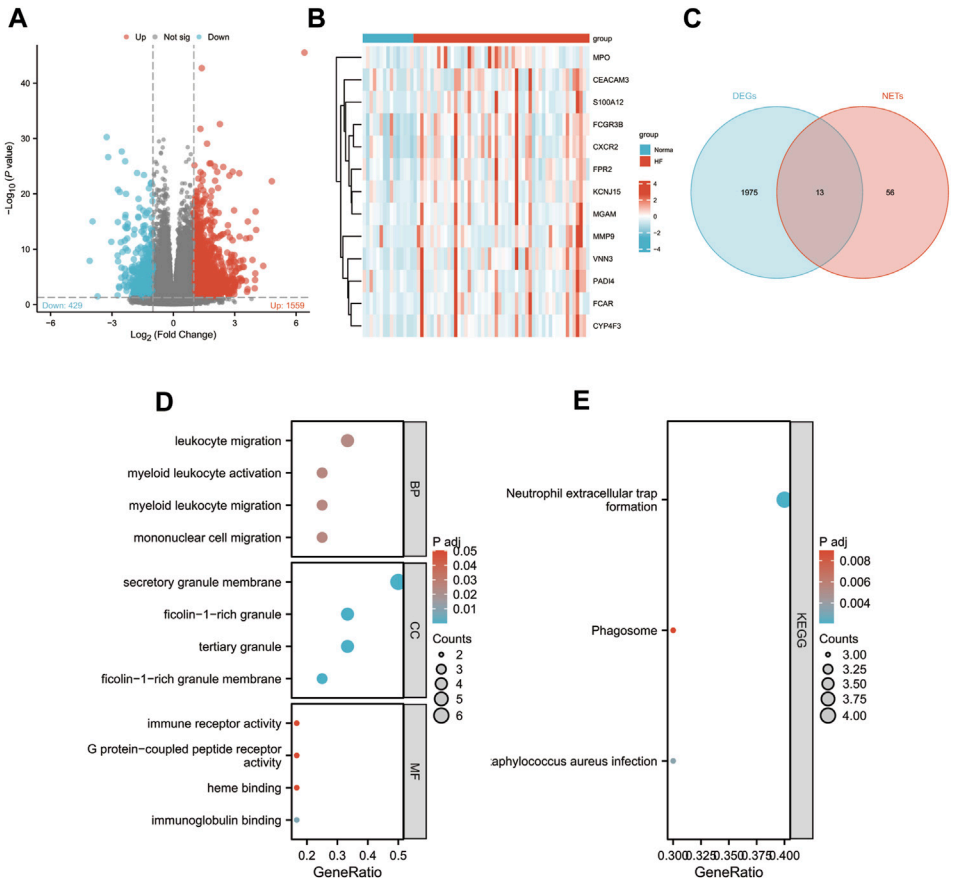


FIGURE 1
Identification of differentially expressed NETs-related genes. **(A)** Volcano plot of DEGs in GSE145154. **(B)** A Venn diagram of GSE145154 DEGs and NETs-related genes.**(C)** Expressions of differentially expressed NETs-related genes. **(D–E)** The enrichment pathway analysis of NETs-related DEGs.

ChamQ SYBR qPCR Master Mix (Low Rox Premixed) (Vazyme, Q331-02) and gene-specific primers. PCR analysis was performed on the QuantStudio™ 5 Real-Time PCR Detection System, with the following thermal cycling conditions: initial denaturation at 95°C for 1 min, followed by denaturation at 95°C for 10 s, annealing at 60°C for 30 s, and a total of 40 cycles. The relative expression levels of individual genes were quantified by the $2^{(-\Delta\Delta Ct)}$ method and normalized to the endogenous expression of glyceraldehyde-3-phosphate dehydrogenase (GAPDH). The sequences of the specific primers utilized for qRT-PCR in this study are provided in the below table.

Statistical analysis

The validation of key gene expression differences between the experimental and control groups in both the training and validation datasets was conducted using the Wilcoxon rank sum test. Additionally, the qRT-PCR validation results were analyzed employing Student's t-test. A significance threshold of $p < 0.05$ was applied to determine statistical significance. Data analysis and graph generation were performed using R software version 4.1.0 (http://www.R-project.org) and GraphPad Prism 8 (GraphPad Software, San Diego, CA, United States).

10.3389/fcell.2023.1258959

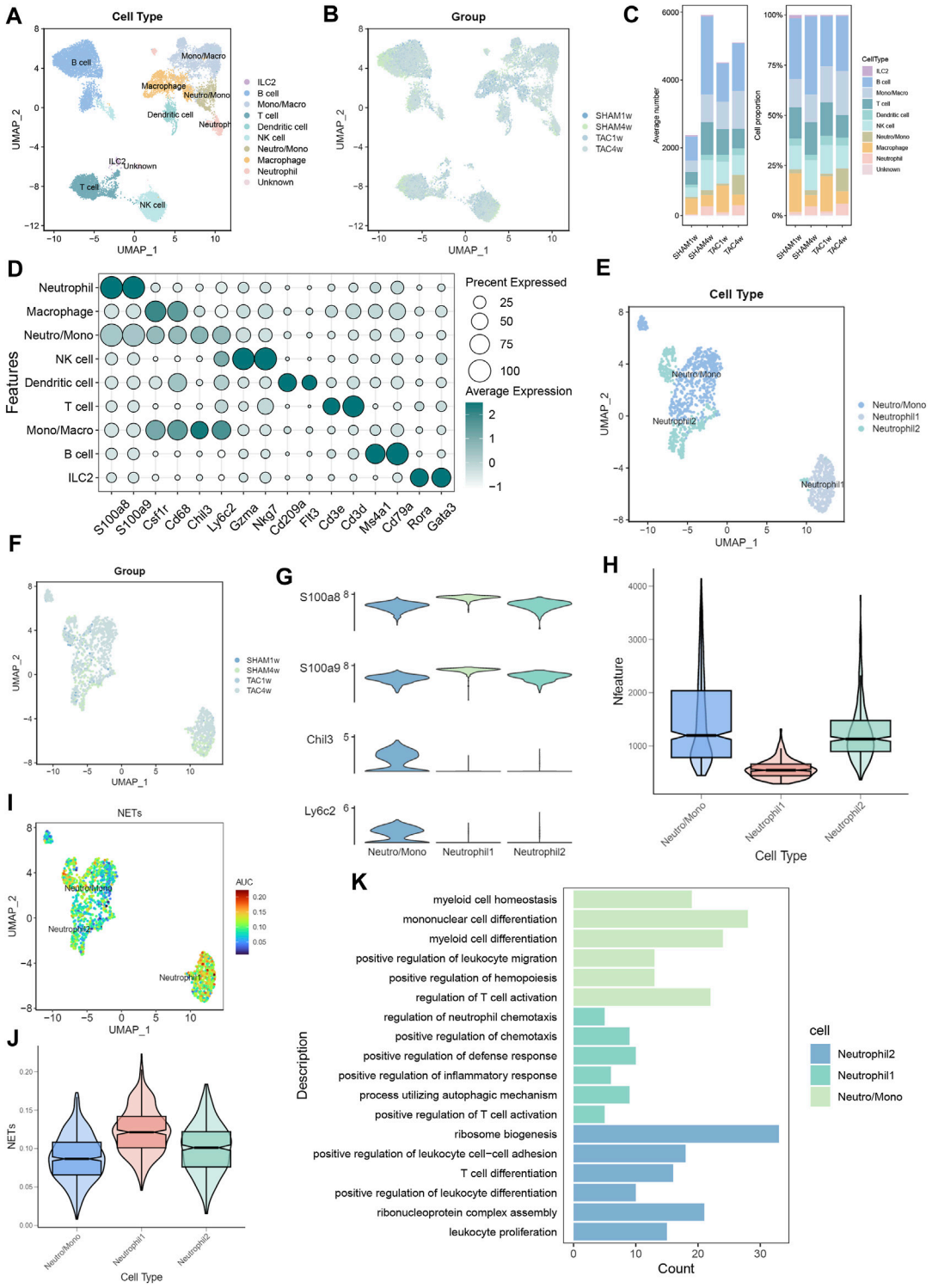


FIGURE 2
Identification of neutrophil subpopulations associated with NETs.**(A)** UMAP plots of 17,918 cells from 7 mouse heart samples.**(B)** UMAP plots showing the distribution of cells in SHAM1w, SHAM4w, TAC1w, and TAC4w heart samples.**(C)** Proportion of each cell subpopulation in different experimental groups.**(D)** Markers for different cell subpopulations.**(E)** UMAP plots of 1,451 neutrophils versus some monocytes from 7 mouse heart samples.**(F)** UMAP plots showing the distribution of neutrophils versus some monocytes in SHAM1w, SHAM4w, TAC1w, and TAC4w heart samples.**(G)** Violin plots showing markers for different neutrophil subpopulations.**(H)** Violin plots demonstrating the number of genes in different neutrophil subpopulations.**(I)** AUCELL quantifies NET activity in different neutrophil subpopulations.**(J)** Violin plot demonstrating NETs scores in different neutrophil subpopulations.**(K)** Bar graph demonstrating the results of enrichment analysis of different neutrophil subpopulations.

10.3389/fcell.2023.1258959

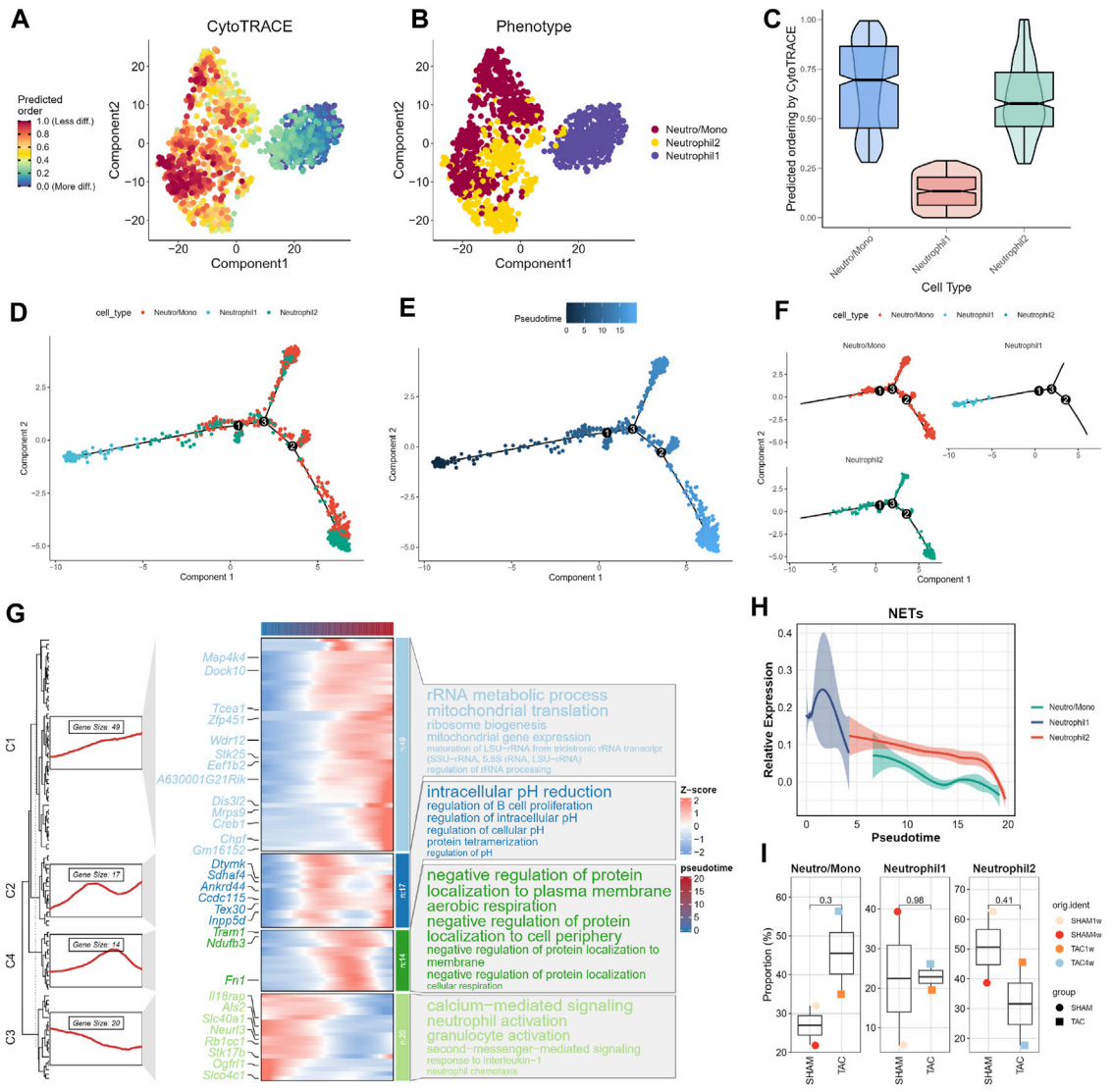


FIGURE 3
Trajectory analysis of neutrophils. **(A–C)** CytoTRACE predicts the cell differentiation potential of neutrophil subpopulations. **(D)** Distribution status of different neutrophil subpopulations in pseudotrajectory analysis. **(E)** Overall trajectory analysis of neutrophil subpopulations. **(F)** Independent distribution status of different neutrophil subpopulations in pseudotrajectories. **(G)** Distribution of key genes in trajectories and heatmap of gene enrichment analysis at different stages. **(H)** Dynamic expression profile of NETs activity. **(I)** Cell proportions of neutrophil subpopulations in different experimental subgroups.

Results

Identification of differentially expressed NETs-related genes

To explore the role of NETs-related genes in HF pathogenesis, we conducted an analysis using the GSE145154 dataset, which consisted of 52 HF patient samples and 15 control samples. Following differential analysis between the two groups, we identified 1998 differentially expressed genes (DEGs), with 1,559 genes upregulated and 429 downregulated in the HF group compared to the control group (Figure 1A). We intersected the list of DEGs with known NETs-related genes, identifying a subset of 13 genes (Figure 1C; Supplementary Table S2). We also visualized the expression of these genes among different groups by heat map (Figure 1B). Through gene ontology (GO) enrichment analysis of the differentially expressed NETs-related genes, we found their involvement in leukocyte, myeloid leukocyte, and mononuclear cell migration. Additionally, the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis suggested the activation of signaling pathways related to NET formation, phagosome, and *Staphylococcus aureus* infection (Figures 1D, E).

Identification of neutrophil subpopulations associated with NETs

To investigate the status of infiltrating neutrophils in both heart failure and normal cardiac tissues, we conducted a comprehensive

search within the GEO database. Subsequently, we acquired the GSE122930 dataset, which comprises 7 mouse hearts. This dataset includes two samples at 1 week post-TAC, two at 4 weeks post-TAC, one at 1 week post-SHAM, and two at 4 weeks post-SHAM, aligning with the objectives of our study. After rigorous quality control and meticulous screening, we identified a total of 10 distinct cell subpopulations from an initial pool of 17,918 cells, which were subsequently subjected to downstream analysis (Figures 2A, D). Importantly, when compared to the sham-operated mice, our cell scale plots revealed a substantial increase in neutrophil infiltration in mice observed 4 weeks after TAC (Figure 2C). Subsequently, the two neutrophil-related subpopulations (Neutro/Mono and Neutrophil) were extracted individually. Unsupervised clustering was then performed again, identifying three distinct cell clusters: Neutro/Mono (638 cells), Neutrophil 1 (409 cells), and Neutrophil 2 (404 cells) (Figures 2E,G). Additionally, the violin plot illustrates the number of genes expressed within these three cell subpopulations, with Neutro/Mono exhibiting the highest gene expression and Neutrophil 1 showing the lowest (Figure 2H). To gain insights into the diffrences between these neutrophil subpopulations, we quantified the activity of the NETs pathway using the AUCELL method, revealing that Neutrophil 1 displayed significantly higher activity in NETs-related gene sets in comparison to Neutrophil 2 and Neutro/Mono (Figures 2I, J). Furthermore, the results of our enrichment analysis shed light on the distinct functions of these three neutrophil-associated subpopulations (Figure 2K). Specifically, Neutro/Mono appeared to be linked with myeloid cell differentiation, Neutrophil 1 exhibited associations with the regulation of neutrophil chemotaxis, and Neutrophil 2 displayed a role in the positive regulation of leukocyte differentiation. Notably, all three cell subpopulations exhibited enrichment for terms related to T cell activation. Additionally, we conducted an assessment of the overall distribution of cell populations within both normal and disease groups within the dataset, employing UMAP plots, and provided characterization for each cell population utilizing known cell markers (Figures 2B, F).

Trajectory analysis of neutrophils

We assessed the differentiation capacity of neutrophil-associated cell subpopulations using CytoTRACE. Neutro/Mono exhibited the highest predicted differentiation potential, while Neutrophil1 was predicted to have the lowest differentiation capacity (Figures 3A–C; Supplementary Table S3). This suggests that the Neutro/Mono cell subpopulation may play a role in initiating the differentiation of the neutrophil population. To further elucidate these findings, we integrated this result with pseudotrajectory analysis. Our analysis revealed that Neutrophil1 was positioned at the end of the differentiation trajectory, Neutrophil2 was located at the pre-mid differentiation stage, and Neutro/Mono was situated at the pre-differentiation stage (Figures 3D–F). This suggests that Neutrophil1 may represent mature neutrophils, whereas Neutrophil2 represents immature neutrophils. We also generated a heatmap illustrating the key genes involved at each stage of the neutrophil differentiation process, along with the results of enrichment analysis at different stages (Figure 3G). Additionally, curve graphs were used to visualize the NETs activities of different cell subpopulations along the pseudotemporal ordering, with Neutrophil1 exhibiting the highest activity (Figure 3H).

Finally, we combined the results of the pseudotrajectory analysis to demonstrate changes in the proportions of different neutrophil-associated subpopulations in various experimental subgroups (Figure 3I). Given that one mouse heart sample was lost due to sample wetting failure 1 week after SHAM surgery in the original study, resulting in a smaller number of cardiac tissues in the SHAM group compared to the TAC group, we focused on the changes in cell proportions in TAC1w and TAC4w. This period represents a critical transition from cardiac hypertrophy to heart failure. During this period, the proportion of Neutrophil1 cells slightly increased, while the Neutro/Mono cell proportion significantly increased. This suggests that bone marrow hematopoiesis remained active during this time. In contrast, the proportion of Neutrophil2 cells significantly decreased, indicating that this period allows for the differentiation of more young neutrophils into mature neutrophils.

Cell communication analysis

We performed CellChat analysis to identify key cell subpopulations and receptor-ligand pairs involved in interactions with neutrophils. Initially, we explored the communication patterns among all immune cell subpopulations and their interactions with neutrophils in TAC mice. Our findings indicated that macrophages were the most active communicating cell subpopulation in TAC mice (Figure 4A). Both macrophages and T cells displayed close communication with neutrophil-associated clusters (Figure 4B). Furthermore, the overall number of immune cell subpopulations and the strength of intercellular communication were increased in TAC mice compared to the sham-operated mouse group (Figure 4C).

To gain insights into the specific signaling pathways associated with neutrophils, we examined the communication between neutrophil-associated subpopulations and other cellular subpopulations, considering neutrophils as receptors and senders, respectively. We focused on understanding the communication between neutrophils and macrophages, taking into account cell subpopulations and experimental groupings. Combining these results with our trajectory analysis, we observed a progressive increase in chemokine expression in neutrophils during maturation, particularly in the hearts of mice in the TAC group (Figure 4D). Previous reports in the literature have indicated that neutrophils are the first immune cells recruited in large numbers into the myocardium after pressure overload. They produce cytokines and chemokines to attract splenic-derived macrophages to migrate into cardiac tissues. Consequently, we paid particular attention to the communication with macrophages when neutrophil 1 served as a sender. Our findings revealed that the Ccl6-Ccr2 receptor-ligand pair exhibited the closest communication. Studies show resident CCR2+ cardiac macrophages promote neutrophil infiltration into injured myocardial tissue (Li et al., 2016).

Moreover, we also focused on the communication between neutrophils and T cells. The results of our previous enrichment analysis suggested that all three types of neutrophil-associated subpopulations might be related to T cell activation. We

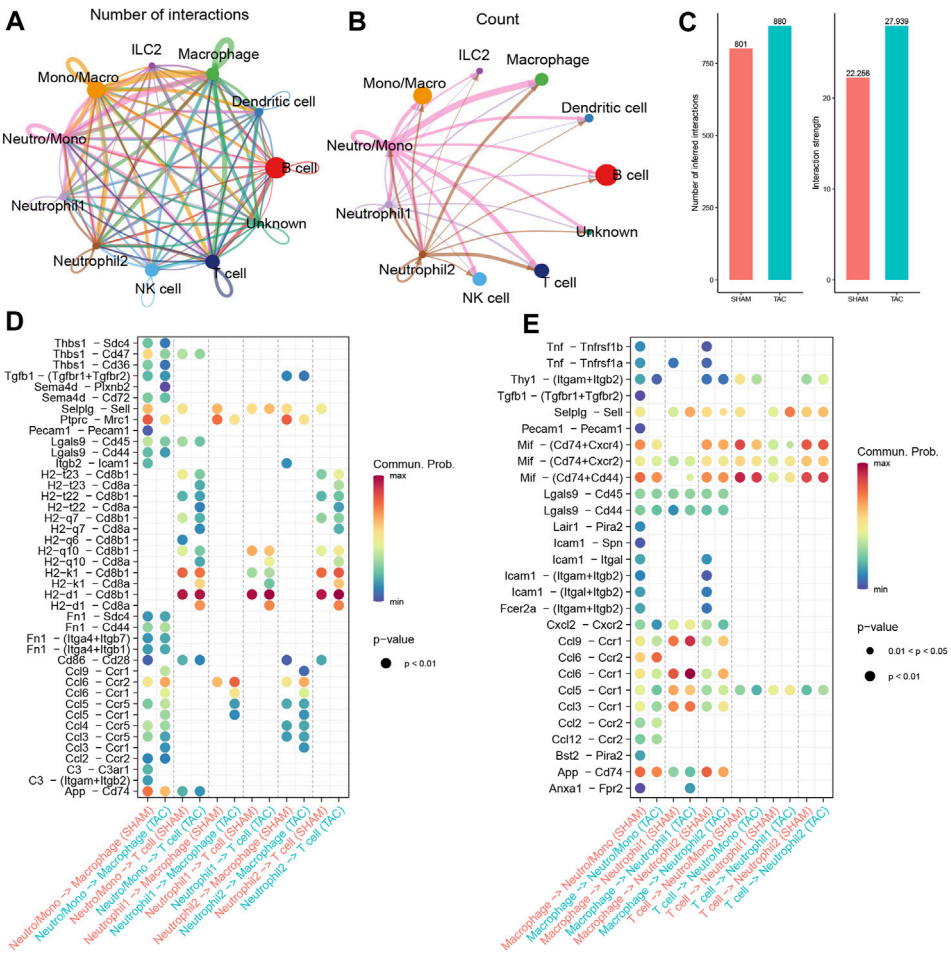


FIGURE 4
Cell Communication Analysis. **(A)** The number of interactions among immune cell subpopulations in the hearts of TAC mice. **(B)** Communication condition between neutrophils and other cell clusters. **(C)** Number and strength of interactions between SHAM group and TAC group. **(D–E)** Receptor-ligand pairs of different neutrophil subsets interacting with macrophages and T cells.

observed a significant number of H2-Cd8 receptor-ligand pairs in the communication signals when neutrophils served as senders (Figure 4E). Cd8, acting as a co-stimulatory molecule, interacts with MHC I molecules to enhance TCR recognition of MHC-antigen peptide complexes (H2 molecules in mice), thereby promoting T-cell activation. However, when combined with trajectory analyses, most of the receptor-ligand pairs either disappeared or showed reduced intensity during the transition from Neutro/Mono to mature neutrophils. This implies that the ability to activate T cells may not be specific to neutrophils.

Screening of NETs-related biomarkers in HF using machine learning

In the training set GSE145154, we employed two machine learning algorithms to identify the featured genes from among the candidate key genes in heart failure patients. In the Lasso regression analysis, we inputted the 13 NETs-related genes and performed a 10-fold cross-validation (Figures 5A, B). We used lambda, determined based on the minimum binomial deviation, as the criterion, ultimately identifying five candidate genes. Additionally, we utilized the RF machine learning algorithm to rank these 13 genes based on their importance variables. Genes with a MeanDecreaseGini greater than 2 were extracted (Figures 5C, D). Through taking the intersection of genes from the LASSO and random forest algorithms, the study found five common signature genes, namely, CXCR2, FCGR3B, VNN3, FPR2, and MPO (Figures 5E; Supplementary Table S4).

Validation of key biomarkers

In this study, we utilized subject working characteristic curves to assess the diagnostic value of five key biomarkers in HF. Our results indicate that HF patients had elevated expression levels of all key genes (Figures 6B, D). In the training set GSE145154, all pivotal genes showed an AUC greater than 0.700, excluding MPO, and VNN3 had the highest diagnostic value with an AUC of 0.774 (Figure 6A; Supplementary Table S5). In the validation set GSE116250, the diagnostic value of the identified key genes was further confirmed, all of the key genes demonstrated significant

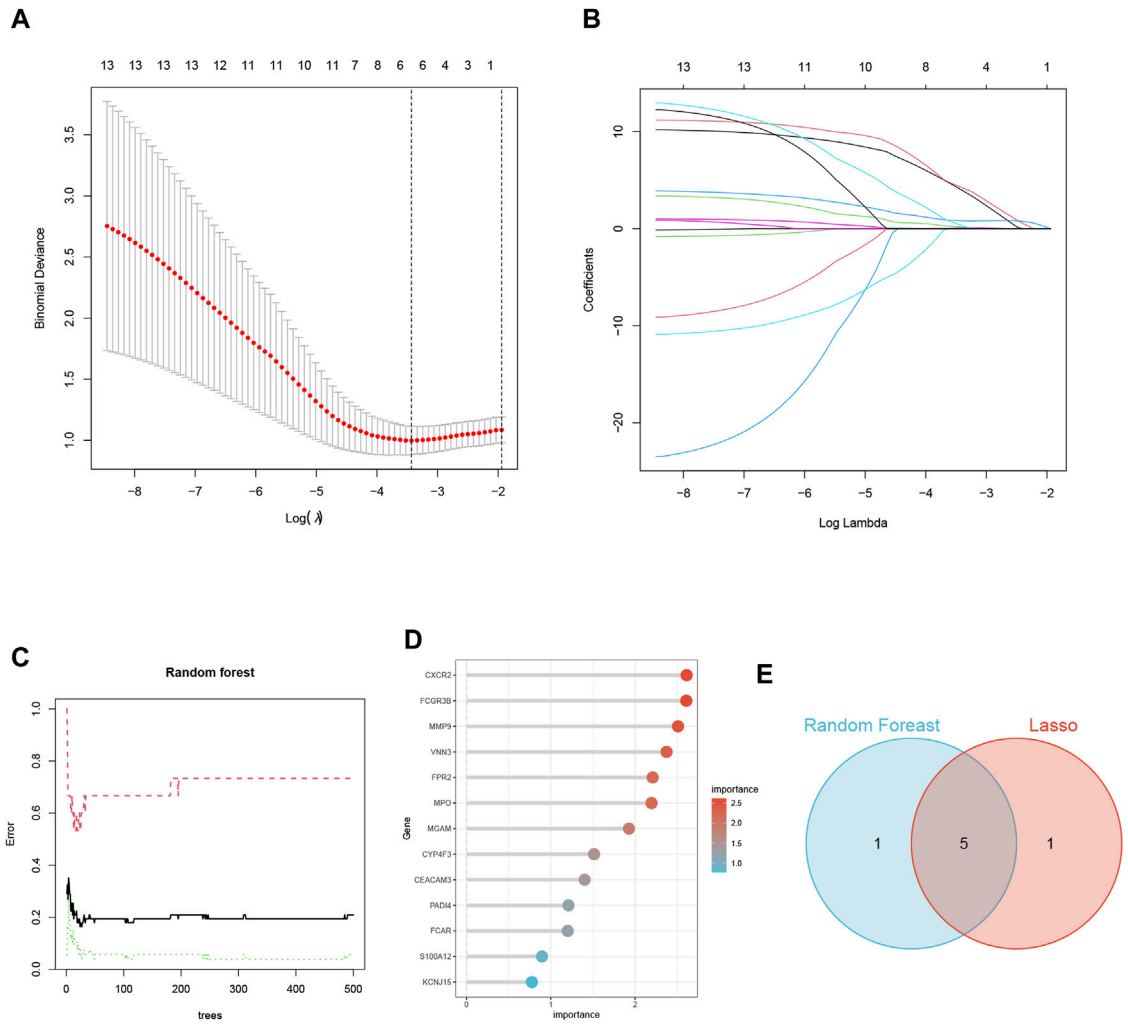


FIGURE 5
Screening of NETs-Related Biomarkers in HF Using Machine Learning. **(A)** Map of the regression coefficients of the 9 genes in LASSO model. **(B)** 6 hub genes screened by 10-fold cross-validation in the LASSO regression model. **(C)** The influence of the number of decision trees on the error rate. **(D)** The Gini Coefficient Method Achieved in Random Forest Classifier Results. **(E)** Venn Diagram Illustrating the Overlap of Two Machine Learning Screens for Genes.

diagnostic value except for Mpo, which had an AUC greater than 0.700 (Figure 6C; Supplementary Table S6). Lastly, we analyzed the expression of these biomarkers in mouse cardiac immune cells. We found that while VNN3 was undetectable in the immune cell population, the other four genes were expressed (Figure 6E). Among these genes, Mpo exhibited predominant expression in ILC2 cells, whereas Fpr2 and Cxcr2 displayed higher expression levels in neutrophil 1 compared to other cell types. Fcgr4, on the other hand, was expressed in neutrophils, monocytes, and macrophages, with slightly higher expression in Neutrophil2 compared to other cells (Figure 6F). To validate these findings, we examined the expression of these five genes in both TAC mice and SHAM mice. The results revealed that Cxcr2, Fpr2, and Vnn3 exhibited significantly higher expression in TAC mice compared to the control group (Figures 6G, I, J). However, Mpo expression was too low to be reliably quantified, and Fcgr4 did not show any significant differences between the groups (Figure 6H; Supplementary Table S7).

Discussion

The treatment of HF is a significant challenge for experts and researchers, given its complex pathogenesis and irreversible nature (Wang et al., 2017). Historically, innovative drugs such as ACE inhibitors have been the primary method of treating HF. However, recent studies suggest that immune cells, particularly neutrophils, may play a role in the disease's progression. In mice with ANGII-induced HF, DNaseI administration resulted in the clearance of NETs and reduced cardiomyocyte death (Tang et al., 2022). Furthermore, neutrophil depletion was shown to reduce TAC-induced hypertrophy and inflammation, thus preserving cardiac function (Wang et al., 2019). While there is some direct experimental evidence implicating NETs in the progression of HF disease, our current understanding of this aspect remains limited, highlighting the need for further research in this area.

This study employed a comprehensive analysis by combining single-cell sequencing and bulk transcriptome sequencing to

10.3389/fcell.2023.1258959

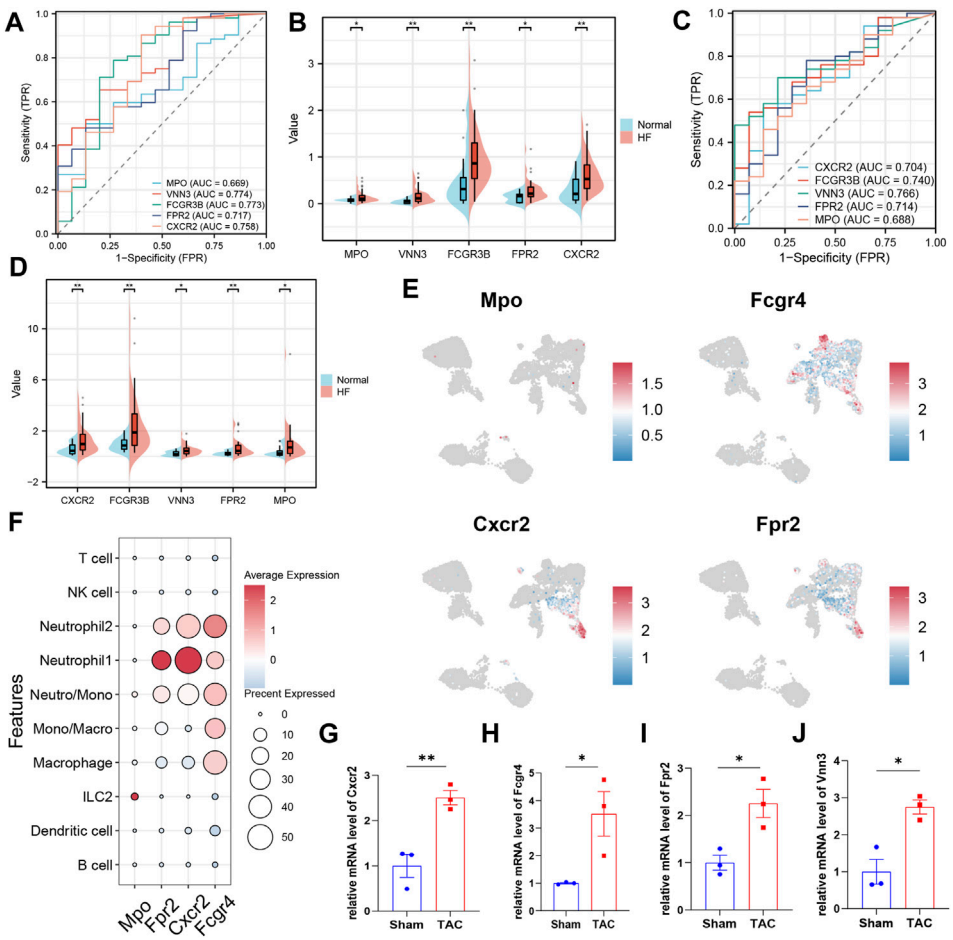


FIGURE 6
Validation of key biomarkers. **(A)** ROC curves were conducted for evaluations of the diagnostic potential of candidate genes in the GSE145154 dataset. **(B)** The pod plot showed gene expression differences between HF and normal groups in GSE145154 (a wilcoxon rank sum test was uesd). "***," "**," "*" represent $p < (0.001, 0.01, 0.05)$.**(C)** ROC curves were conducted for evaluations of the diagnostic potential of candidate genes in the GSE116250 dataset. **(D)** The pod plot showed gene expression differences between HF and normal groups in GSE116250 (a wilcoxon rank sum test was uesd). "***," "**," "*" represent $p < (0.001, 0.01, 0.05)$. **(E)** The UMAP plot illustrates the distribution of key genes across subpopulations of immune cells in the mouse heart. **(F)**The dot plots displayed the expression of key genes across various subpopulations of immune cells. **(G–J)** qRT-PCR to verify the expression of key genes (3 mice per group, a Student's t-test was used). "***," "**," "*" represent $p < (0.001, 0.01, 0.05)$.

elucidate the biological significance of NETs in HF from multiple perspectives. Using two machine learning algorithms, LASSO and random forest, the study identified five NET-related biomarkers in HF patients: CXCR2, FCGR3B, VNN3, FPR2, and MPO, and experimental verification confirmed that the expression of the first four of these genes was significantly elevated in heart failure mice.

Among these biomarkers,FcγRIIIb is a glycosphingolipid (GPI)-anchored receptor exclusively expressed on neutrophils and plays a crucial role in the activation of NETs released by neutrophils (David et al., 2005). Previous studies have reported that the kinases Syk and TAK1 are involved in the signaling pathway that leads to the formation of NETs upon FcγRIIIb stimulation in neutrophils (Fonseca et al., 2018). CXCR2 is a prominent chemokine receptor expressed on neutrophils. Studies have demonstrated that CXCR2 and its downstream pathways through IL8 agonism mediate the classical pathway of NETosis (Chen et al., 2021). Moreover, the CXCL1-CXCR2 axis mediates cardiac hypertrophy and remodeling in HF model mice by regulating monocyte infiltration (Wang et al., 2018). FPR2 is a multifunctional G protein-coupled receptor with a 7-transmembrane structural domain (Dahlgren et al., 2020). Previous studies have reported that FPR2 can reduce hyperosmolarity-induced NETosis, which helps alleviate dry eye (Tibrewal et al., 2014). In acute HF, Fpr2 triggers increased infiltration of immature and inactive neutrophils in the heart (Kain et al., 2019). VNN3 is a secreted and membrane-bound exoenzyme involved in the conversion of pantothenic acid and cysteamine. A previous study utilizing blood transcriptome-based molecular signatures identified VNN3 as a potential diagnostic biomarker for ST-segment elevation myocardial infarction. Out of the five NET-related genes, we discovered that CXCR2, FPR2, and FCGR4 demonstrated significantly higher expression in Neutrophil 1 as compared to Neutrophil 2. This finding corroborates our observations in the mouse single-cell dataset.

The single-cell sequencing data used in this study were derived from previous studies, but our study differs from the original analysis. The original study classified neutrophils into two classes

of cells by unsupervised clustering and identified CXCR2 as a gene specifically expressed in Neutrophil1, but failed to further discuss the function and significance of neutrophils. In our present study, we harnessed advanced cellular annotation tools to achieve precise annotation of key cellular subpopulations. Subsequently, we employed various methods to uncover distinct functions within neutrophil subpopulations, delve into the intricacies of neutrophil differentiation, and identify pivotal receptor-ligand pairs regulating intercellular interactions. These findings significantly contribute to our understanding of neutrophils in the context of heart failure.

The present study is subject to limitations, one of which is the relatively low number of neutrophils obtained from single-cell sequencing data. This could be attributed to several factors, such as the fact that neutrophils are not resident in cardiac tissue and only migrate to sites of injury or inflammation to perform their functions. Moreover, neutrophils have lower levels of gene expression and are more sensitive to the experimental environment, which may affect their ability to be captured by single-cell sequencing methods. It is challenging to obtain a sufficient population of neutrophils in single-cell sequencing studies without specifically targeting this cell type. Therefore, future experimental studies are still required to confirm our findings.

Conclusion

In this study, we conducted an in-depth investigation into the functions of neutrophil subpopulations that infiltrate cardiac tissue in TAC mice. Additionally, we identified four biomarkers (CXCR2, FCGR3B, VNN3, and FPR2)associated with NETs in HF. Our findings enhance the understanding of immunology in HF.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

Ethics statement

Ethical approval was not required for the study involving humans in accordance with the local legislation and institutional requirements. Written informed consent to participate in this study was not required from the participants or the participants' legal guardians/next of kin in accordance with the national legislation and the institutional requirements. The animal study was approved by Experimental Animal Care and Use Committee of Nanjing Medical University. The study was conducted in accordance with the local legislation and institutional requirements.

Author contributions

XL: Data curation, Software, Writing–original draft. XC: Validation. QL: Data curation, Supervision, Writing–review and editing. QS: Supervision, Writing–review and editing. LZ: Data curation, Funding acquisition, Writing–original draft.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Acknowledgments

We would like to express our gratitude to the original research scholars who uploaded the following datasets for subsequent analysis: GSE145154(36), GSE116250(37), and GSE122930(38).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcell.2023.1258959/full#supplementary-material

SUPPLEMENTARY TABLE S1
69 initial biomarkers of NETs.

SUPPLEMENTARY TABLE S2
13 NETs-related differential genes were obtained between HF and control samples.

SUPPLEMENTARY TABLE S3
The CytoTRACE values of two clusters of neutrophils.

SUPPLEMENTARY TABLE S4
Biomarkers related to NETs identified through machine learning.

SUPPLEMENTARY TABLE S5
The expression status of five NETs-related biomarkers in the GSE145154 dataset.

SUPPLEMENTARY TABLE S6
The expression status of five NETs-related biomarkers in the GSE116250 dataset.

SUPPLEMENTARY TABLE S7
Relative expression of four key genes in the SHAM and TAC groups.

References

Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. *Genome Biol.* 11 (10), R106. doi:10.1186/gb-2010-11-10-r106

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

Carmona-Rivera, C., Khaznadar, S. S., Shwin, K. W., Irizarry-Caro, J. A., O'Neil, L. J., Liu, Y., et al. (2019). Deficiency of adenosine deaminase 2 triggers adenosine-mediated NETosis and TNF production in patients with DADA2. *Blood* 134 (4), 395–406. doi:10.1182/blood.2018892752

Chen, T., Li, Y., Sun, R., Hu, H., Liu, Y., Herrmann, M., et al. (2021). Receptor-Mediated NETosis on neutrophils. *Front. Immunol.* 12, 775267. doi:10.3389/fimmu.2021.775267

Dahlgren, C., Holdfeldt, A., Lind, S., Mårtensson, J., Gabl, M., Björkman, L., et al. (2020). Neutrophil signaling that challenges dogmata of G protein-coupled receptor regulated functions. *ACS Pharmacol. Transl. Sci.* 3 (2), 203–220. doi:10.1021/acsptsci.0c00004

David, A., Fridlich, R., and Aviram, I. (2005). The presence of membrane Proteinase 3 in neutrophil lipid rafts and its colocalization with FcgammaRIIIb and cytochrome b558. *Exp. Cell Res.* 308 (1), 156–165. doi:10.1016/j.yexcr.2005.03.034

de Bont, C. M., Stokman, M. E. M., Faas, P., Thurlings, R. M., Boelens, W. C., Wright, H. L., et al. (2020). Autoantibodies to neutrophil extracellular traps represent a potential serological biomarker in rheumatoid arthritis. *J. Autoimmun.* 113, 102484. doi:10.1016/j.jaut.2020.102484

Domínguez Conde, C., Xu, C., Jarvis, L. B., Rainbow, D. B., Wells, S. B., Gomes, T., et al. (2022). Cross-tissue immune cell analysis reveals tissue-specific features in humans. *Science* 376 (6594), eabl5197. doi:10.1126/science.abl5197

Fonseca, Z., Díaz-Godínez, C., Mora, N., Alemán, O. R., Uribe-Querol, E., Carrero, J. C., et al. (2018). Entamoeba histolytica induce signaling via raf/MEK/ERK for neutrophil extracellular trap (NET) formation. *Front. Cell Infect. Microbiol.* 8, 226. doi:10.3389/fcimb.2018.00226

Frangou, E., Chrysanthopoulou, A., Mitsios, A., Kambas, K., Arelaki, S., Angelidou, I., et al. (2019a). REDD1/autophagy pathway promotes thromboinflammation and fibrosis in human systemic lupus erythematosus (SLE) through NETs decorated with tissue factor (TF) and interleukin-17A (IL-17A). *Ann. Rheum. Dis.* 78 (2), 238–248. doi:10.1136/annrheumdis-2018-213181

Frangou, E., Vassilopoulos, D., Boletis, J., and Boumpas, D. T. (2019b). An emerging role of neutrophils and NETosis in chronic inflammation and fibrosis in systemic lupus erythematosus (SLE) and ANCA-associated vasculitides (AAV): implications for the pathogenesis and treatment. *Autoimmun. Rev.* 18 (8), 751–760. doi:10.1016/j.autrev.2019.06.011

Gray, R. D., Hardisty, G., Regan, K. H., Smith, M., Robb, C. T., Duffin, R., et al. (2018). Delayed neutrophil apoptosis enhances NET formation in cystic fibrosis. *Thorax* 73 (2), 134–144. doi:10.1136/thoraxjnl-2017-210134

Guerra, M., Halls, V. S., Schatterny, J., Hagner, M., Mall, M. A., and Schultz, C. (2020). Protease FRET reporters targeting neutrophil extracellular traps. *J. Am. Chem. Soc.* 142, 20299–20305. doi:10.1021/jacs.0c08130

Gulati, G. S., Sikandar, S. S., Wesche, D. J., Manjunath, A., Bharadwaj, A., Berger, M. J., et al. (2020). Single-cell transcriptional diversity is a hallmark of developmental potential. *Science* 367 (6476), 405–411. doi:10.1126/science.aax0249

Jin, S., Guerrero-Juarez, C. F., Zhang, L., Chang, I., Ramos, R., Kuan, C. H., et al. (2021). Inference and analysis of cell-cell communication using CellChat. *Nat. Commun.* 12 (1), 1088. doi:10.1038/s41467-021-21246-9

Kain, V., Jadapalli, J. K., Tourki, B., and Halade, G. V. (2019). Inhibition of FPR2 impaired leukocytes recruitment and elicited non-resolving inflammation in acute heart failure. *Pharmacol. Res.* 146, 104295. doi:10.1016/j.phrs.2019.104295

Kanehisa, M., Furumichi, M., Sato, Y., Kawashima, M., and Ishiguro-Watanabe, M. (2023). KEGG for taxonomy-based analysis of pathways and genomes. *Nucleic Acids Res.* 51 (D1), D587–d592. doi:10.1093/nar/gkac963

Khandpur, R., Carmona-Rivera, C., Vivekanandan-Giri, A., Gizinski, A., Yalavarthi, S., Knight, J. S., et al. (2013). NETs are a source of citrullinated autoantigens and stimulate inflammatory responses in rheumatoid arthritis. *Sci. Transl. Med.* 5 (178), 178ra40. doi:10.1126/scitranslmed.3005580

Kraaij, T., Kamerling, S. W. A., de Rooij, E. N. M., van Daele, P. L. A., Bredewold, O. W., Bakker, J. A., et al. (2018). The NET-effect of combining rituximab with belimumab in severe systemic lupus erythematosus. *J. Autoimmun.* 91, 45–54. doi:10.1016/j.jaut.2018.03.003

Li, W., Hsiao, H. M., Higashikubo, R., Saunders, B. T., Bharat, A., Goldstein, D. R., et al. (2016). Heart-resident CCR2(+) macrophages promote neutrophil extravasation through TLR9/MyD88/CXCL5 signaling. *JCI Insight* 1 (12), e87315. doi:10.1172/jci.insight.87315

Martini, E., Kunderfranco, P., Peano, C., Carullo, P., Cremonesi, M., Schorn, T., et al. (2019). Single-cell sequencing of mouse heart immune infiltrate in pressure overload-driven heart failure reveals extent of immune activation. *Circulation* 140 (25), 2089–2107. doi:10.1161/circulationaha.119.041694

Martinod, K., Witsch, T., Erpenbeck, L., Savchenko, A., Hayashi, H., Cherpokova, D., et al. (2017). Peptidylarginine deiminase 4 promotes age-related organ fibrosis. *J. Exp. Med.* 214 (2), 439–458. doi:10.1084/jem.20160530

Morán, G., Uberti, B., and Quiroga, J. (2022). Role of cellular metabolism in the formation of neutrophil extracellular traps in airway diseases. *Front. Immunol.* 13, 850416. doi:10.3389/fimmu.2022.850416

O'Neil, L. J., Oliveira, C. B., Wang, X., Navarrete, M., Barrera-Vargas, A., Merayo-Chalico, J., et al. (2023). Neutrophil extracellular trap-associated carbamylation and histones trigger osteoclast formation in rheumatoid arthritis. *Ann. Rheum. Dis.* 82 (5), 630–638. doi:10.1136/ard-2022-223568

Papayannopoulos, V. (2018). Neutrophil extracellular traps in immunity and disease. *Nat. Rev. Immunol.* 18 (2), 134–147. doi:10.1038/nri.2017.105

Qiu, X., Mao, Q., Tang, Y., Wang, L., Chawla, R., Pliner, H. A., et al. (2017). Reversed graph embedding resolves complex single-cell trajectories. *Nat. Methods* 14 (10), 979–982. doi:10.1038/nmeth.4402

Rao, M., Wang, X., Guo, G., Wang, L., Chen, S., Yin, P., et al. (2021). Resolving the intertwining of inflammation and fibrosis in human heart failure at single-cell level. *Basic Res. Cardiol.* 116 (1), 55. doi:10.1007/s00395-021-00897-1

Sakkas, L. I., Bogdanos, D. P., Katsiari, C., and Platsoucas, C. D. (2014). Anti-citrullinated peptides as autoantigens in rheumatoid arthritis-relevance to treatment. *Autoimmun. Rev.* 13 (11), 1114–1120. doi:10.1016/j.autrev.2014.08.012

Skopelja, S., Hamilton, B. J., Jones, J. D., Yang, M. L., Mamula, M., Ashare, A., et al. (2016). The role for neutrophil extracellular traps in cystic fibrosis autoimmunity. *JCI Insight* 1 (17), e88912. doi:10.1172/jci.insight.88912

Tang, X., Wang, P., Zhang, R., Watanabe, I., Chang, E., Vinayachandran, V., et al. (2022). KLF2 regulates neutrophil activation and thrombosis in cardiac hypertrophy and heart failure progression. *J. Clin. Invest.* 132 (3), e147191. doi:10.1172/jci147191

The Gene Ontology Resource, (2019). The gene ontology Resource: 20 years and still GOing strong. *Nucleic Acids Res.* 47 (D1), D330–d338. doi:10.1093/nar/gky1055

Tibrewal, S., Ivanir, Y., Sarkar, J., Nayeb-Hashemi, N., Bouchard, C. S., Kim, E., et al. (2014). Hyperosmolar stress induces neutrophil extracellular trap formation: implications for dry eye disease. *Invest. Ophthalmol. Vis. Sci.* 55 (12), 7961–7969. doi:10.1167/iovs.14-15332

Wang, H., Anstrom, K., Ilkayeva, O., Muehlbauer, M. J., Bain, J. R., McNulty, S., et al. (2017). Sildenafil treatment in heart failure with preserved ejection fraction: targeted metabolomic profiling in the relax trial. *JAMA Cardiol.* 2 (8), 896–901. doi:10.1001/jamacardio.2017.1239

Wang, L., Zhang, Y. L., Lin, Q. Y., Liu, Y., Guan, X. M., Ma, X. L., et al. (2018). CXCL1-CXCR2 axis mediates angiotensin II-induced cardiac hypertrophy and remodelling through regulation of monocyte infiltration. *Eur. Heart J.* 39 (20), 1818–1831. doi:10.1093/eurheartj/ehy085

Wang, Y., Sano, S., Oshima, K., Sano, M., Watanabe, Y., Katanasaka, Y., et al. (2019). Wnt5a-Mediated neutrophil recruitment has an obligatory role in pressure overload-induced cardiac dysfunction. *Circulation* 140 (6), 487–499. doi:10.1161/circulationaha.118.038820

Yamaguchi, T., Sumida, T. S., Nomura, S., Satoh, M., Higo, T., Ito, M., et al. (2020). Cardiac dopamine D1 receptor triggers ventricular arrhythmia in chronic heart failure. *Nat. Commun.* 11 (1), 4364. doi:10.1038/s41467-020-18128-x

Yu, G., Wang, L. G., Han, Y., and He, Q. Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics* 16 (5), 284–287. doi:10.1089/omi.2011.0118

Zhang, Y., Guo, L., Dai, Q., Shang, B., Xiao, T., Di, X., et al. (2022). A signature for pan-cancer prognosis based on neutrophil extracellular traps. *J. Immunother. Cancer* 10 (6), e004210. doi:10.1136/jitc-2021-004210

frontiers | Frontiers in Cardiovascular Medicine

TYPE Original Research
PUBLISHED 14 November 2023
DOI 10.3389/fcvm.2023.1221620

OPEN ACCESS

EDITED BY
Michel Puceat,
Institut National de la Santé et de la Recherche Médicale (INSERM), France

REVIEWED BY
Yifan Lu,
Temple University, United States
Chris Platsoucas,
Old Dominion University, United States

*CORRESPONDENCE
Victor Greiff
✉ victor.greiff@medisin.uio.no
Norbert Gerdes
✉ gerdes@hhu.de

RECEIVED 12 May 2023
ACCEPTED 16 October 2023
PUBLISHED 14 November 2023

CITATION
Elster C, Ommer-Bläsius M, Lang A, Vajen T, Pfeiler S, Feige M, Yau Pang T, Böttenberg M, Verheyen S, Lê Quý K, Chernigovskaya M, Kelm M, Winkels H, Schmidt SV, Greiff V and Gerdes N (2023) Application and challenges of TCR and BCR sequencing to investigate T- and B-cell clonality in elastase-induced experimental murine abdominal aortic aneurysm.
Front. Cardiovasc. Med. 10:1221620.
doi: 10.3389/fcvm.2023.1221620

Application and challenges of TCR and BCR sequencing to investigate T- and B-cell clonality in elastase-induced experimental murine abdominal aortic aneurysm

Christin Elster[1], Miriam Ommer-Bläsius[1], Alexander Lang[1], Tanja Vajen[1], Susanne Pfeiler[1], Milena Feige[1], Tin Yau Pang[1,2], Marius Böttenberg[1], Sarah Verheyen[1], Khang Lê Quý[3], Maria Chernigovskaya[3], Malte Kelm[1,4], Holger Winkels[5], Susanne V. Schmidt[6,7], Victor Greiff[3]* and Norbert Gerdes[1,4]*

[1]Division of Cardiology, Pulmonology, and Vascular Medicine, Medical Faculty and University Hospital, Heinrich Heine University, Düsseldorf, Germany, [2]Department of Biology, Institute for Computer Science, Heinrich Heine University, Düsseldorf, Germany, [3]Department of Immunology, University of Oslo and Oslo University Hospital, Oslo, Norway, [4]Cardiovascular Research Institute Düsseldorf (CARID), Medical Faculty, Heinrich Heine University, Düsseldorf, Germany, [5]Department of Cardiology, Faculty of Medicine and University Hospital Cologne, University of Cologne, Cologne, Germany, [6]Institute of Innate Immunity, Medical Faculty and University Hospital, Rheinische Friedrich-Wilhelms-University, Bonn, Germany, [7]Institute of Clinical Chemistry and Clinical Pharmacology, University Hospital Bonn, Bonn, Germany

Background: An abdominal aortic aneurysm (AAA) is a life-threatening cardiovascular disease. Although its pathogenesis is still poorly understood, recent evidence suggests that AAA displays autoimmune disease characteristics. Particularly, T cells responding to AAA-related antigens in the aortic wall may contribute to an initial immune response. Single-cell RNA (scRNA) T cell receptor (TCR) and B cell receptor (BCR) sequencing is a powerful tool for investigating clonality. However, difficulties such as limited numbers of isolated cells must be considered during implementation and data analysis, making biological interpretation challenging. Here, we perform a representative single-cell immune repertoire analysis in experimental murine AAA and show a reliable bioinformatic processing pipeline highlighting opportunities and limitations of this approach.

Methods: We performed scRNA TCR and BCR sequencing of isolated lymphocytes from the infrarenal aorta of male C57BL/6J mice 3, 7, 14, and 28 days after AAA induction via elastase perfusion of the aorta. Sham-operated mice at days 3 and 28 and non-operated mice served as controls.

Results: Comparison of complementarity-determining region (CDR3) length distribution of 179 B cells and 796 T cells revealed neither differences between AAA and control nor between the disease stages. We found no clonal expansion of B cells in AAA. For T cells, we identified several clones in 11 of 16 AAA samples and one of eight control samples. Immune receptor repertoire comparison indicated that only a few clones were shared between the individual AAA samples. The most frequently used V-genes in the TCR beta chain in AAA were TRBV3, TRBV19, and the splicing variant TRBV12-2 + TRBV13-2.

Conclusion: We found no clonal expansion of B cells but evidence for clonal expansion of T cells in elastase-induced AAA in mice. Our findings imply that a more precise characterization of TCR and BCR distribution requires a more extensive number of lymphocytes to prevent undersampling and potentially detect rare clones. Thus, further experiments are necessary to confirm our findings. In summary, this paper examines TCR and BCR sequencing results, identifies limitations and pitfalls, and offers guidance for future studies.

KEYWORDS

aortic aneurysm, single-cell sequencing (scRNA-seq), T cell receptor (TCR), B cell receptor (BCR), clonality analysis

Introduction

An abdominal aortic aneurysm (AAA) is a cardiovascular disease characterized by a permanent dilation of the abdominal aorta greater than 50% or 3 cm. Most AAAs develop in the infrarenal region between the renal veins and the aortic bifurcation (1). The prevalence of AAA is 4%–8% in men older than 60 years and 0.5%–1.5% in women, with the rupture of AAA conferring a high mortality rate (2). AAA is a multifactorial and progressive disease. Genetic factors and inflammation strongly contribute to AAA development (3), and several studies have revealed recently that autoimmunity may contribute to the pathogenesis of AAA (4–8).

Inflammation and immune cell recruitment are characteristics of AAA. Accordingly, T and B cells are among the predominant infiltrating immune cells in human AAA tissue (4, 8, 9). The presence of these lymphocytes in AAA tissue was confirmed in several experimental mouse models of AAA including the porcine pancreatic elastase (PPE) perfusion model, which was used for this study (10–14). The PPE model produces infrarenal aortic aneurysms and is considered the experimental mouse model most resembling human AAA, although the aneurysms do not form intraluminal thrombus or rupture (15). Several experimental interventions (e.g., HIF-1α inhibitors, PI3Kγ inhibitors) preventing elastase-induced AAA are associated with decreased numbers of lymphocytes in the aneurysmal tissue (16–18) (**Supplementary Table S1**), suggesting an important role for lymphocytes in AAA development in this model.

Previous reports have implicated both T helper-1 (Th1) and T helper-2 (Th2) cells in various stages of AAA development (19). Th2 cells release inflammatory mediators and cytokines such as interleukins 4, 5, 9, 10, and 13 and Fas ligand that may contribute to the regulation of AAA progression, whereas Th1-derived interferon-gamma and CD40 ligand are associated with macrophage activation, regulation of vascular smooth muscle cell apoptosis, and aortic wall remodeling (2, 19–22). Investigation of the T cell receptor (TCR)/antigen/human leukocyte antigen (HLA) complex revealed evidence that AAA encompasses a specific antigen-driven T cell response (4, 5). Studies discovered the clonal expansion of T cells in AAA lesions, linked AAA to specific HLA Class I and Class II types, and identified self- or non-self-antigens that may be associated with AAA (4, 6, 23).

The role of B cells in AAA is controversially discussed. B cell-derived immunoglobulins (Ig), such as IgM and IgG, localize in AAA tissue, where they promote inflammation and tissue degradation (10, 24). B cell depletion can prevent AAA growth in experimental models. One study showed that B cell depletion with an anti-CD20 antibody suppressed AAA growth in the angiotensin-II- and the elastase perfusion AAA model (25). Another study showed that AAA was induced by periaortic application of $CaCl_2$, and significantly smaller AAAs in B cell-deficient muMT mice compared to wild-type (WT) mice were observed (10). Injection of polyclonal IgG antibodies into muMT mice resulted in AAA size comparable to WT mice indicating that IgG alone is sufficient to promote AAA development (10). In contrast, Meher et al. (11) could not observe differences in experimental AAA formation between muMT mice and WT mice and showed that adoptive transfer of B2 cells suppressed AAA formation and decreased infiltration of mononuclear cells into aneurysmal tissue. However, there is evidence that an autoimmune process directed against self-antigens in the aortic wall may play a role in AAA pathogenesis. Zhou et al. (23) identified a natural IgG antibody against fibrinogen in aortic tissues of elastase-induced AAA that induced AAA formation by activating the complement lectin pathway. Other findings suggest that a collagen-associated 80-kDA protein from the aneurysm wall is a potential target of the autoimmune response in AAA disease (26). IgG antibodies, including autoantibodies, have been isolated from the aortic wall of patients with AAA, and eight of 10 of these AAA wall IgGs reacted with an 80-kDa protein from aortic microfibrillar extracts shown by Western blotting. This protein was found to be located in the adventitial connective tissue matrix confirmed by immunohistochemistry (26). Further investigation of the role of T cells, B cells, and Ig involved in AAA is essential to improve the understanding of AAA pathogenesis. Clonality and diversity analysis of the adaptive immune receptor repertoire (AIRR) provide insights into disease mechanisms. Such analysis may further define the immunological status of an individual, thus proving useful for disease diagnosis and risk stratification (27).

AIRR sequencing is increasingly used to investigate lymphocyte dynamics in pathological contexts such as autoimmune and sterile inflammatory diseases, cancer, and infections (28). TCR and B cell receptors (BCR) are highly diverse heterodimers that recognize an immense variety of antigens (29). The receptors consist of a combination of heavy and light chains in the case of BCRs and a combination of α/β or γ/δ chains in the case of TCRs. The

majority of TCRs expressed in T cells consist of a combination of α and β chains. The receptors are formed by variable, diversity, and joining (VDJ) recombination, which is the rearrangement of the V-, D-, and J-gene segments. For TCR, α chains, and BCR light chains, only V- and J-genes are involved in the recombination. Additional diversity is achieved by adding or deleting random nucleotides at the junction sites between the gene segments and by the chain pairing. Somatic hypermutation results in greater BCR diversity. Each receptor chain contains three hypervariable loops termed complementarity determining regions (CDR) that are required for the interaction of the receptors with the antigen. CDR3 is commonly used as a region of interest to determine T and B cell clones due to its high diversity and essential role in antigen binding (29–32). A clone is a set of cells expressing the same immune receptor, which implies that the receptors consist of the same V-, D-, and J-genes and encode the identical CDR3 nucleotide sequence. The AIRR is the union of all TCRs and BCRs of one individual and can change greatly with the onset and progression of diseases (29). The TCR repertoire within one individual is estimated to comprise 10^7 in humans and 10^6 in mice (33), whereas the estimated size of the B cell repertoire is 10^{18} in humans and 10^{13} in mice (34, 35).

Previously, TCRs were analyzed in aortic aneurysms, and their function has been investigated in mice and humans, yet no study addressed B cell clonality in aortic aneurysms. Li et al. (36) found clonal expansion of regulatory T cells (Treg) in mouse aortae after elastase-induced AAA formation, and several studies showed the presence of clonally expanded TCRs in aneurysmal lesions of patients with AAA or ascending thoracic aortic aneurysms, supporting the notion that AAA may be promoted by specific antigen-driven T cells (5, 37–39).

Single-cell RNA (scRNA) sequencing of TCRs and BCRs is a powerful tool for investigating the AIRR involved in AAA pathology. In comparison to bulk RNA sequencing, which yields a mixture of different gene expression profiles from the material studied, scRNA sequencing offers several advantages (40). scRNA sequencing provides information on TCR chain pairing and higher resolution and is more suitable for investigating the TCR specificity for an antigen of interest (41). However, there are also some limitations of scRNA sequencing. These challenges include the isolation of living single cells out of tissues, the lower output of sequenced cells compared to bulk sequencing, and higher costs (41). In particular, for scRNA sequencing of human AAA, only a small number of cells of interest is available for analysis, as only small sections of AAA can be collected during surgery (30). In mouse models, the whole AAA can be used, but the total amount of T and B cells is small for scRNA TCR and BCR sequencing. Zhao et al. (42) obtained approximately 3,000 cells, encompassing all present cell types, from topical elastase-induced AAA of 10 mice. Due to the small number of cells, biological interpretation of the sequenced TCR and BCR repertoire in AAA is challenging. In addition, standardized and uniform sample preparation, preprocessing of data, and bioinformatics workflow for data analysis are important to obtain robust and comparable data. There are already several guides (43), tools (44), and pipelines (45) for data analysis. However, in this paper, we highlight the limitations of scRNA TCR and BCR sequencing specifically in experimental AAA and provide a strategy for performing these experiments and for subsequent data analysis using a dataset we generated. The objective of this study is to guide and encourage fellow researchers to generate and evaluate scRNA TCR and BCR sequencing data, thereby enabling them to draw more significant conclusions.

Methods

Mice

Male C57BL/6J mice at the age of 10–11 weeks that were purchased from Janvier Labs (Saint-Berthevin, France) were used for experiments. All animal experiments were performed according to Animal Research: Reporting of In Vivo Experiments (ARRIVE) II guidelines and approved by LANUV (North Rhine-Westphalia State Agency for Nature, Environment and Consumer Protection) in accordance with the European Convention for the Protection of Vertebrate Animals used for Experimental and other Scientific Purposes (license approval number: 81-02.04.2018.A408). The mice were housed under standard laboratory conditions with a 12 h light/dark cycle and had *ad libitum* access to drinking water and standard chow.

PPE perfusion model

To induce AAA in the mice, the PPE perfusion model was used as previously described by Pyo et al. (46). Briefly, the mice received analgesics by injecting 0.1 mg/kg body weight (bw) buprenorphine subcutaneously prior to surgery. The mice were anesthetized with isoflurane (initial 3%, then 1.5%) and oxygenated air. After the absence of the toe reflex, laparotomy was performed, and the proximal and distal infrarenal aorta was isolated and temporarily ligated. The aorta was punctured, a catheter was inserted, and the infrarenal part was perfused with sterile isotonic saline containing type I PPE (2.5–3 U/ml #E1250 Sigma-Aldrich, Burlington, MA, USA) or 0.9% NaCl (sham surgery) under 120 mmHG for 5 min. Elastase concentrations ranged from 2.5 to 3 U/ml depending on the batch number, as different concentrations were necessary to trigger the same AAA incidence and size. The aortic puncture was sutured, the ligations were removed, and the abdomen was closed. Afterward, the mice received buprenorphine (0.1 mg/kg bw, subcutaneously) if required in the first eight hours. In addition, the mice received buprenorphine (0.01 mg/ml) via the drinking water for three days. The mice were monitored regularly until the end of the experiment.

Ultrasound imaging

Ultrasound was used to measure the aortic diameter prior to surgery and the AAA progression weekly. The Vevo 3100 high-resolution *in vivo* imaging system with a 25–55 MHz transducer (MX550D) (VisualSonics Inc., FUJIFILM, Toronto, Canada) was used for imaging. The mice were anesthetized with isoflurane and placed on a heated pad at 37°C. Aspiration rate,

electrocardiogram, and body temperature were monitored during the entire time of imaging. To assess the aortic diameter, longitudinal B-mode images of the infrarenal aorta were acquired. The aortic diameter was analyzed from leading to leading edge (LTL) in three cardiac cycles at end-diastole using the Vevo LAB 5.6.0 software.

Organ harvesting

The infrarenal aortae were harvested on days 3 ($n = 5$), 7 ($n = 5$), 14 ($n = 2$), and 28 ($n = 4$) after PPE surgery and on days 3 ($n = 3$) and 28 ($n = 3$) after sham surgery. In addition, the infrarenal aortae from non-treated C57BL/6J mice ($n = 3$) were pooled for two control samples. In total, we obtained 24 samples for scRNA sequencing. Approximately 10 min prior to organ harvesting, the mice were injected i.v. with 100 µl CD45-FITC antibody (#553079, BioLegend, dilution 1:1,000) to label circulating leukocytes. The mice were anesthetized and received analgesia with ketamine (100 mg/kg bw) and xylazine (10 mg/kg bw). After the absence of the toe reflex, blood was collected from the heart with a heparinized syringe. The thorax and abdomen were opened, the vena cava was cut, and the cardiovascular system was perfused with cold PBS through the left ventricle of the heart. The infrarenal part of the aorta was isolated by carefully removing all fatty tissue, collected, and stored in PBS on ice until further processing.

Digestion of aortic tissue into single cells

The isolated infrarenal aortae were digested into single cells based on the protocol from Hu et al. (47). Briefly, the aortae were cut and transferred into an enzyme mix containing 500 U/ml collagenase I (Sigma-Aldrich, #C0130-100MG), 120 U/ml collagenase XI (Sigma-Aldrich, #C7657-25MG), 60 U/ml hyaluronidase I-S (Sigma-Aldrich, #H3506-100MG), and 60 U/ml Dnase I (Sigma-Aldrich, #11285932001) in Dulbecco's phosphate buffered saline (DPBS) containing calcium and magnesium supplemented with 20 mM HEPES (Thermo Fisher Scientific, #15630106). The aortae were incubated in the enzyme mix for 50 min on a shaker (600 rpm) at 37°C. The cell suspension was filtered through a 100 µm cell strainer (pluriSelect Life Science, Leipzig). The remaining aortic tissue was mashed with a syringe plunger through the cell strainer, which was rinsed several times with DPBS. After centrifugation (10 min, 450 × *g*, 4°C), the cells were resuspended in cold PBS and transferred into a 96-well plate.

For additional flow cytometric analysis, the cells were subsequently resuspended in RPMI 1,640 (Sigma-Aldrich) supplemented with 10% fetal calf serum (Sigma-Aldrich) and incubated on a shaker (600 rpm, 12 min, 37°C). Finally, cells were centrifuged (10 min, 450 × *g*, 4°C), resuspended in PBS, and transferred into a 96-well plate.

Staining of single cells

The 96-well plate was centrifuged for 5 min at 500 × *g* and 4°C. The cells were stained with a staining mix containing Fc receptor blocker (TruStain FcX™, BioLegend, Amsterdam, Netherlands, 1:100), viability stain (Zombie Aqua™ and Zombie Green™ Fixable Viability Kit, BioLegend, 1:500), CD45-APC/cyanine 7 (BioLegend, clone 30-F11, dilution 1:200), TER-119-FITC (BioLegend, clone TER-119, 1:200), and C0443 CD41 (BioLegend, Barcode Sequence ACTTGGATGGACACT, 1:1,400) in PBS. In addition, an individual TotalSeq Hashtag antibody (BioLegend, TotalSeq™-C) was added to the single-cell suspension of each mouse. The hashtag antibodies allow the combination of samples from several mice in the same 10X sequencing run and are needed to demultiplex cells from individual mice. The cells from 10 mice were hashtagged with an antibody from the TotalSeq™-C series (BioLegend), respectively. Two additional hashtag antibodies were built by combining the antibodies MHC I-biotin (BioLegend, clone 28-8-6) and CD45-biotin (BioLegend, clone 30-F11) with the streptavidin-conjugated barcodes TotalSeq™ C971 or C972. The samples were stained with the staining mix and hashtag antibodies for 15 min at room temperature (RT) in the dark. After centrifugation (5 min, 500 × *g*, 4°C), the supernatant was discarded, and the cells were resuspended in MACS buffer (Miltenyi Biotec, #130-091-221) for following cell sorting.

For flow cytometric analysis, CountBright™ Absolute Counting Beads (Thermo Fisher Scientific) were added to every sample prior to staining to determine cell counts. The cells were centrifuged for 5 min at 500 × *g* and 4°C and stained with Fc receptor blocker (TruStain FcX™, BioLegend, 1:100) and viability stain (Zombie Aqua™ Fixable Viability Kit, BioLegend, 1:500) at RT for 10 min in the dark. After centrifugation (5 min, 500 × *g*, 4°C), the cells were stained with the following conjugated antibodies for 20 min at RT in the dark: CD3-APC (BioLegend, clone REA613, 1:200), CD11b-APC-Cy7 (BioLegend, clone M1/70, 1:200), CD19-FITC (BioLegend, clone MB19-1, 1:100), CD45-V450 (BioLegend, clone 30-F11, 1:200), Ly6C-PE (BioLegend, clone HK1.4, 1:133), Ly6G-PerCP-Cy5 (BioLegend, clone 1A8, 1:100), and NK1.1-PE-Cy7 (BioLegend, clone PK136, 1:200). The cells were centrifuged (5 min, 500 × *g*, 4°C) and resuspended in PBS with 0.5% bovine serum albumin. The samples were acquired with the BD FACSVerse™ Cell Analyzer (BD, Heidelberg, Germany), and data were analyzed with the FlowJo software v10.5.3.

Cell sorting

Cell sorting was performed on a MoFlo XDP (Beckman Coulter, Krefeld, Germany). For every sample, up to 3,000 living $CD45^+$ cells were sorted. If the samples did not reach the appropriate cell count, this was compensated by sorting more cells of other samples. The cells from all samples were combined into one reaction tube and centrifuged for 5 min at 500 × *g* at RT. The supernatant was removed, and the cells were resuspended in the MACS buffer. Recounting of the cells confirmed approximately 60,000 living cells.

Generation of single-cell library

Single-cell libraries were generated with the 10X Chromium Controller system utilizing the Chromium Next GEM Single Cell 5′ Kit v2 (10X Genomics, Pleasanton, CA, USA) according to

the instructions of the manufacturer. Sequencing was carried out on a NextSeq 550 system (Illumina Inc. San Diego, CA, USA) with a mean sequencing depth of ~50,000 reads/cell for gene expression. The T cell and B cell libraries and the hashtag libraries were sequenced at ~5,000 reads/cells.

Processing of 10X genomics single-cell data

Raw sequencing data was processed using the 10X Genomics CellRanger software (v6.0.2). Raw BCL files were demultiplexed and processed to Fastq files using the CellRanger *mkfastq* pipeline. Alignment of reads to the mm10 genome and the corresponding VDJ gene references and UMI counting was performed via the CellRanger *multi* pipeline to generate a gene–barcode matrix.

scRNA sequencing data analysis

R package *Seurat v4.0* (48) was used for analysis. First, the hashtag library was added as an assay to the metadata of the RNA library. The cells with less than 200 RNA counts and more than 30% mitochondrial RNA were excluded. Data were normalized and scaled with the functions *NormalizeData()* with the factor of 10.000, and *ScaleData()* was performed with all genes. Principal component analysis (75 dimensions), variable gene finding, cell clustering, and Uniform Manifold Approximation and Projection (UMAP) dimensional reduction (30 dimensions) were performed. Doublets were removed with *DoubletFinder v2.0* (49) and cells with positivity for more than one hashtag. T cells were defined by expression of *Cd3e*, *Cd3d*, *Cd3g*, and *Cd28*. B cells were defined by the expression of *Cd19*, *Cd79a*, and *Cd79b*. T and B cells were isolated bioinformatically and merged with the preprocessed scRNA TCR and BCR sequencing data.

scRNA TCR and BCR sequencing analysis

Preprocessing of the scRNA TCR and BCR sequencing data included quality control and adding the library of the hashtag antibodies. Only receptors with exactly one alpha/heavy chain and one beta/light chain were used for analysis. Sequences of single chains, more than two chains per receptor, and non-matching chains were excluded from the analysis. The hashtag information was merged with the 10× output file to enable the assignment of TCRs and BCRs to the different mice. The *Immunarch package v 0.6.9* (50) was used to analyze CDR3 length distribution, clone abundance, repertoire overlap, germline gene V-gene usage, clonal expansion, and diversity estimation. We defined a clone as a set of T or B cells expressing the same receptor that consists of the same V-, D-, and J-genes and encodes an identical CDR3 nucleotide sequence.

Statistics

Data are presented as absolute numbers or mean ± SD. Two-sample permutation-based Kolmogorov–Smirnov test was used to compare CDR3 length distributions with the function ks_test from R package "twosamples". Bonferroni correction was used for multiple comparisons. To compare the correlation strength of the V-gene usage in the TCR alpha chain with the V-gene usage in the TCR beta chain, we compared the individual correlation coefficients with a two-tailed Mann–Whitney *U* test. For comparison of our AAA data with public databases, one-sided Fisher's exact test with Bonferroni correction was performed using the R v4.0 software. A detailed description of the database comparison can be found in the **Supplementary Material**. The results with $p < 0.05$ were considered significant.

Results

scRNA sequencing workflow

The experimental workflow started with the induction of AAA by perfusing the infrarenal aorta of male C57BL/6J mice with elastase or NaCl (**Figure 1A**). AAA formation was monitored and analyzed via ultrasound imaging prior to surgery, at days 7, 14, and 28. The infrarenal aortae were harvested on days 3, 7, 14, and 28 after PPE surgery and on days 3 and 28 after sham surgery, and images were taken for macroscopic analysis. In addition, the infrarenal aortae from six non-treated C57BL/6J mice were harvested for sequencing, of which three were pooled into one sample. In total, 24 samples were subjected to scRNA sequencing and immunoreceptor analysis. All mice that underwent PPE surgery developed an AAA confirmed by ultrasound or macroscopic analysis of the infrarenal aorta (**Figure 1B**, **Supplementary Figure S1**). The aortae were harvested and digested into single-cell suspensions. Single cells were stained and sorted and then subjected to scRNA sequencing as well as TCR and BCR scRNA sequencing.

Consistent preprocessing of immune receptor scRNA sequencing data is crucial for comparable data. We suggest using only those immune receptor data, with all fragments intact and both chains (α/β for TCR or heavy/light for BCR) present. Receptors with only one, or more than two chains, that can appear in the data frame due to sequencing errors, were excluded. After that, the information of the hashtag antibodies was added to the data to assign an immune receptor specifically to one cell of a specific mouse. Subsequently, the receptors were assigned to the corresponding cell in our scRNA sequencing dataset. The majority of the immune receptor analysis such as clonality, repertoire overlap, V-gene usage, and diversity was performed with the R package *immunarch* (50). In addition, a comparison of the AAA data with public databases to identify disease-associated receptors was performed (**Figure 1C**).

Analysis of scRNA TCR and BCR sequencing data

To ensure adequate data quality, we processed the TCR and BCR sequencing data and evaluated basic statistics. The raw data

10.3389/fcvm.2023.1221620

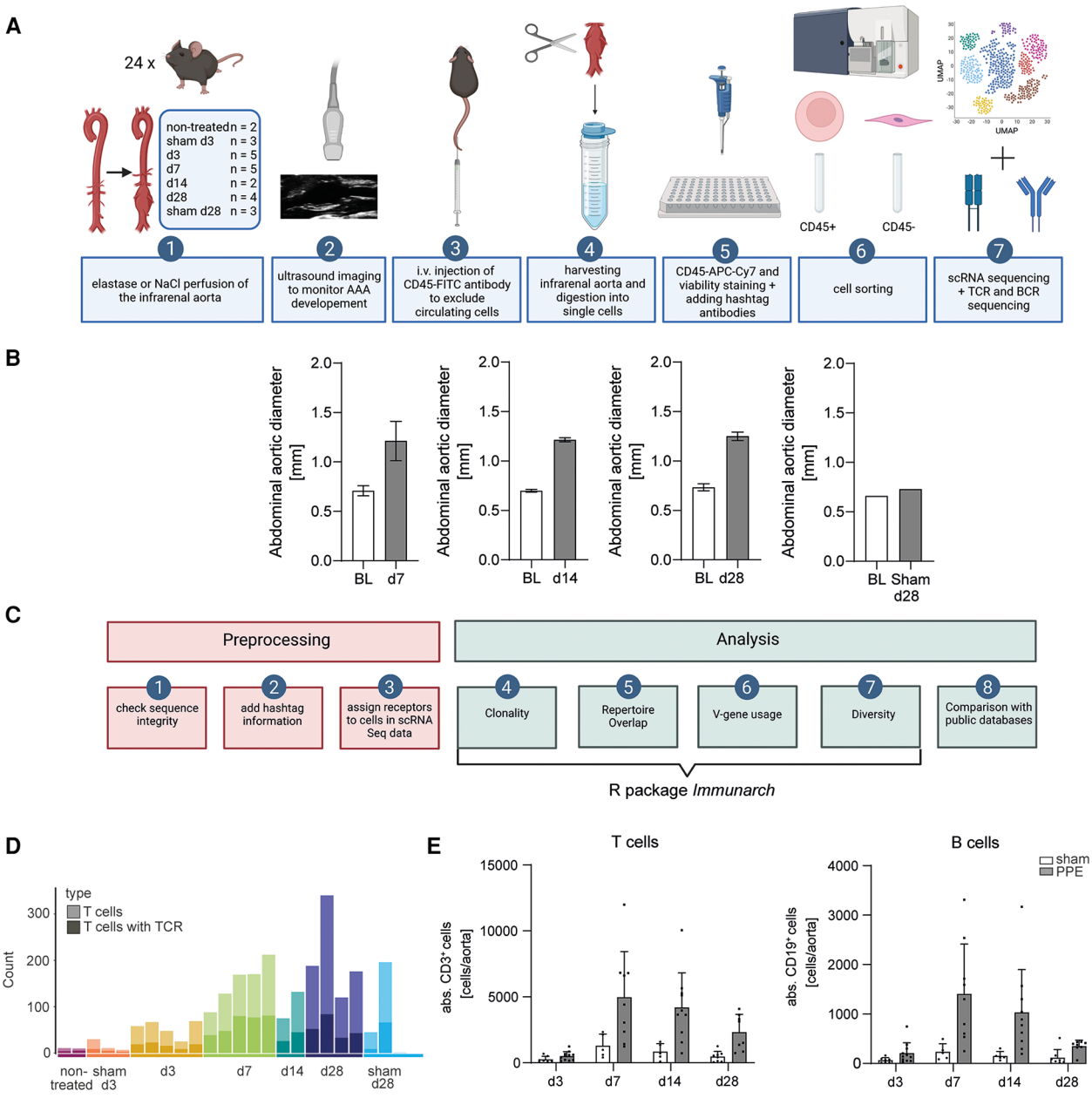


FIGURE 1
Experimental and bioinformatic workflow and quality control of the data. (**A**) Experimental workflow. An AAA was induced in mice via perfusion of the infrarenal aorta with elastase. NaCl perfusion of the aorta served as a sham-operated control. AAA development was monitored by ultrasound imaging. Prior to organ harvesting, the mice were intravenously injected with a fluorophore-labeled CD45 antibody. The harvested infrarenal aortae were digested into single cells and stained with antibodies and hashtags. The cells were sorted for living leukocytes and living non-leukocytes. scRNA sequencing and TCR and BCR scRNA sequencing were performed. (**B**) The abdominal aortic diameter was analyzed using ultrasound images of baseline (BL) and days 7 ($n = 5$), 14 ($n = 2$), and 28 ($n = 2$) post-PPE surgery and BL and day 28 ($n = 1$) post-sham surgery. (**C**) Bioinformatic workflow including preprocessing and data analysis. Three preprocessing steps were performed as quality control for the data. Only fully intact sequences were retained for analysis. The information for the hashtag antibodies was added to the TCR and BCR scRNA sequencing data. The receptors were assigned to the corresponding cells in the scRNA sequencing dataset. The main data analysis including clonality, repertoire overlap, V-gene usage, and diversity was performed with the immunarch R package. In addition, the data were compared with the public databases. (**D**) T cell amounts (light bar color) and T cell amounts exhibit an intact TCR (dark bar color) in aortic tissue on different AAA disease stages received from scRNA sequencing data. The amount of T cells increases with AAA progression. (**E**) Flow cytometric analysis of T and B cell amounts of aortic tissue 3 ($n = 7$), 7 ($n = 5$), 14 ($n = 5$), and 28 ($n = 8$) days after sham operation and 3 ($n = 11$), 7 ($n = 9$), 14 ($n = 10$), and 28 ($n = 8$) days after PPE-induced AAA formation.

contained 3,370 TCR sequences [1,484 TCR alpha chains (TRA), 1,886 TCR beta chains (TRB)] and 1,745 BCR sequences [570 Ig heavy chains (IgH), 1,131 Ig kappa (IgK) light chains, 44 Ig lambda (IgL) light chains]. We observed a high number of immune receptors with only one sequenced chain and some receptors with more than two chains, which were excluded from subsequent analysis (**Supplementary Figure S2**). Only T cells with both productive TRA and TRB chains and B cells with both

productive heavy and light chains were used for analysis. After that step, 2,296 TCR chains (1,148 pairs of TRA and TRB chains) and 980 BCR chains (490 pairs of heavy and light chains) remained. We next filtered for receptors that could be associated with a hashtagged cell and retained 2012 TCR chains and 770 BCR chains. Assignment of the immune receptors to the corresponding cells in our scRNA sequencing data revealed that only 47% of BCRs were expressed in B cells (defined by mRNA expression of *Cd19*, *Cd79a*, *Cd79b*), whereas the majority of TCRs (79%) was expressed in T cells (defined by mRNA expression of *Cd3e*, *Cd3d*, *Cd3g*, *Cd28*), and the remaining immune receptors were found on other cell types (**Supplementary Figure S3**). TCRs and BCRs not expressed in the respective lineage were excluded to avoid analysis of false positive receptors due to sequencing artifacts. The final analysis included 1,592 TCR chains (796 pairs of TCRs) and 358 BCR chains (179 pairs of BCRs).

We next compared the number of immune receptors with that of T and B cells, which were present in our scRNA sequencing dataset and displayed the distribution across the time points and samples (**Figure 1D**, **Supplementary Figure S4** and **Supplementary Table S2**). Overall, there were fewer B cells (325) than T cells (2,376) in AAA tissue, and three out of 24 samples did not contain B cells (non-treated, d3, sham d28) (**Supplementary Figure S4** and **Supplementary Table S2**). The total number of T and B cells increased with AAA progression until day 7 (**Figure 1D**, **Supplementary Figure S4**). We corroborated our findings by flow cytometry revealing a peak of lymphocytes at day 7 in AAA. Of note, only a few cells were detected in sham-operated mice (**Figure 1E**). A fully productive TCR could be assigned to 33.5% of the present T cells (2,376 T cells, 796 TCRs) (**Figure 1D**). Thus, a large proportion of TCR sequences present in AAA were missing due to inefficient sequencing. In our data, 55.1% of B cells had a matching BCR (325 B cells, 179 BCR). In four out of 24 samples, no BCRs could be detected, and these were sham-operated or early time points (non-treated, sham d3, d3, sham d28) (**Supplementary Figure S4B** and **Supplementary Table S2**).

Estimating clonal expansion by spectratyping

Spectratyping identifies the pattern of the CDR3 length distribution (51). Comparing the shape of the CDR3 length distribution between control and disease can indicate the presence of clonal expansion in a repertoire. Deviations from the normal pattern might be due to the high frequency of a specific CDR3 sequence and are therefore associated with clonal expansion (31). We compared the CDR3 length distribution of TCRs in AAA across all time points and controls using the two-sample permutation-based Kolmogorov–Smirnov test (**Figure 2A**). The resulting *p*-value of 0.92 suggests that the CDR3 length followed the same distribution in AAA and controls. To compare the CDR3 length distribution between the different disease stages, we performed pairwise two-sample permutation-based Kolmogorov–Smirnov tests and used Bonferroni correction for multiple comparisons. We did not observe significant differences between the CDR3 length distribution at the different time points (**Figure 2B**).

Receptor clonality can also be investigated by determining the number of unique clones and the clone abundance. A clone was defined as a set of cells expressing the same receptor that consists of the same V-, D-, and J-genes and encodes an identical CDR3 nucleotide sequence. The majority of TCRs in AAA, sham-operated, and non-treated aortae were unique (**Figure 2C**, **Supplementary Table S3**). Only one sample of the sham-operated and non-treated aortae and 11 AAA samples contained T cell clones (**Figure 2C**, **Supplementary Table S3**). However, the clones were infrequent in AAA samples (occurring 2–5 times), whereas one of the four T cell clones present encompassed 25 cells in one sham-operated sample (**Figure 2D**, **Supplementary Table S3**). The frequencies of the T cell clones, their V-, D-, and J-genes of alpha and beta chain, and their CDR3 nucleotide and amino acid sequences are shown in **Supplementary Table S4**. The extent of receptor clonality can be indicated with an evenness profile of the repertoire (27). The alpha values represent different diversity indices with different weights on expanded clones. Higher alpha values give more weight to expanded clones, while alpha = 0 weights every clone equally regardless of its frequency. Therefore, high receptor clonality is indicated as a highly uneven curve, and no receptor clonality is associated with a completely even curve (**Figure 2E**). In our study, the sham-operated d28 sample, in which one clone was identified 25 times (**Figure 2D**), exhibited also the highest clonality, whereas all other sham-operated or non-treated samples showed no clonality (**Figure 2E**). In addition, 11 AAA samples from different time points showed a lower extent of receptor clonality.

Investigating the TCR repertoire similarity

Repertoire overlap analysis is commonly used to identify "public" TCRs that are shared between individuals (52). The R package *immunarch* provides several methods to measure receptor similarities between individuals. Using the function "public" specified, the exact number of shared immune receptors between different repertoires, thereby revealing that in seven instances, a TCR sequence was shared by two AAA samples (**Figure 3A,B**). In addition, repertoire similarity can be investigated by identifying TCRs of different individuals containing the same V region genes. Fragments of V region genes are classified into families according to their nucleotide sequence similarity (at least ~70%). Specific V-gene usage patterns have been associated with different diseases and were shown to change in response to therapeutic approaches (53, 54). In our data, the V-gene usage of the beta chain (TRBV) correlated stronger than the V-gene usage of the alpha chain (TRAV) of the TCR (two-tailed Mann–Whitney *U* test, $p = <0.0001$, **Figures 3C,D**). A deeper analysis of the distribution and frequency of used TRBV genes revealed a high usage of TRBV3, TRBV19, and TRBV12-2 + TRBV13-2 in AAA samples at days 7, 14, and 28 (**Figure 3E**). TRBV12-2 + TRBV13-2 is a term for a common alternate splicing between the first exon of TRBV12-2 and the second exon of TRBV13-2. TRBV19 was

10.3389/fcvm.2023.1221620

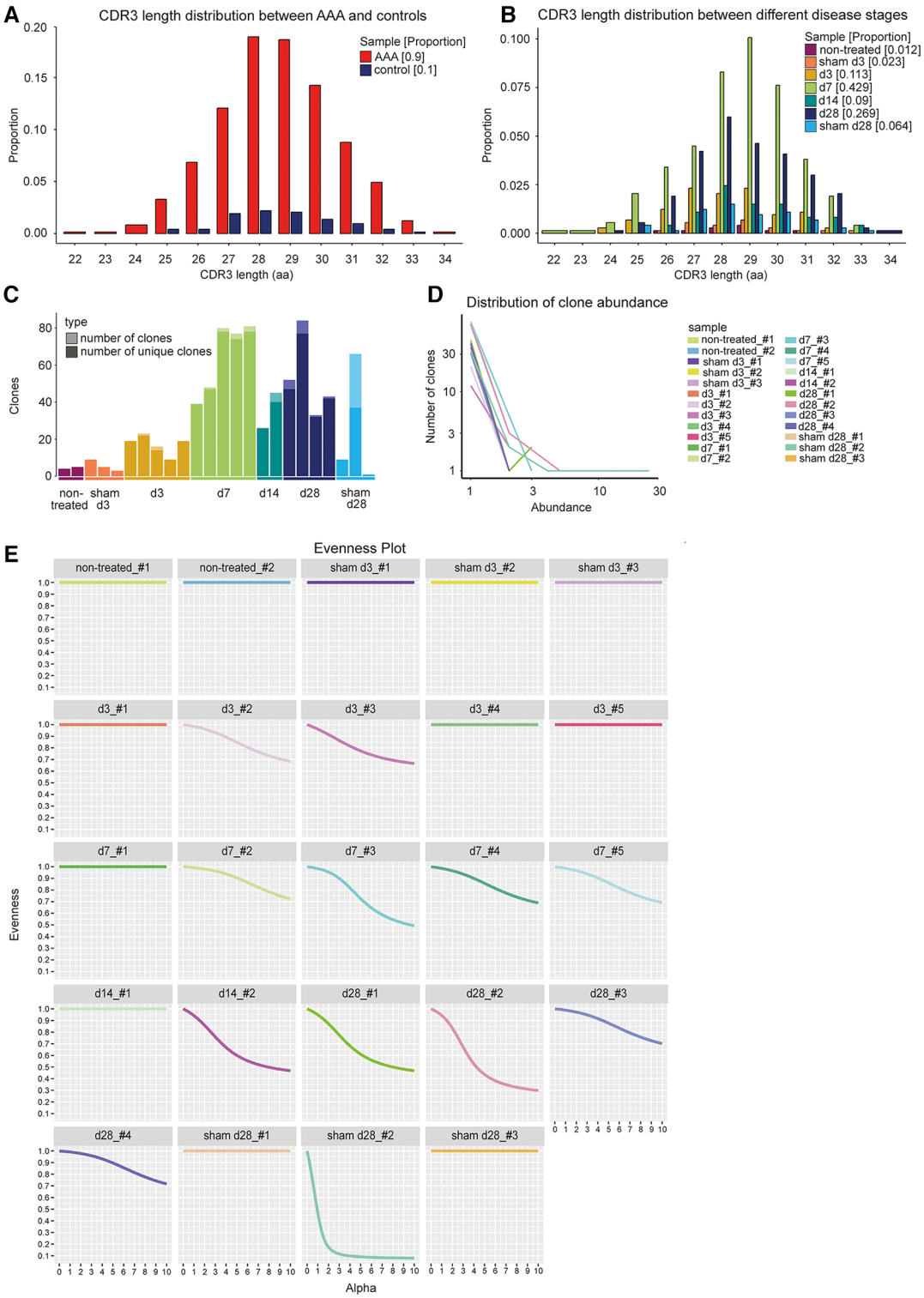


FIGURE 2
CDR3 length distribution and clone abundance indicate expanded T cell clones in elastase-induced aneurysm in mice. (**A**) No changes in CDR3 length distribution of TCRs (paired chains) between AAA including all different disease stages (red) and control samples including sham-operated and non-treated mice (blue). The proportion is plotted against the amino acid (aa) CDR3 length. (**B**) No alterations in CDR3 length distribution of TCRs (paired chains) between samples of different disease stages (days 3, 7, 14, 28) and sham-operated and non-treated samples (different colors). The proportion is plotted against the amino acid (aa) CDR3 length. (**C**) Amount of all TCR clones per sample (light bar color) including the amount of unique clones (dark bar color). The majority of TCR clones were found to be unique. Multiple copies of one clone appear only in one of the sham_d28 samples and in 11 of the AAA samples. (**D**) Line plot indicating the number and abundance of clones per sample. In the sham_d28 sample exhibiting clones with multiple copies, one clone is present 25 times. In contrast, in AAA samples, one clone only appears 2–5 times (**E**) Evenness plots indicating the extent of clonal expansion for every sample. One sham-operated sample 28 days after perfusion exhibits the highest clonality. The other control samples show no receptor clonality. Eleven AAA samples show some clonal expansion.

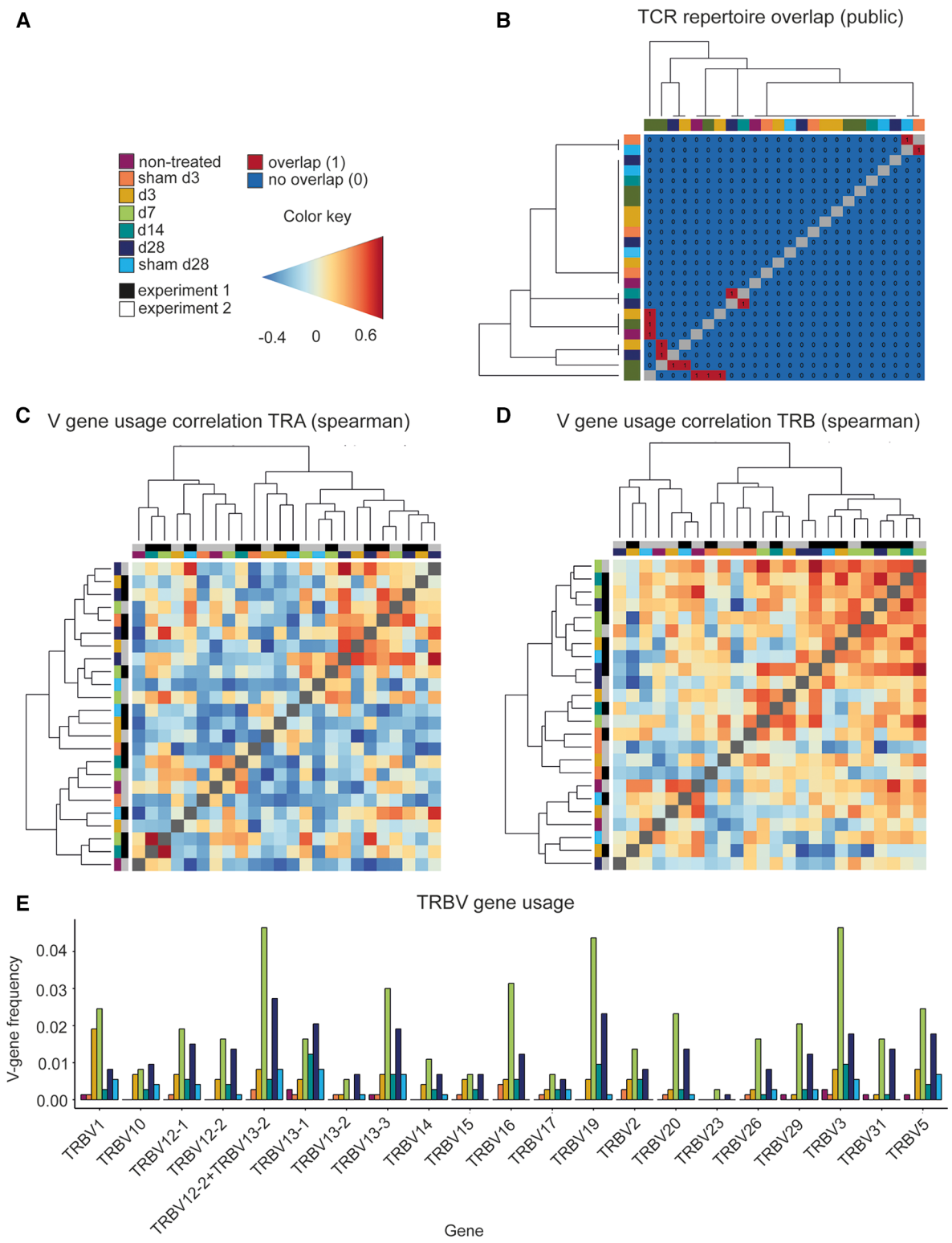


FIGURE 3
Comparing the different TCR repertoires reveals shared TCR sequences, a high correlation of the V-gene usage and several frequently used TRBV genes in AAA. (**A**) Legend and color code. (**B**) Analysis of the TCR repertoire overlap shows that in 7 instances two AAA samples contain one equal TCR. The color code in the heatmap indicates the different samples. The V-gene usage of the TCR alpha chain (**C**) and of the TCR beta chain (**D**) is correlated and hierarchical clustered between the different samples using spearman correlations. Color gradient indicates the level of correlation (blue = negative correlation, red = positive correlation). The color code on the axes indicates the different samples. (**E**) Distribution and frequency of TRBV genes occurring in all samples. Frequently used TRBV genes in AAA samples are TRBV3, TRBV19, and TRBV12-2+TRBV13-2 at day 7, 14 and 28. TRA, TCR alpha chain; TRAV, TCR alpha chain v gene; TRB, TCR beta chain; TRBV, TCR beta chain v gene.

present in 5 and TRBV3 in 2 of the expanded clones. Further Vbeta genes that were used by 2–3 of the expanded TCRs are TRBV10, TRBV13-1, TRBV13-5, TRBV2, TRBV20, and TRBV29 (**Supplementary Table S4**).

Dataset comparison with public TCR and BCR databases

We compared the presence of CDR3 sequences in our dataset with the two public TCR databases VDJdb (55) and McPAS-TCR (56) to investigate if TCR clones in our dataset are associated with other diseases or antigens (57). VDJdb is a curated database of TCR sequences with known antigen specificities containing TCR information of three different species (*Homo sapiens*, *Macaca mulatta*, and *Mus musculus*) and various diseases (55). McPAS-TCR is a database of TCR sequences found in T cells that were associated with various pathological conditions in humans and mice (56). TCRs with less than four amino acids and diseases with less than five TCRs were excluded. Accordingly, 5,206 TCRs were found in influenza (3,156 TCRs), lymphocytic choriomeningitis virus (LCMV) (151 TCRs), murine cytomegalovirus (MCMV) (1,463 TCRs), *Plasmodium berghei* (245 TCRs), respiratory syncytial virus (RSV) (125 TCRs), and vesicular stomatitis virus (VSV) (66 TCRs) for VDJdb, and 3,530 TCRs that were assigned to 21 different diseases/pathogens were used for analysis for McPAS-TCR (**Supplementary Table S5**). After merging, the two databases and filtering for unique CDR3 sequences, we obtained 4,331 CDR3 sequences for comparison with the AAA dataset. One-sided Fisher's exact test was used to examine the overrepresentation of TCR clones in our dataset that are associated with diseases or antigens according to the two databases (**Figure 4A**). Our dataset shared 55 CDR3 sequences with the public databases, which were assigned to MCMV, LCMV, influenza, *Plasmodium berghei*, VSV, diabetes type 1, and tumor. The obtained *p*-values were adjusted for multiple testing using Bonferroni correction. Bonferroni correction resulted in no significant *p*-values indicating there were no TCR clones overrepresented in our dataset that are associated with diseases or antigens according to the two databases (**Figure 4B**, **Supplementary Table S6**).

No clonal expansion or repertoire overlap of BCRs in elastase-induced aneurysm in mice

The CDR3 length distribution of the BCRs showed no differences between AAA and control at the disease stages (**Figure 5A**,**B**, pairwise two-sample permutation-based Kolmogorov–Smirnov test resulted in no significant differences). Furthermore, 98% of BCR clones were unique (176 of 179 clones were unique). As mentioned before, BCRs were present in only 20 of 24 samples in our dataset. The samples that lacked BCRs were control samples or early disease stages that are known to contain few B cells overall (non-treated, sham d3, sham d28, d3). BCRs that appear more than once were found in two of these 20 samples. One day 28 sample contained one BCR that was present thrice, and one day 7 sample had one BCR that appeared twice (**Figure 5C**). The evenness profile likewise indicated a higher clonality for these two AAA samples in comparison to all other samples that showed no clonality (**Figure 5D**, **Supplementary Figure S5**). Next, we investigated the isotype distribution of the BCR heavy chains. The most frequent isotype was IGHM, followed by IGHD (**Figure 5E**). The similarity measurement of the BCR repertoire present in the different samples showed that two AAA samples (days 7 and 28) share one BCR (**Figure 5F**). Otherwise, there was no similarity between the different samples. The correlation of V-gene usage between the different samples was likewise low (**Figure 5G**). The strongest, yet

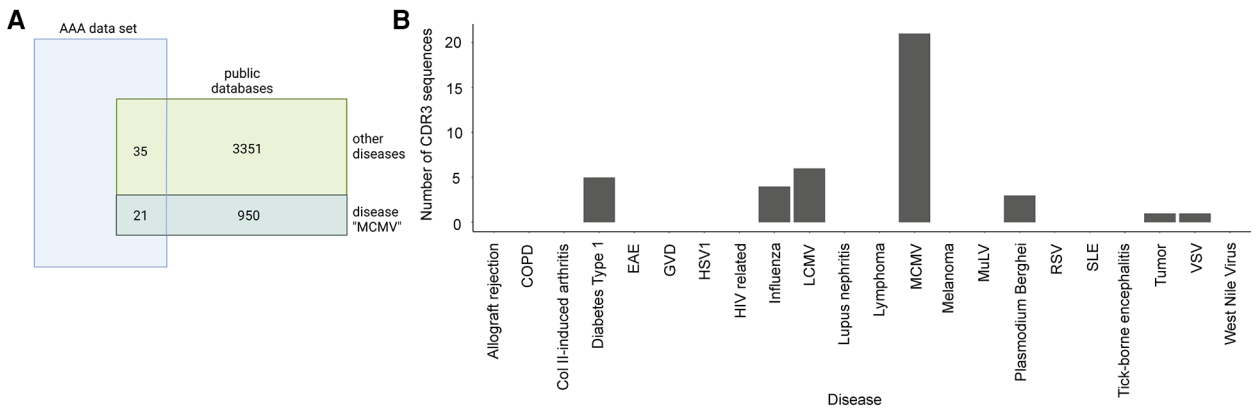


FIGURE 4
Overlap of AAA-associated CDR3 sequences with public databases. (A) Venn diagram indicating the overlap of our TCR data with the public TCR databases (VDJdb and McPAS-TCR) representative for the disease MCMV. Our data shares 21 CDR3 sequences with the public databases that are associated with MCMV and 35 CDR3 sequences that are associated with other diseases. One-sided Fisher's exact test with Bonferroni correction for multiple comparisons was used to examine the overrepresentation of TCR clones in our dataset that are associated with diseases or antigens. (B) Barplot displaying the number of sequences per disease in our data. COPD, chronic obstructive pulmonary disease; Col II, collagen II; EAE, experimental autoimmune encephalomyelitis; GVD, graft vs. host disease; HSV1, herpes simplex virus 1; HIV, human immunodeficiency virus; LCMV, lymphocytic choriomeningitis virus; MCMV, murine cytomegalovirus; MuLV, murine leukemia virus; RSV, respiratory syncytial virus; SLE, systemic lupus erythematosus; VSV, vesicular stomatitis virus.

10.3389/fcvm.2023.1221620

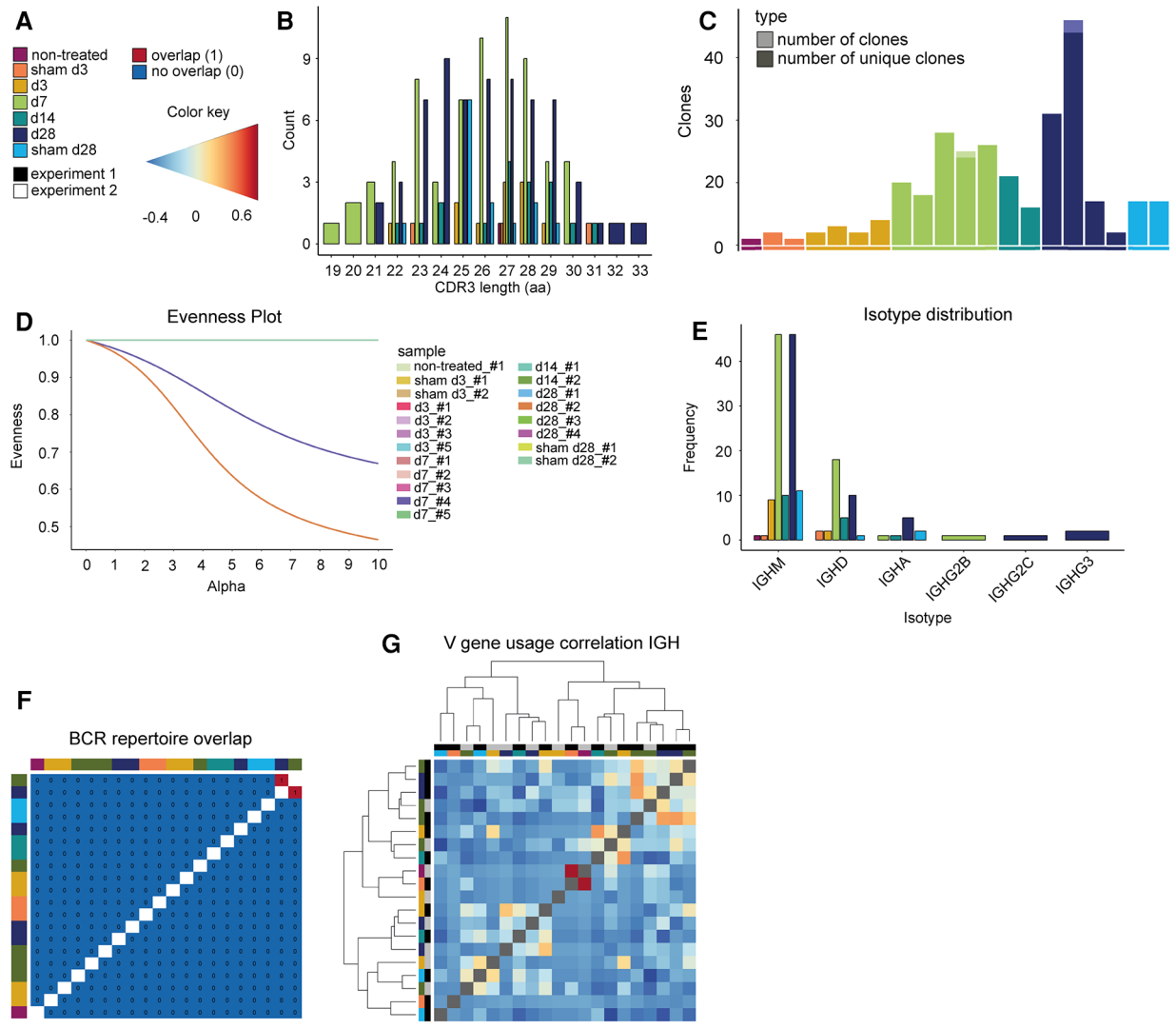


FIGURE 5
Elastase-induced AAA shows no BCR clonality. (**A**) Legend and color code. (**B**) Histogram indicating the CDR3 length distribution of the BCRs (paired chains). The proportion is plotted against the amino acid (aa) CDR3 length. The color indicates whether the TCRs belong to an aorta isolated at days 3, 7, 14, or 28 after PPE or sham surgery or to a non-treated aorta. (**C**) Barplot displaying the number of all BCR clones (pale color) and only unique clones (vibrant color) per mouse. Almost all BCR clones are unique. (**D**) Evenness plot indicating the extent of clonal expansion for every sample. (**E**) Barplot indicating the isotype distribution of the BCR heavy chain (IGH) in the different conditions. The most frequent isotypes for all conditions are IGM and IGD. (**F**) BCR repertoire overlap is displayed in a heatmap. The color bars on the top and left side of the heatmap indicate the time point. (**G**) Heatmap presenting the Spearman correlation and hierarchical clustering of the BCR heavy chain (IGH) V-gene usage of the different samples. The color gradient indicates the level of correlation (blue = weak correlation, red = strong correlation). The color bars on the top and left side of the heatmap indicate the experiment and the time point.

still weak correlation, was found between one non-treated control and one d3 sample ($r = 0.701$). Overall, these data revealed no evidence for clonality among B cells in AAA.

Discussion

We assessed TCR and BCR clonality in elastase-induced AAA in mice at different disease stages using scRNA TCR and scRNA BCR sequencing. Our results show no differences in CDR3 length distribution of TCRs and BCRs between the different disease stages, indicating no strong clonal expansion of immune cell receptors in elastase-induced AAA. The clone abundance analysis likewise revealed no clonal expansion of BCRs in AAA. We found expanded T cell clones in 68% of AAA samples and no clonality in control samples except for one. A comparison of the immune receptor repertoires showed a low similarity between the individual samples. Spearman correlation to compare the V-gene usage between the different AAA samples and controls revealed that the V-gene usage of the TCR beta chain correlates stronger than the V-gene usage of the TCR alpha chain. The most frequently used V-genes in the TCR beta chain in AAA are TRBV3, TRBV19, and TRBV12-2 + TRBV13-2. A comparison of TCR clones identified by us revealed no overrepresentation of TCR clones associated with diseases or antigens annotated in two public databases. The main Ig isotype in our BCR dataset is IgM followed by IgD. Although this may prompt

speculation of enrichment of B1 cells that predominantly express IgM, the overall scRNA-sequencing dataset suggests that B2 cells are approximately 20-fold more frequent than B1 cells. Notably, this corroborates previous reports of B2 dominating the B cell pool in mouse AAA (11).

Clonal expansion, public TCRs, and convergent T cells

Antigen recognition by immune cell receptors activates naive lymphocytes prompting them to proliferate. This process is termed clonal expansion and enables a targeted, adaptive immune response. However, the term clonal expansion remains strongly debated as there is no clear and consensus definition. Lu et al. investigated T cell clonality in aneurysmal lesions of AAA patients and defined clonal expansion as the presence of multiple identical copies of TCR transcripts (37). They reasoned that the size of the T cell repertoire makes it unlikely that multiple identical copies of a TCR transcript would be found by chance in an independent sample of T cells (37). According to this definition, we found clonal expansion in 11 of 16 AAA samples and one of eight control samples.

Public TCRs are shared across different individuals due to VDJ recombination biases and might target common antigens (52). The repertoire similarity of the AAA samples was low. In seven instances, a TCR sequence and only one BCR sequence were shared by two individual AAA samples. Next to clonal expansion and public TCRs, there is also T cell convergence. Convergent T cells are cells expressing TCRs with identical CDR3 amino acid sequences and variable genes but different CDR3 nucleotide sequences (58). Convergent T cells arise due to codon degeneracy and can be observed in almost every individual. Pan and Li (58) showed that convergent T cells are different from public TCRs and seem to be antigen-specific. According to their results, TCR convergence might be a better indicator of antigen specificity than clonal expansion. Since convergent T cells constitute only a small proportion of the total population of T cells, studies of TCR convergence require a large number of sequenced T cells. We did not find convergent T cells in our dataset, probably due to the small number of T cells but expect that investigating T cell convergence in larger datasets may be a feasible and worthwhile approach to address this important component of an antigen-specific T cell response.

Clonal expansion in human AAA and atherosclerosis

Studies with patients demonstrated the presence of clonally expanded TCRs in AAA or ascending thoracic aortic aneurysms (TAA), supporting the paradigm of AAA as a disease driven by an antigen-specific T cell response (5, 37–39). In particular, clonal expansion of TCR beta (5, 37) and alpha (38) chains was demonstrated in AAA lesions of patients while others reported clonal expansion of γ/δ T cells in AAA (5). Furthermore, TCRs were investigated in different types of TAA (patients with Marfan syndrome, familial TAA, and sporadic aneurysm), and the results indicate a similar clonal nature of the TCRs present in TAA (39). He et al. (39) found a preferential usage of the V-genes Vb22 and Vb25 in lesions from patients with TAA. Lu et al. (37) reported multiple appearances (at least twice) of TRBV3 in 60% of AAA patients. Atherosclerotic vascular disease, which is also a risk factor for AAA development while it also shares some common (immuno-) pathophysiological pathways (59), is likewise associated with T cell expansion and clonality. In particular, TCRs containing Vβ6 are expanded in atherosclerotic lesions of mice (60). Moreover, a decreased diversity of the TCR β chain repertoire was shown in human atherosclerotic plaques due to the expansion of a few T cell subclones (61).

Limitations of the current study

The main limitation of the current study is the relatively small number of lymphocytes resulting in potential undersampling. The limited number of T and B cells resulted from the naturally scarce source (i.e., minimal aneurysm size in mice), from additional sorting procedures (sorting of all leukocytes and not specifically T and B cells), and not fully efficient sequencing. Indeed, we had to exclude many TCR and BCR sequences from the data due to inefficient sequencing. The undersampling leads to the issue that the TCR and BCR copies in our data do not represent the real absolute number of copies present in AAA and even the ratio of the various clones to each other does not reflect the real ratio (62). Accordingly, this data should be interpreted cautiously and presents restricted value for biological interpretation. Further experiments are needed to verify the evidence of T cell clonality in elastase-induced AAA. Until now, mouse models have been standard for studying mechanisms of human pathophysiology. However, there are considerable differences between species regarding genetics, physiology, and immunology, which have to be considered. Although the PPE model is the mouse model most closely resembling human AAA, it does not fully mimic the complexity of AAA development in humans (15, 63). Human AAA features a complex and long-lasting disease development that is only partially resembled in experimental rodent models that aim to recapitulate disease patterns in a few weeks. Thus, effects observed in mice have to be extrapolated with caution to human aneurysmal and atherosclerotic disease.

Future perspectives and recommendations

To overcome the problem of undersampling, we suggest sorting at least 5,000–10,000 B and T cells instead of including all leukocytes. We detected by flow cytometry on average 120–430 B cells and 400–1,600 T cells per mg AAA tissue depending on the stage of AAA development, while lymphocyte numbers are lower in control conditions (e.g., native or -sham-operated mice; on average 120–300 B cells and 400–500 T cells/mg aortic tissue). Thus, pooling of aneurysms from several mice is necessary to obtain a sufficient number of lymphocytes. In this case, we recommend the use of Hashtag antibodies before pooling to enable the assignment of the lymphocytes to the corresponding mouse and to monitor clonality

for each individual. Next, the choice of experimental model should be carefully considered, as each model has limitations and mimics specific features of human AAA (15). Li et al. (36) induced AAA in mice with elastase and $CaPO_4$, performed scRNA sequencing combined with TCR sequencing of 41,341 CD4+ T cells isolated from AAA, and found a clonal expansion of Treg. This suggests that the number of analyzed cells is an important factor for investigating TCR clonality and a high number of cells facilitates the identification of clonal expansion. Higher cell counts increase the likelihood of detecting rare TCR and BCR clones and allow additional investigation of T cell convergence. We would like to recommend prioritized sequencing of the TCR and BCR libraries and to include DNA-barcoded antibodies against B cells (CD19) and T cells (CD3), which allows for superior identification of subpopulations in comparison to identification by mRNA expression of feature genes. In addition to analyzing the aneurysmatic tissue itself, future studies may also include sequencing of secondary lymphoid organs (e.g., draining lymph nodes) to identify changes in lymphocyte clonality, migration, and activity. scRNA TCR and BCR sequencing has the advantage of a high-throughput, multi-parametric analysis of target cells. Drop-sequencing approaches, such as the commercially available 10X Genomics solution used here, allow us to interrogate the transcriptome, TCR, and BCR of 1,000 cells simultaneously. However, there is a high dropout in detecting lowly expressed genes, which bears limitations: (1) this technology is particularly advantageous in describing a diverse cell population, while other approaches might be superior in studying transcriptional changes in related subpopulations, (2) full-length transcripts of TCRs and BCRs for both chains might not be detectable in all cells (we identified full-length TCRs and BCRs in 33.5% and 55.1% of all cells). To uncover detailed transcriptional changes in T cell subpopulations, this approach could be complemented by sorting these cells and performing bulk transcriptomics, which delivers a deeper insight. In addition, beta repertoire sequencing can be used to approximate T cell clonality on a global level and confirm observations made by scRNA TCR sequencing. However, this method does not provide information about the transcriptome of an individual cell or the corresponding paired TCR alpha chain, thus not reflecting the true complex clonality.

For the comparison of datasets from different research groups, it is important to have standardized workflows for sample preparation and the preprocessing of the data. Accordingly, we present an example and detailed workflow for the preprocessing steps. We recommend including only immune receptors with two productive chains that can be assigned to a B or T cell. In addition to clonal expansion and repertoire similarity, future analyses should also address T cell convergence.

Conclusion

In conclusion, we found evidence of clonal expansion of T cells, but not of B cells, in experimental elastase-induced AAA. Due to the small number of cells further experiments are needed to verify the evidence of T cell clonality. Since other studies found TCR clonality in AAA lesions of patients and considering the paradigm of an autoimmune response in aneurysmal disease, further examination of TCR and BCR clonality is important. Our findings imply that a precise characterization of TCR and BCR distribution requires a more extensive number of lymphocytes to prevent undersampling and to allow for the detection of rare clones and convergent T cells. This paper provides an in-depth analysis of TCR and BCR sequencing data, emphasizes the potential drawbacks and constraints of these experiments, and offers recommendations for future investigations in this area.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found here: https://doi.org/10.5281/zenodo.7942455.

Ethics statement

The animal study was approved by LANUV (North Rhine-Westphalia State Agency for Nature, Environment and Consumer Protection). The study was conducted in accordance with the local legislation and institutional requirements.

Author contributions

NG and AL designed the study. CE, MO-B, MF, AL, SV, and SP performed the experiments. VG designed and supervised the data analysis. MC designed Fisher's exact test and Kolmogorov–Smirnov test analysis. CE analyzed the data and wrote the manuscript. KL helped with bioinformatic analysis. VG and NG supervised the manuscript write-up. TV and MB designed the figures. MK, HW, SS, and TP critically reviewed the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This study was supported by the following grants: Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)—grant no. 397484323—CRC/TRR259; project A04 to HW and project A05 to NG; MODS project funded from the program "Profilbildung 2020" (grant no. PROFILNRW-2020-107-A), an initiative of the Ministry of Culture and Science of the State of North Rhine Westphalia; Research Commission of the Medical Faculty of Heinrich-Heine University to AL (grant no. 2021-10). We acknowledge the support of the Susanne-Bunnenberg-Stiftung at the Düsseldorf Heart Center.

Acknowledgments

We would like to acknowledge Julia Odendahl and Joscha Mulorz for performing surgeries on the mice and the

assistance from Katarina Raba at the Core Flow Cytometry Facility at the Institute for Transplantation Diagnostics and Cell Therapeutics Düsseldorf. We thank Tobias Lautwein for performing the single-cell sequencing and analyzing scRNA-seq primary data. Computational infrastructure and support were provided by the Centre for Information and Media Technology at Heinrich Heine University Düsseldorf. Figures were created with Biorender.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcvm.2023.1221620/full#supplementary-material

References

1. Sakalihasan N, Michel J-B, Katsargyris A, Kuivaniemi H, Defraigne J-O, Nchimi A, et al. Abdominal aortic aneurysms. *Nat Rev Dis Primers*. (2018) 4(1):34. doi: 10.1038/s41572-018-0030-7

2. Yuan Z, Lu Y, Wei J, Wu J, Yang J, Cai Z. Abdominal aortic aneurysm: roles of inflammatory cells. *Front Immunol*. (2020) 11:609161. doi: 10.3389/fimmu.2020.609161

3. Kuivaniemi H, Platsoucas CD, Tilson MD 3rd. Aortic aneurysms: an immune disease with a strong genetic component. *Circulation*. (2008) 117(2):242–52. doi: 10.1161/CIRCULATIONAHA.107.690982

4. Lu S, White JV, Nwaneshiudu I, Nwaneshiudu A, Monos DS, Solomides CC, et al. Human abdominal aortic aneurysm (AAA): evidence for an autoimmune antigen-driven disease. *Autoimmun Rev*. (2022) 21(10):103164. doi: 10.1016/j.autrev.2022.103164

5. Platsoucas CD, Lu S, Nwaneshiudu I, Solomides C, Agelan A, Ntaoula N, et al. Abdominal aortic aneurysm is a specific antigen-driven T cell disease. *Ann N Y Acad Sci*. (2006) 1085:224–35. doi: 10.1196/annals.1383.019

6. Piacentini L, Werba JP, Bono E, Saccu C, Tremoli E, Spirito R, et al. Genome-wide expression profiling unveils autoimmune response signatures in the perivascular adipose tissue of abdominal aortic aneurysm. *Arterioscler Thromb Vasc Biol*. (2019) 39(2):237–49. doi: 10.1161/ATVBAHA.118.311803

7. Brophy CM, Reilly JM, Smith GJ, Tilson MD. The role of inflammation in nonspecific abdominal aortic aneurysm disease. *Ann Vasc Surg*. (1991) 5(3):229–33. doi: 10.1007/BF02329378

8. Koch AE, Haines GK, Rizzo RJ, Radosevich JA, Pope RM, Robinson PG, et al. Human abdominal aortic aneurysms. Immunophenotypic analysis suggesting an immune-mediated response. *Am J Pathol*. (1990) 137(5):1199–213.

9. Forester ND, Cruickshank SM, Scott DJ, Carding SR. Functional characterization of T cells in abdominal aortic aneurysms. *Immunology*. (2005) 115(2):262–70. doi: 10.1111/j.1365-2567.2005.02157.x

10. Furusho A, Aoki H, Ohno-Urabe S, Nishihara M, Hirakata S, Nishida N, et al. Involvement of B cells, immunoglobulins, and Syk in the pathogenesis of abdominal aortic aneurysm. *J Am Heart Assoc*. (2018) 7(6):e007750. doi: 10.1161/JAHA.117.007750

11. Meher AK, Johnston WF, Lu G, Pope NH, Bhamidipati CM, Harmon DB, et al. B2 cells suppress experimental abdominal aortic aneurysms. *Am J Pathol*. (2014) 184(11):3130–41. doi: 10.1016/j.ajpath.2014.07.006

12. Saraff K, Babamusta F, Cassis LA, Daugherty A. Aortic dissection precedes formation of aneurysms and atherosclerosis in angiotensin II-infused, apolipoprotein E-deficient mice. *Arterioscler Thromb Vasc Biol*. (2003) 23(9):1621–6. doi: 10.1161/01.ATV.0000085631.76095.64

13. Tian K, Xia C, Liu H, Xu B, Wei P, Fu W, et al. Temporal and quantitative analysis of aortic immunopathologies in elastase-induced mouse abdominal aortic aneurysms. *J Immunol Res*. (2021) 2021:6297332. doi: 10.1155/2021/6297332

14. Griepke S, Grupe E, Lindholt JS, Fuglsang EH, Steffensen LB, Beck HC, et al. Selective inhibition of soluble tumor necrosis factor signaling reduces abdominal aortic aneurysm progression. *Front Cardiovasc Med*. (2022) 9:942342. doi: 10.3389/fcvm.2022.942342

15. Busch A, Bleichert S, Ibrahim N, Wortmann M, Eckstein HH, Brostjan C, et al. Translating mouse models of abdominal aortic aneurysm to the translational needs of vascular surgery. *JVS Vasc Sci*. (2021) 2:219–34. doi: 10.1016/j.jvssci.2021.01.002

16. Wang W, Xu B, Xuan H, Ge Y, Wang Y, Wang L, et al. Hypoxia-inducible factor 1 in clinical and experimental aortic aneurysm disease. *J Vasc Surg*. (2018) 68(5):1538–50.e2. doi: 10.1016/j.jvs.2017.09.030

17. Li G, Zhou H, He Y, Sun S, Wu X, Yuan H. Ulinastatin inhibits the formation and progression of experimental abdominal aortic aneurysms. *J Vasc Res*. (2020) 57(2):58–64. doi: 10.1159/000504848

18. Liu R, Huang J, Ge Y, Liu S, Huang T, Cai H, et al. Inhibition of phosphatidylinositol 3-kinase gamma by Ipi-549 attenuates abdominal aortic aneurysm formation in mice. *Eur J Vasc Endovasc Surg*. (2020) 60(2):254–63. doi: 10.1016/j.ejvs.2020.03.042

19. Dale MA, Ruhlman MK, Baxter BT. Inflammatory cell phenotypes in AAAs: their role and potential as targets for therapy. *Arterioscler Thromb Vasc Biol*. (2015) 35(8):1746–55. doi: 10.1161/ATVBAHA.115.305269

20. Kusters PJH, Seijkens TTP, Beckers L, Lievens D, Winkels H, de Waard V, et al. Cd40l deficiency protects against aneurysm formation. *Arterioscler Thromb Vasc Biol*. (2018) 38(5):1076–85. doi: 10.1161/ATVBAHA.117.310640

21. Schonbeck U, Sukhova GK, Gerdes N, Libby P. T(H)2 predominant immune responses prevail in human abdominal aortic aneurysm. *Am J Pathol*. (2002) 161(2):499–506. doi: 10.1016/S0002-9440(10)64206-X

22. Lindholt JS, Shi GP. Chronic inflammation, immune response, and infection in abdominal aortic aneurysms. *Eur J Vasc Endovasc Surg*. (2006) 31(5):453–63. doi: 10.1016/j.ejvs.2005.10.030

23. Zhou HF, Yan H, Bertram P, Hu Y, Springer LE, Thompson RW, et al. Fibrinogen-specific antibody induces abdominal aortic aneurysm in mice through complement lectin pathway activation. *Proc Natl Acad Sci U S A*. (2013) 110(46):E4335–44. doi: 10.1073/pnas.1315512110

24. Zhang L, Wang Y. B lymphocytes in abdominal aortic aneurysms. *Atherosclerosis*. (2015) 242(1):311–7. doi: 10.1016/j.atherosclerosis.2015.07.036

25. Schaheen B, Downs EA, Serbulea V, Almenara CC, Spinosa M, Su G, et al. B cell depletion promotes aortic infiltration of immunosuppressive cells and is protective of experimental aortic aneurysm. *Arterioscler Thromb Vasc Biol*. (2016) 36(11):2191–202. doi: 10.1161/ATVBAHA.116.307559

26. Chew DK, Knoetgen J, Xia S, Tilson MD. The role of a putative microfibrillar protein (80 kDa) in abdominal aortic aneurysm disease. *J Surg Res*. (2003) 114(1):25–9. doi: 10.1016/S0022-4804(03)00208-7

27. Greiff V, Bhat P, Cook SC, Menzel U, Kang W, Reddy ST. A bioinformatic framework for immune repertoire diversity profiling enables detection of immunological Status. *Genome Med*. (2015) 7(1):49. doi: 10.1186/s13073-015-0169-8

28. Liu X, Wu J. History, applications, and challenges of immune repertoire research. *Cell Biol Toxicol*. (2018) 34(6):441–57. doi: 10.1007/s10565-018-9426-0

29. Rosati E, Dowds CM, Liaskou E, Henriksen EKK, Karlsen TH, Franke A. Overview of methodologies for T cell receptor repertoire analysis. *BMC Biotechnol*. (2017) 17(1):61. doi: 10.1186/s12896-017-0379-9

30. Davis FM, Tsoi LC, Ma F, Wasikowski R, Moore BB, Kunkel SL, et al. Single-cell transcriptomics reveals dynamic role of smooth muscle cells and enrichment of

immune cell subsets in human abdominal aortic aneurysms. *Ann Surg*. (2022) 276 (3):511–21. doi: 10.1097/SLA.0000000000005551

31. Miqueu P, Guillet M, Degauque N, Dore JC, Soulillou JP, Brouard S. Statistical analysis of Cdr3 length distributions for the assessment of T and B cell repertoire biases. *Mol Immunol*. (2007) 44(6):1057–64. doi: 10.1016/j.molimm.2006.06.026

32. Minervina A, Pogorelyy M, Mamedov I. T cell receptor and B cell receptor repertoire profiling in adaptive immunity. *Transpl Int*. (2019) 32(11):1111–23. doi: 10.1111/tri.13475

33. Zarnitsyna VI, Evavold BD, Schoettle LN, Blattman JN, Antia R. Estimating the diversity, completeness, and cross-reactivity of the T cell repertoire. *Front Immunol*. (2013) 4:485. doi: 10.3389/fimmu.2013.00485

34. Greiff V, Menzel U, Miho E, Weber C, Riedel R, Cook S, et al. Systems analysis reveals high genetic and antigen-driven predetermination of antibody repertoires throughout B cell development. *Cell Rep*. (2017) 19(7):1467–78. doi: 10.1016/j.celrep.2017.04.054

35. Elhanati Y, Sethna Z, Marcou Q, Callan CG Jr., Mora T, Walczak AM. Inferring processes underlying B cell repertoire diversity. *Philos Trans R Soc Lond B Biol Sci*. (2015) 370(1676):20140243. doi: 10.1098/rstb.2014.0243

36. Li J, Xia N, Li D, Wen S, Qian S, Lu Y, et al. Aorta regulatory T cells with a tissue-specific phenotype and function promote tissue repair through Tff1 in abdominal aortic aneurysms. *Adv Sci*. (2022) 9(9):e2104338. doi: 10.1002/advs.202104338

37. Lu S, White JV, Lin WL, Zhang X, Solomides C, Evans K, et al. Aneurysmal lesions of patients with abdominal aortic aneurysm contain clonally expanded T cells. *J Immunol*. (2014) 192(10):4897–912. doi: 10.4049/jimmunol.1301009

38. Lu S, White JV, Judy RI, Merritt LL, Lin WL, Zhang X, et al. Clonally expanded alpha-chain T cell receptor (TCR) transcripts are present in aneurysmal lesions of patients with abdominal aortic aneurysm (AAA). *PLoS One*. (2019) 14(7):e0218990. doi: 10.1371/journal.pone.0218990

39. He R, Guo DC, Sun W, Papke CL, Duraisamy S, Estrera AL, et al. Characterization of the inflammatory cells in ascending thoracic aortic aneurysms in patients with Marfan syndrome, familial thoracic aortic aneurysms, and sporadic aneurysms. *J Thorac Cardiovasc Surg*. (2008) 136(4):922–9, 929.e1. doi: 10.1016/j.jtcvs.2007.12.063

40. Li X, Wang CY. From bulk, single-cell to spatial RNA sequencing. *Int J Oral Sci*. (2021) 13(1):36. doi: 10.1038/s41368-021-00146-0

41. Mazzotti L, Gaimari A, Bravaccini S, Maltoni R, Cerchione C, Juan M, et al. T cell receptor repertoire sequencing and its applications: focus on infectious diseases and cancer. *Int J Mol Sci*. (2022) 23(15):8590. doi: 10.3390/ijms23158590

42. Zhao G, Lu H, Chang Z, Zhao Y, Zhu T, Chang L, et al. Single-cell RNA sequencing reveals the cellular heterogeneity of aneurysmal infrarenal abdominal aorta. *Cardiovasc Res*. (2021) 117(5):1402–16. doi: 10.1093/cvr/cvaa214

43. Marquez S, Babrak L, Greiff V, Hoehn KB, Lees WD, Luning Prak ET, et al. Adaptive immune receptor repertoire (AIRR) community guide to repertoire analysis. *Methods Mol Biol*. (2022) 2453:297–316. doi: 10.1007/978-1-0716-2115-8_17

44. Marcou Q, Mora T, Walczak AM. High-throughput immune repertoire analysis with igor. *Nat Commun*. (2018) 9(1):561. doi: 10.1038/s41467-018-02832-w

45. Rubio T, Chernigovskaya M, Marquez S, Marti C, Izquierdo-Altarejos P, Urios A, et al. A Nextflow pipeline for T cell receptor repertoire reconstruction and analysis from RNA sequencing data. *ImmunoInformatics*. (2022) 6:100012. doi: 10.1016/j.immuno.2022.100012

46. Pyo R, Lee JK, Shipley JM, Curci JA, Mao D, Ziporin SJ, et al. Targeted gene disruption of matrix metalloproteinase-9 (gelatinase B) suppresses development of experimental abdominal aortic aneurysms. *J Clin Invest*. (2000) 105(11):1641–9. doi: 10.1172/JCI8931

47. Hu D, Yin C, Mohanta SK, Weber C, Habenicht AJ. Preparation of single cell suspensions from mouse aorta. *Bio Protoc*. (2016) 6(11):e1832. doi: 10.21769/bioprotoc.1832

48. Hao Y, Hao S, Andersen-Nissen E, Mauck WM 3rd, Zheng S, Butler A, et al. Integrated analysis of multimodal single-cell data. *Cell*. (2021) 184(13):3573–87.e29. doi: 10.1016/j.cell.2021.04.048

49. McGinnis CS, Murrow LM, Gartner ZJ. Doubletfinder: doublet detection in single-cell RNA sequencing data using artificial nearest neighbors. *Cell Syst*. (2019) 8(4):329–37.e4. doi: 10.1016/j.cels.2019.03.003

50. Team I. Immunarch: an R package for painless bioinformatics analysis of T cell and B cell immune repertoires. Zenodo (2019). doi: 10.5281/zenodo.3367200

51. Currier JR, Robinson MA. Spectratype/immunoscope analysis of the expressed TCR repertoire. *Curr Protoc Immunol*. (2001) Chapter 10:10.28.1–8.4. doi: 10.1002/0471142735.im1028s38

52. Li H, Ye C, Ji G, Han J. Determinants of public T cell responses. *Cell Res*. (2012) 22(1):33–42. doi: 10.1038/cr.2012.1

53. Simark-Mattsson C, Bergenholtz G, Jontell M, Tarkowski A, Dahlgren UI. T cell receptor V-gene usage in oral lichen planus; increased frequency of T cell receptors expressing V alpha 2 and V beta 3. *Clin Exp Immunol*. (1994) 98(3):503–7. doi: 10.1111/j.1365-2249.1994.tb05519.x

54. Glusman G, Rowen L, Lee I, Boysen C, Roach JC, Smit AF, et al. Comparative genomics of the human and mouse T cell receptor loci. *Immunity*. (2001) 15 (3):337–49. doi: 10.1016/S1074-7613(01)00200-X

55. Goncharov M, Bagaev D, Shcherbinin D, Zvyagin I, Bolotin D, Thomas PG, et al. VDJdb in the pandemic era: a compendium of T cell receptors specific for SARS-CoV-2. *Nat Methods*. (2022) 19(9):1017–9. doi: 10.1038/s41592-022-01578-0

56. Tickotsky N, Sagiv T, Prilusky J, Shifrut E, Friedman N. McPAS-TCR: a manually curated catalogue of pathology-associated T cell receptor sequences. *Bioinformatics*. (2017) 33(18):2924–9. doi: 10.1093/bioinformatics/btx286

57. Amoriello R, Chernigovskaya M, Greiff V, Carnasciali A, Massacesi L, Barilaro A, et al. TCR repertoire diversity in multiple sclerosis: high-dimensional bioinformatics analysis of sequences from brain, cerebrospinal fluid and peripheral blood. *EBioMedicine*. (2021) 68:103429. doi: 10.1016/j.ebiom.2021.103429

58. Pan M, Li B. T cell receptor convergence is an indicator of antigen-specific T cell response in cancer immunotherapies. *Elife*. (2022) 11:e81952. doi: 10.7554/eLife.81952

59. Toghill BJ, Saratzis A, Bown MJ. Abdominal aortic aneurysm-an independent disease to atherosclerosis? *Cardiovasc Pathol*. (2017) 27:71–5. doi: 10.1016/j.carpath.2017.01.008

60. Paulsson G, Zhou X, Törnquist E, Hansson GK. Oligoclonal T cell expansions in atherosclerotic lesions of apolipoprotein E-deficient mice. *Arterioscler Thromb Vasc Biol*. (2000) 20(1):10–7. doi: 10.1161/01.ATV.20.1.10

61. Lin Z, Qian S, Gong Y, Ren J, Zhao L, Wang D, et al. Deep sequencing of the T cell receptor beta repertoire reveals signature patterns and clonal drift in atherosclerotic plaques and patients. *Oncotarget*. (2017) 8(59):99312–22. doi: 10.18632/oncotarget.19892

62. Greiff V, Menzel U, Haessler U, Cook SC, Friedensohn S, Khan TA, et al. Quantitative assessment of the robustness of next-generation sequencing of antibody variable gene repertoires from immunized mice. *BMC Immunol*. (2014) 15:40. doi: 10.1186/s12865-014-0040-5

63. Golledge J, Krishna SM, Wang Y. Mouse models for abdominal aortic aneurysm. *Br J Pharmacol*. (2022) 179(5):792–810. doi: 10.1111/bph.15260

frontiers | Frontiers in Cardiovascular Medicine

TYPE Original Research
PUBLISHED 08 April 2024
DOI 10.3389/fcvm.2024.1265378

OPEN ACCESS

EDITED BY
Hanjoong Jo,
Emory University, United States

REVIEWED BY
Brett David Hambly,
Torrens University Australia, Australia
Maria Grazia Andreassi,
National Research Council (CNR), Italy

*CORRESPONDENCE
Chang-Nan Wang
wangchangnan@shu.edu.cn
Jun Li
li.jun@zs-hospital.sh.cn
Jian Xiao
18916066266@189.cn

†These authors contributed equally to this work and share first authorship

RECEIVED 26 July 2023
ACCEPTED 26 February 2024
PUBLISHED 08 April 2024

CITATION
Liu X-W, Wang P, Zhang L, Zhu Y, Zhai J-Y, Wang C-N, Li J and Xiao J (2024) Single-cell RNA sequencing and ATAC sequencing identify novel biomarkers for bicuspid aortic valve-associated thoracic aortic aneurysm. Front. Cardiovasc. Med. 11:1265378. doi: 10.3389/fcvm.2024.1265378

Single-cell RNA sequencing and ATAC sequencing identify novel biomarkers for bicuspid aortic valve-associated thoracic aortic aneurysm

Xu-Wen Liu[1†], Pei Wang[2†], Li Zhang[1], Yu Zhu[3], Jun-Yu Zhai[4], Chang-Nan Wang[5*], Jun Li[4*] and Jian Xiao[1,2*]

[1]School of Medicine, Guangxi University, Nanning, China, [2]Department of Cardiothoracic Surgery, Changzheng Hospital, Naval Medical University, Shanghai, China, [3]School of Health Sciences and Engineering, University of Shanghai for Science and Technology, Shanghai, China, [4]Department of Cardiac Surgery, Zhongshan Hospital, Fudan University, Shanghai, China, [5]School of Life Sciences, Shanghai University, Shanghai, China

Introduction: Bicuspid aortic valve (BAV) is the most prevalent congenital cardiovascular defect and known to cause thoracic aortic aneurysms (TAAs). To improve our understanding of BAV pathogenesis, we characterized the cellular composition of BAV tissues and identified molecular changes in each cell population.
Methods: Tissue samples from two patients with BAV and two heart transplant donors were analyzed using single-cell RNA sequencing, assay for transposase-accessible chromatin using sequencing, and weighted gene coexpression network analysis for differential gene analysis. TAA-related changes were evaluated by comparing the proportion of each cell type and gene expression profiles between TAA and control tissues. Further, by combining our single-cell RNA sequencing data with publicly available data from genome-wide association studies, we determined critical genes for BAV.
Results: We found 20 cell subpopulations in TAA tissues, including multiple subtypes of smooth muscle cells, fibroblasts, macrophages, and T lymphocytes. This result suggested that these cells play multiple functional roles in BAV development. Several differentially expressed genes, including CD9, FHL1y, HSP90AA1, GAS6, PALLD, and ACTA2, were identified.
Discussion: We believe that this comprehensive assessment of the cellular composition of TAA tissues and the insights into altered gene expression patterns can facilitate identification of novel diagnostic biomarkers and therapeutic targets for BAV-associated TAA.

KEYWORDS

bicuspid aortic valve, thoracic aortic aneurysm, scRNA-seq, ATAC-seq, diagnostic biomarkers, UMAP, WGCNA

1 Introduction

Bicuspid aortic valve (BAV), the most common congenital cardiovascular defect with an incidence of 1%–2% in the general population (1), can cause aortic disease and increase the risk of aortic-related diseases in affected patients (2–5). BAV is one of the causes of thoracic aortic aneurysms (TAAs) (6–10), which occur when the aortic blood vessel expands, causing the aorta to become >1.5 times the standard arterial diameter. Aortic

aneurysms are typically asymptomatic during progressive enlargement until acute aortic dissection occurs, which is linked to a high mortality rate. TAA is a dangerous, potentially fatal, and asymptomatic condition that can involve one or more segments of the thoracic aorta (4, 6).

Recent studies have revealed the complex and unique pathogenesis of BAV-associated TAA, encompassing gene mutations, hemodynamics, mechanical stress, oxidation and inflammation, as well as their interactions (10). However, previous research has paid limited attention to cell-specific changes in the aortic wall of patients with BAV-associated TAA.

To address this gap, herein we investigated TAA tissues from patients with BAV using single-cell RNA sequencing (scRNA-seq) and assay for transposase-accessible chromatin using sequencing (ATAC-seq) to elucidate the molecular processes underlying BAV-associated TAA occurrence and development. scRNA-seq is a powerful tool to characterize gene expression in individual cells; moreover, it allows for the identification of intercellular variations and can reveal complex cell populations and regulatory relationships between genes. ATAC-seq is a method to assess open chromatin and evaluate genome-wide chromatin accessibility in individual cells. We also performed weighted gene coexpression network analysis (WGCNA), a method increasingly employed in disease and gene association analysis, to identify key genes and better comprehend disease regulatory mechanisms. We believe that our findings will facilitate the identification of novel diagnostic biomarkers and therapeutic targets for BAV-associated TAA and enhance our understanding of its pathogenesis.

2 Materials and methods

2.1 Tissue collection

Aortic aneurysm patient tissues were obtained from two patients with BAV, and control tissues were obtained from two heart transplant recipients. This study was approved by the Ethics Committee of Shanghai Changzheng Hospital, all participants were informed about our research (Table 1).

TABLE 1 Control group patient information form.

	Patient 1	Patient 2
Age	64	32
Sex	Male	Male
BMI	23.71	24.32
Primary cause	Myocardiopathy	Myocardiopathy
Risk factors		
Diabetes	No	No
Hypertension	Yes	Yes
Cerebrovascular disease	No	No
Previous history of cardiac surgery	No	No
Smoking	No	Yes
Drinking	Yes	Yes
Aortic valve condition	Normal	Normal
Aortic diameter	31 mm	34 mm
Left ventricular end diastolic diameter	59 mm	61
Pulmonary arterial pressure	48 mmHg	44 mmHg

In the control group, the aortic wall of patients undergoing heart transplantation was examined. During the surgical procedure of removing the donor heart, the aortic wall was excised and trimmed. The aortic valve of the recipient patient was tricuspid, and the aorta showed no significant dilation, with no other vascular lesions.

2.2 ATAC-seq

We lysed 50,000 cells in cold ATAC resuspension buffer containing 0.1% NP40, 0.1% Tween-20, and 0.01% digitonin, followed by incubation on ice for 3 min. Subsequently, 1 ml precooled resuspension buffer containing only 0.1% Tween-20 was added, mixed by inversion three times, and centrifuged at 500 ×g for 5 min at 4°C. The supernatant thus obtained was discarded; 50 µl transposition mix was then added, followed by incubation at 37°C on a rotating mixer at 1,000 RPM for 30 min. DNA fragments were purified using the Qiagen MinElute PCR Purification Kit and amplified by 5–9 PCR cycles. After library detection and quality control, sequencing was performed on the Illumina HiSeq/NextSeq platform. Raw reads obtained upon ATAC-seq were filtered to remove adapters and contaminants before being aligned to the reference genome. High-quality mapped reads (MPAQ ≥30) were used for subsequent analyses.

2.3 Dataset collection and processing

The following four gene expression profile datasets were retrieved from Gene Expression Omnibus (GEO) (http://www.ncbi.nlm.nih.gov/geo): GSE5180, GSE83675, GSE61128, and GSE2615 (11–13). In addition, we utilized self-test datasets, which included two aneurysm patients and two control groups. Single-cell data were also derived from GSE155468 (14).

The original dataset was subjected to background correction and normalization using the Affy R package. Datasets requiring data conversion were $\log_2$ transformed. In cases where multiple probes corresponded to the same gene, we calculated the average value to determine its expression level. After merging the five datasets, we used the Bioconductor sva package in R to eliminate batch effects, resulting in a batch effect-free matrix for subsequent analyses.

2.4 Weighted gene coexpression network construction

WGCNA is employed to analyze gene modules with high biological significance and investigate the relationship between gene networks and diseases. The WGCNA R package is used for constructing a weighted gene coexpression network. First, hierarchical clustering is performed to detect outliers and remove abnormal samples. A soft threshold value of $\beta = 12$ is selected based on $R^2 > 0.9$ to achieve scale-free topology. The minimum number of module genes is then set to 25 and module merging threshold is set to 0.25 to construct a weighted coexpression

network and identify gene modules. Finally, the correlation between modules and phenotypes is determined, and modules highly correlated with phenotypes are subjected to further analysis.

2.5 Chromatin accessibility analysis

The FastQC tool was applied for quality control of clean data after removing adapters and low-quality data. Clean data were aligned to the reference genome using Bowtie2. The ATACseqQC R package was used to create an insertion fragment length distribution map of the sample, which facilitated the preliminarily evaluation of the quality of the ATAC experiment. The deepTools multiBamSummary and plotCorrelation programs were employed for correlation analyses of duplicate samples. deepTools was also used to generate transcriptional start site (TSS) enrichment heatmaps, indicating the enrichment status of data on TSS and comparing the enrichment status of different samples in the TSS region. The alignment results were then processed using MACS2 for peak calling, with $q < 0.05$ set as the threshold. The ChIPseeker R package was utilized for annotation and visualization of peak calling data.

2.6 Gene ontology (GO) and Kyoto encyclopedia of genes and genomes (KEGG) pathway enrichment analyses

The GO system contains structured and computable information on gene and gene product functions. As a knowledge base, KEGG facilitates evaluating functions of genes and pertinent signaling pathways. We performed functional enrichment analysis using the clusterProfiler R package, and the results of the enrichment analysis were visualized using the ggplot2 R package. $p < 0.05$ indicated significant enrichment.

2.7 Single-cell data processing and cell subpopulation identification

We used the Seurat R package for data processing, ensuring that each gene was expressed in at least 3 cells, with each cell expressing at least 250 genes. Cells containing >5% mitochondrial genes were filtered out, and data were normalized using the log-normalization function. Subsequently, clustering based on differences in gene expression between cells was performed to analyze variations among different cell populations. To achieve this, highly variable genes were selected using the FindVariableFeatures function, and their expression matrix dimension was reduced using the RunUMAP function. The two-dimensional uniform manifold approximation and projection (UMAP) algorithm was applied using the RunUMAP function in Seurat to visualize cells. The FindAllMarkers function was employed to identify cell marker genes and cell subtypes, and the FindMarkers function was utilized to identify differentially expressed genes (DEGs).

2.8 Screening of potential diagnostic biomarkers

To identify novel and key biological markers for TAA, we employed three methods: single-cell differential gene analysis, ATAC differential gene analysis, and WGCNA. DEGs obtained on applying these methods were intersected and sorted by differential expression multiple. From this analysis, we chose the top six genes exhibiting the most significant differences as the final hub genes.

2.9 Pseudo-time analysis

To investigate changes in cell state, we utilized the Monocle2 R package. DEGs over the pseudo-time among cluster cell transitions were calculated with the differentialGeneTest function. DDRTree was employed for dimensionality reduction and visualization, while plot_cell_trajectory was utilized to plot the minimum spanning tree on cells. The SMC Contractile, SMC Proliferating, Fibroblast, and MSC clusters were partitioned and visualized using Monocle3, highlighting hub gene expression changes from the start to the end of the pseudo-time process.

2.10 Statistical analysis

GraphPad Prism was used for statistical analysis. An independent sample test was applied to assess the significance of the differences in hub gene expression among patients with BAV. Statistical significance was set at $p < 0.05$.

3 Results

3.1 Weighted gene coexpression network construction and module preservation analysis

To investigate potential differential expression patterns during TAA formation and development, we utilized the WGCNA method, employing an unordered network to modularize and enrich the genes of 28 patients with TAA (Figure 1). The genes were classified based on their respective expression levels; after batch processing (Figures 1A,B), we determined that a soft threshold (β) value of 12 yielded the most suitable connectivity between genes in the gene network (Figure 1C). Subsequently, 11 coexpressed modules were identified, with the gray module containing genes that were not assigned to any module (Figure 1D). Among these modules, the blue, turquoise, brown, and yellow modules exhibited Zsummary statistics >10 and exhibited the highest stability. The blue and turquoise modules also demonstrated relatively small median rank statistics, indicating relatively good repeatability (Figure 1E) and a positive correlation with TAA.

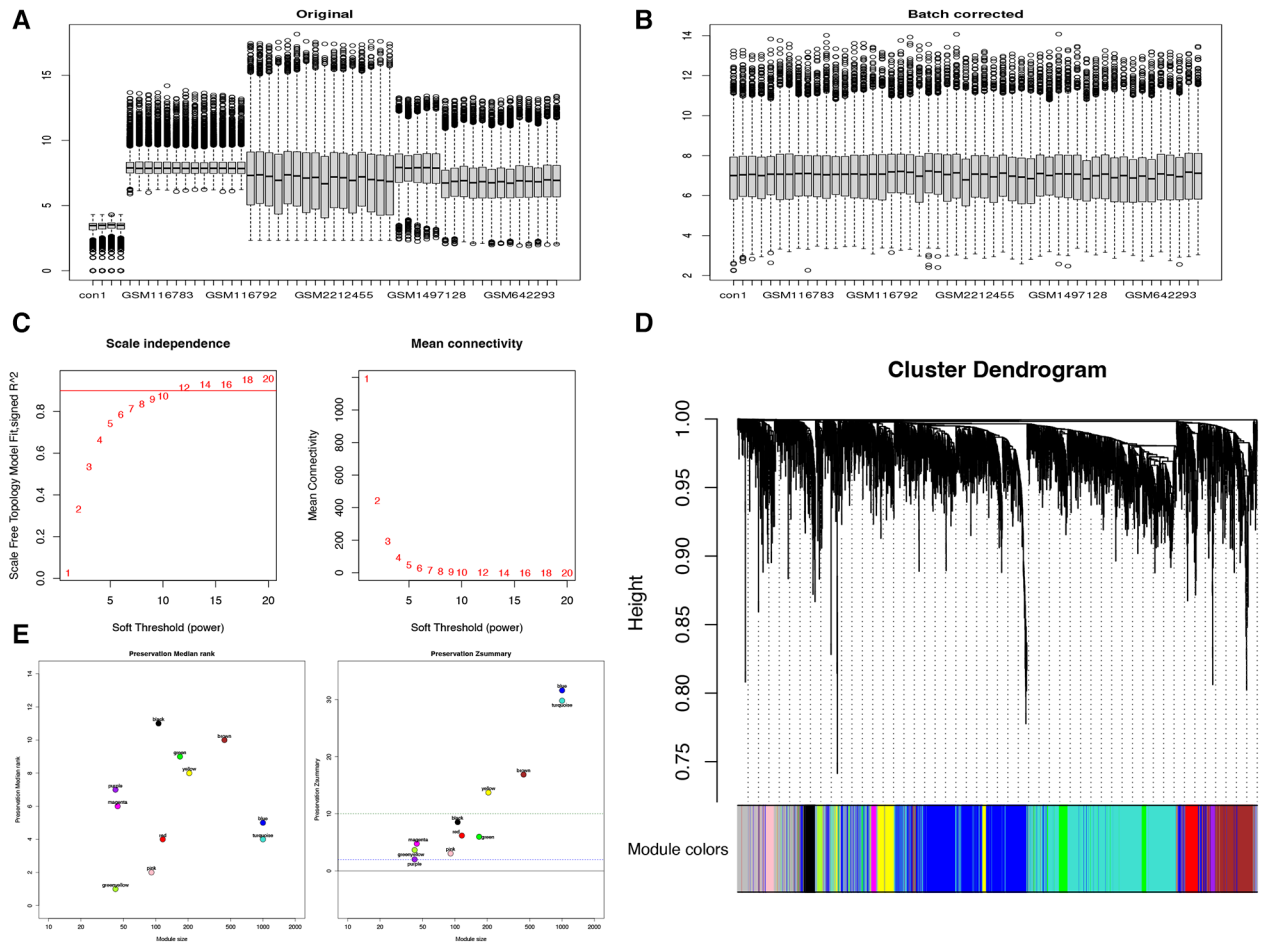


FIGURE 1
WGCNA reveals gene coexpression network in 28 patients with TAA. (**A**) Before and (**B**) after de-batching. (**C**) Analysis of the scale-free fit index (left) and mean connectivity (right) for different soft-thresholding powers. (**D**) Clustering dendrogram of TAA-related DEGs. (**E**) Module reproducibility.

3.2 Correlation analysis between module characteristics and TAA

To further validate the relevance of each gene module to TAA, we performed a correlation analysis between each module and TAA. Based on the heatmap of the correlation between the modules and TAA and the module significance map of the coexpression modules related to TAA, we chose the MEblue module for subsequent analyses. We analyzed the correlation between the genes in the blue module and TAA (Figure 2C). Utilizing three criteria, i.e., correlation with TAA >0.3, gene significance >0.3, and module membership >0.7, we identified hub genes, which were then subjected to KEGG pathway and GO functional enrichment analysis. The hub genes were found to be significantly enriched in pathways such as complement and coagulation cascades, PI3K–Akt signaling pathway, cytokine–cytokine receptor interaction, cell adhesion molecules, Ras signaling pathway, neuroactive ligand–receptor interaction, and chemokine signaling pathway (Figure 2D). Furthermore, the hub genes were significantly enriched in diverse biological processes, such as leukocyte migration, extracellular structure organization, positive regulation of cell adhesion, regulation of peptide secretion, T cell activation, positive regulation of cytokine production, muscle tissue development, and muscle contraction. The cellular components enriched included, for example, extracellular side of plasma membrane, neuronal cell body, membrane region, membrane microdomain, cell–substrate junction, cell–substrate adherens junction, vesicle lumen, and cytoplasmic vesicle lumen, and the molecular functions enriched included, for example, receptor ligand activity; cell adhesion molecule binding; DNA-binding transcription activator activity, RNA polymerase II-specific; cytokine receptor binding; G protein-coupled receptor binding; and enzyme inhibitor activity (Figure 2E).

3.3 ATAC-seq

For chromatin accessibility analysis, ATAC-seq was performed on tissue samples obtained from two patients with BAV and control tissue samples obtained from two heart transplant

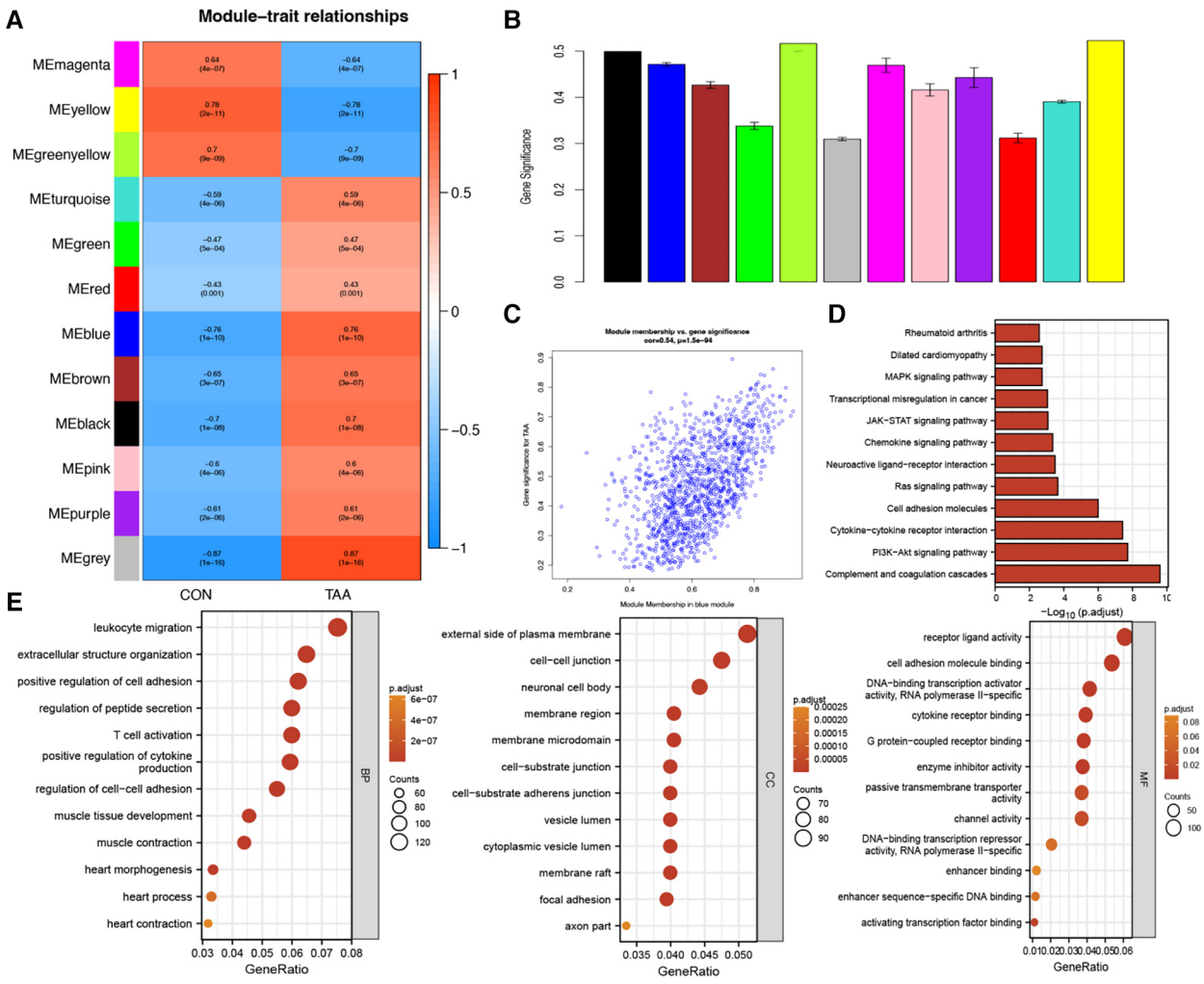


FIGURE 2
Module–trait correlation analyses. (A) Heatmap showing the correlation between modules and TAA. (B) Module significance values of coexpression modules associated with TAA. (C) Gene significance for TAA in the blue module. (D,E) KEGG pathway and GO functional enrichment analysis of hub genes.

donors. The determined accessible regions were observed to be mainly enriched within 3 kb of the TSS (Figure 3A), and in the presumed accessible regions, >70% promoters were located 1 kb upstream of the TSS (Figure 3B). Moreover, the distribution and enrichment of all peaks on chromosomes, which were plotted using Circos, indicated that BAV-associated TAA chromatin accessibility was generally higher than that of the control group, possibly reflecting increased cell type diversity in TAA (Figures 3A–C). KEGG pathway and GO functional enrichment analyses of transcription factor binding site-associated genes revealed significant enrichment in pathways such as Rap1 signaling pathway, MAPK signaling pathway, Ras signaling pathway, regulation of actin cytoskeleton, and EGFR tyrosine kinase inhibitor resistance (Figure 3D); furthermore, the associated genes were significantly enriched in various biological processes, such as regulation of cell morphogenesis, regulation of GTPase activity, extracellular structure organization, actin filament organization, and extracellular matrix organization. The cellular components enriched included, for example, cell–substrate junction, cell–substrate adherens junction, focal adhesion, collagen-containing extracellular matrix, and axon part, and the molecular functions enriched included, for example, actin binding, cell adhesion molecule binding, small GTPase binding, Ras GTPase binding, nucleotide-triphosphatase regulator activity, and GTPase regulator activity (Figure 3E).

3.4 Single-cell data processing and cell subpopulation identification

Unsupervised clustering based on Seurat identified 19 different cell populations (Figure 4A). Cell subpopulations were further identified based on the top five marker genes for each cell subpopulation (Figure 4B). Comparing TAA and control samples, 20 cell subpopulations were classified according to characteristic markers (Figure 4C), and their distribution

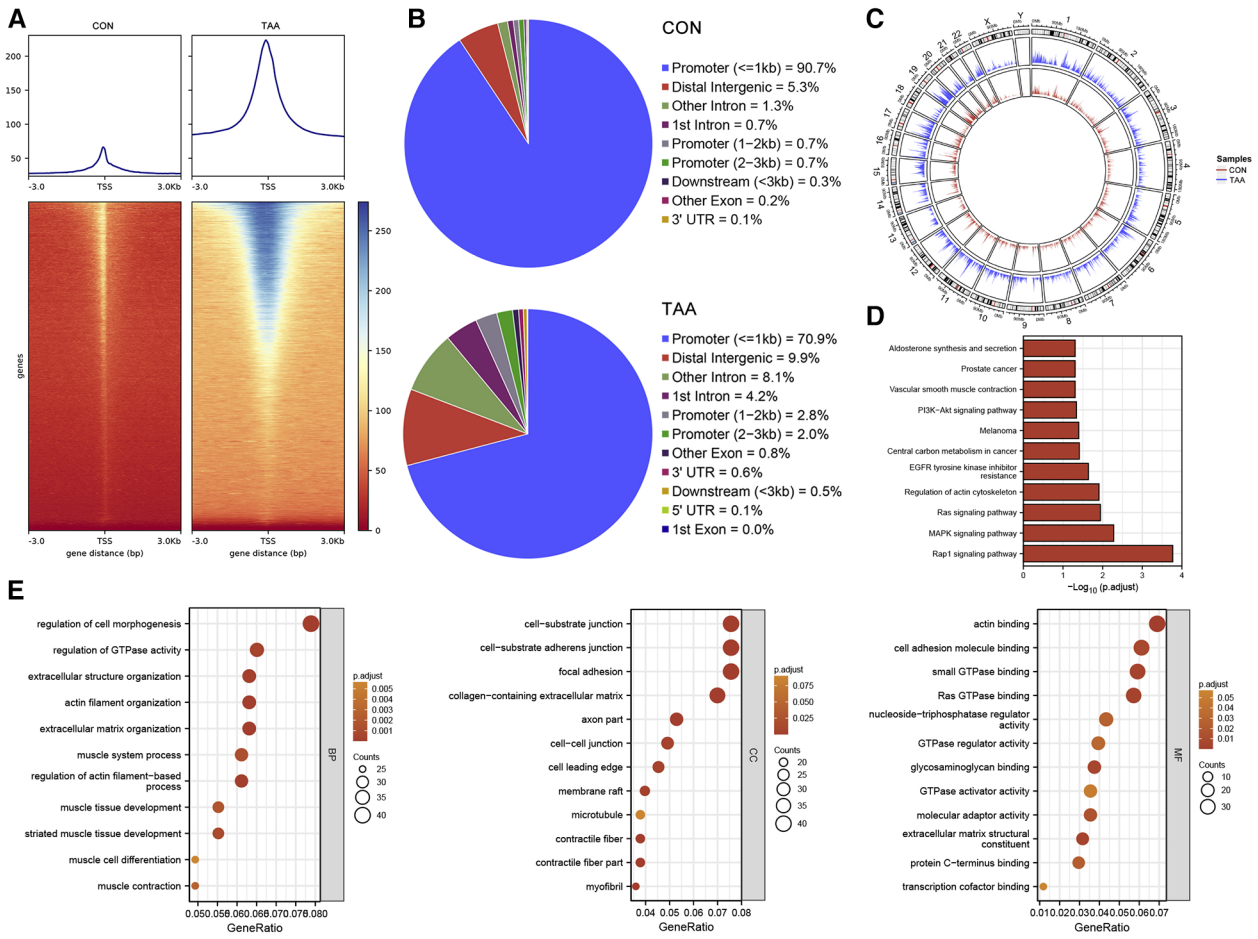


FIGURE 3
ATAC-seq results. (**A**) TSS enrichment heatmap. (**B**) Distribution proportion of chromatin accessible regions. (**C**) Circos plot, with the outer circles indicating chromosomal positions and the inner circles representing different samples. (**D**,**E**) KEGG and GO functional enrichment analysis of transcription factor binding site-associated genes.

differences were analyzed (Figure 4D). Relative to the control group, the TAA group exhibited fewer SMC_Proliferating cells, SMC_Contractile cells, ECs, MSCs, M2like1s, fibroblasts, Plasma cells, Mast cells, and B cells but higher CD4_active cells, T_HSP cells, CD8_active cells, Tregs, M1like1s, M2like2s, T_GIMAPs. The relative numbers of MonoMaphDCs, CD8_TEMRAs, M1like2s, and Monocytes were unaffected. Collectively, these findings suggested that cells with considerable differences, such as CD4_active, SMC_Contractile, and SMC_Proliferating cells, are involved in the process of BAV-induced TAA onset.

3.5 Differential gene analysis

By comparing the intersection of DEGs identified from WGCNA and ATAC-seq, we obtained 34 DEGs (Figure 5A), which were subsequently annotated for each cell subpopulation of TAA (Figure 5B). Among them, the expression of CD9, FHL1y, and heat shock protein 90 alpha family class A member 1 (HSP90AA1) was upregulated and that of GAS6, PALLD, and ACTA2 was downregulated. Further investigation of their expression levels in various cell types revealed significant expression in cells such as SMC_Contractile cells, SMC_Proliferating cells, Fibroblasts, and MSCs (Figures 5C,D).

3.6 Pseudotemporal analysis of cell trajectory changes in TAA

During TAA development, cellular phenotype and functions undergo continuous changes. Pseudo-time analysis results revealed that SMC_Contractile cells were the predominant cell type in the early stage of TAA development. As the cells progressed through the branching node, SMC_Proliferating cells, Fibroblasts, and MSCs became the main cells in TAA development (Figures 6A,B). To investigate the developmental process of these cells in TAA, we performed trajectory analysis using Monocle3, sorting the clusters along the differentiation stage to visualize the process of cell differentiation in TAA (Figures 6C,D). In addition, using Monocle3, we plotted the

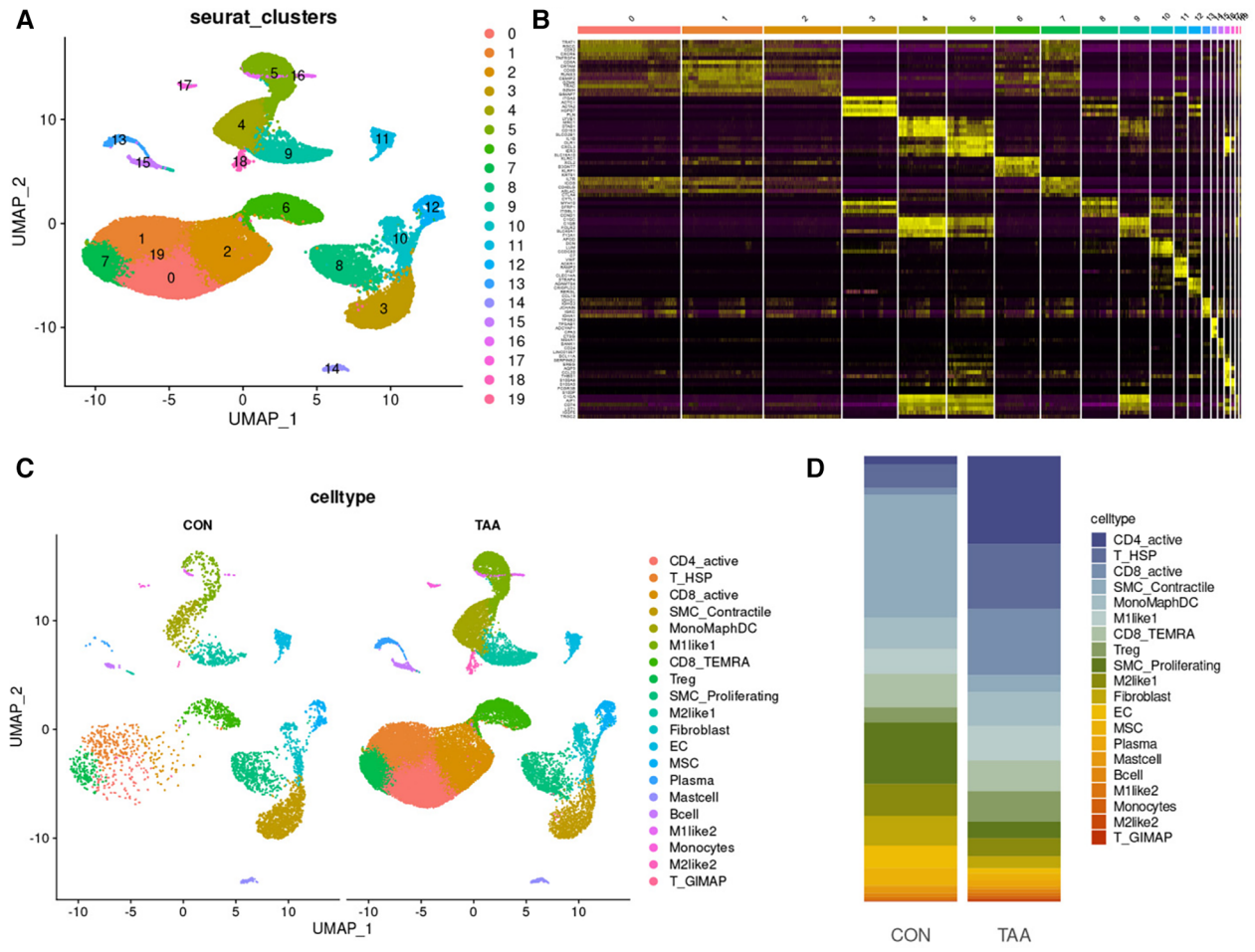


FIGURE 4
Cell-type classification in TAA. (A) UMAP plot of 20 cell clusters. (B) Heatmap of the top five marker genes in each cell cluster. (C) Cell distribution in TAA and control tissues. (D) Bar plot of the distribution of all identified cell clusters: proportion of each cell in TAA and control tissues.

relative expression levels of six genes to highlight their changing trends in the cell development trajectory (Figures 6E,F).

4 Discussion

TAA is a life-threatening condition characterized by continuous expansion and chronic progression. Approximately 20% TAAs are attributed to specific genetic mutations (7–10, 15–19). The pathogenesis of most TAAs caused by genetic defects can be linked to abnormal development or functional defects of connective tissue, as observed in Marfan syndrome and Loeys–Dietz syndrome (10, 20). In addition, thoracic aortic dilation is frequently observed in patients with BAV (4, 21).

In this study, we utilized single-cell sequencing technology to comprehensively analyze key genes and biological processes involved in the pathogenesis of BAV-associated TAA. Our analyses led to the identification of pivotal coexpression modules related to functional pathways as well as cells that play a crucial role in TAA development, including SMC_Contractile cells, SMC_Proliferating cells, Fibroblasts, and MSCs. Furthermore, genes that were markedly up- (CD9, FHL1y, and HSP90AA1) or downregulated (GAS6, PALLD, and ACTA2) were identified. Our comprehensive bioinformatics analysis also led to the identification of differentially expressed transcription factors, downstream marker pathways/immune-related pathways, and immune cells, highlighting key immune-related genes in BAV-associated TAA.

Herein we observed significant differences in the expression of CD9 (22) in MonoMaphDCs, CD8_TEMRAs, Tregs, and M2like2s. CD9 plays a chief role in various biological activities, including cell adhesion, migration, metastasis, growth, signal transduction, and differentiation; moreover, it participates in disease regulation and mediation. On the other hand, FHL1 (23, 24) showed significant differences in its expression in SMC_Proliferating cells, M2like1s, and ECs. This could be attributed to its role in inhibiting the proliferation of aortic vascular SMCs and influencing aortic wall remodeling. HSP90AA1 (25–27) showed significant differences in its expression in CD4_active cells, CD8_active cells, SMC_Contractile cells, M1like1s, and Tregs. HSP90AA1 is a

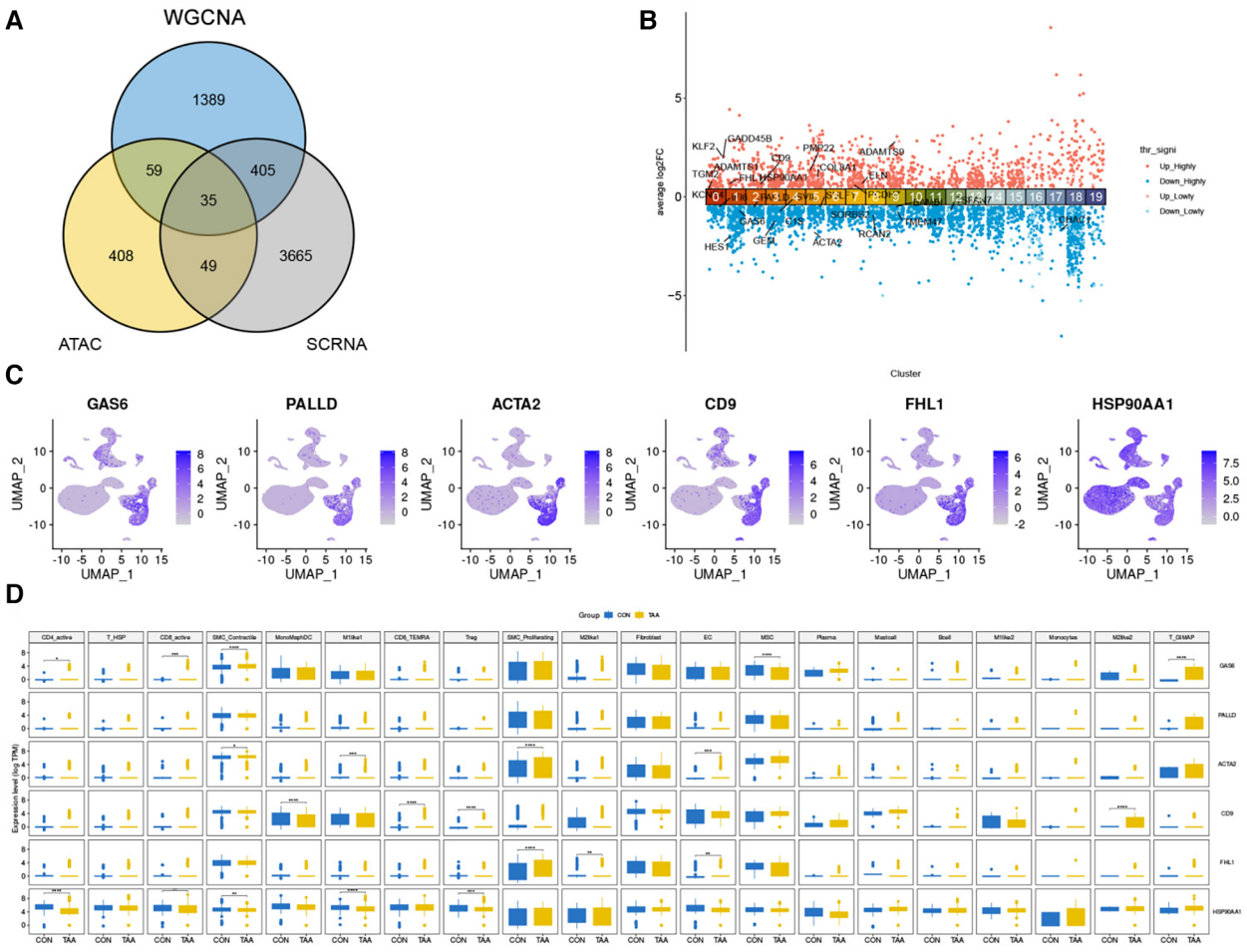


FIGURE 5
Differential genes in cell populations of TAA. (**A**) Venn diagram showing the number of differential genes identified from WGCNA and ATAC-seq, resulting in 34 genes. (**B**) DEG annotation for each cell subpopulation of TAA (adjusted $p < 0.01$ = highly significant). (**C**) UMAP plot of the six genes with higher expression levels. (**D**) Expression levels of these six genes in each cell type.

critical molecular chaperone that is highly conserved in evolution and expressed under the stimulatory circumstances of trauma, infection, and tumors. It is extracellularly expressed and involved in several cell protection mechanisms. Besides, HSP90AA1 participates in tumor progression and cancer cell invasion, and is known to promote cancer cell proliferation and metastasis. In the pathogenesis of TAA, the expression of GAS6 (28) showed significant differences in CD4_active cells, CD8_active cells, SMC_Contractile cells, MSCs, and T_GIMAPs. GAS6 is a cytokine that is associated with various biological processes, including cell proliferation, apoptosis, migration, and inflammatory response, and expressed in many cell types, including ECs, vascular SMCs, and fibroblasts. It also plays a vital role in the occurrence and development of aortic aneurysms by activating the Axl receptor; furthermore, GAS6 affects aortic aneurysm formation by regulating platelet activation and coagulation function. GAS6 protein expression was originally found to be upregulated in growth-arrested fibroblasts. GAS6 stimulation has been reported to rescue serum-starved fibroblasts and vascular SMCs from apoptosis. PALLD (29) encodes a cytoskeletal protein involved in actin reorganization and plays an important role in heart development. ACTA2 (1, 30) expression showed a significant difference in SMC_Contractile cells, M1like1s, SMC_Proliferating cells, and ECs. ACTA2 encodes α-2 smooth muscle actin, which is involved in cell migration and muscle contraction. In aortic aneurysms, mutations in ACTA2 have been linked to weakening and expansion of the aortic wall. ACTA2 is the most common mutation gene responsible for TAA and dissection. Specific ACTA2 mutations have also been linked to an increased risk of early-onset stroke or coronary artery disease.

GO enrichment analysis revealed that associated genes were involved in diverse biological processes, including cell proliferation, indicating extensive proliferation and activation of immune cells during the pathogenesis of aortic aneurysm. Moreover, these genes were enriched in molecular functions related to extracellular matrix organization, suggestive of potential weakening of the aortic structure due to alterations in the extracellular matrix, making it more susceptible to expansion and rupture.

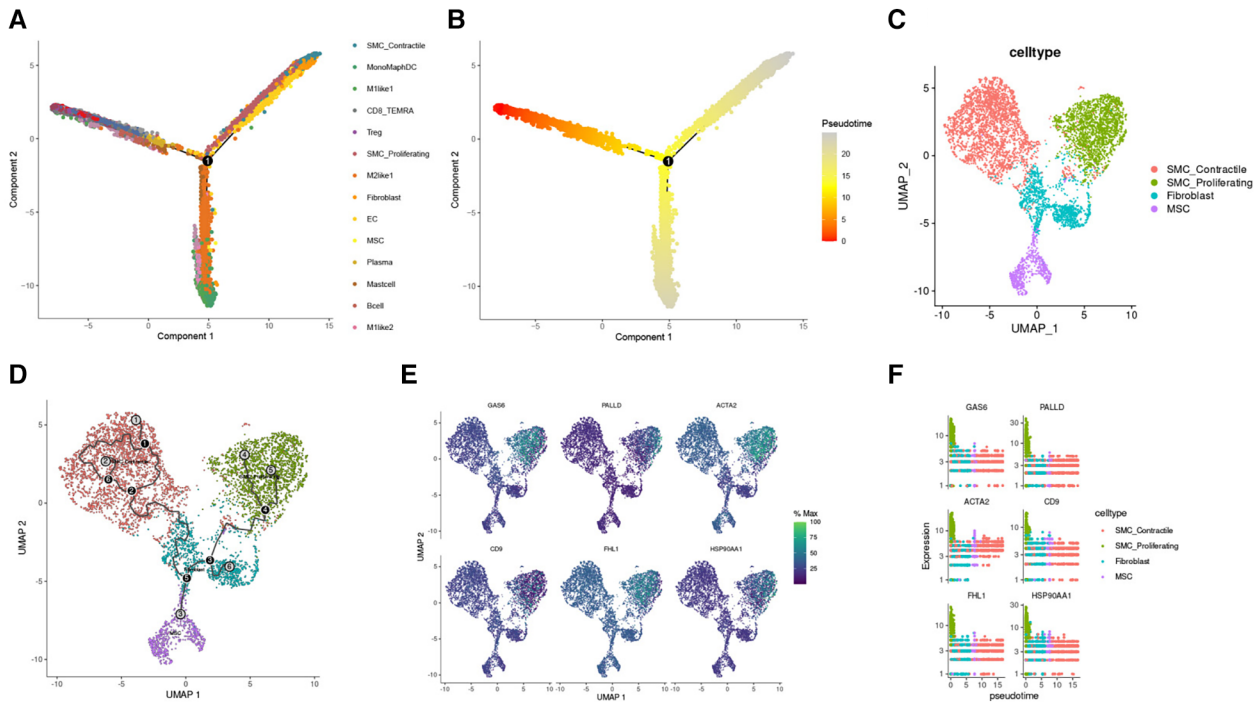


FIGURE 6
Reverse chronological analysis of single-cell samples. (A) Different colors represent different cell clusters. (B) The proposed temporal order, indicated by color from dark to light. (C) Partitioning of SMC Contractile, SMC Proliferating, Fibroblast, and MSC clusters using Monocle3. (D) Identification of cell trajectories. (E) FeaturePlot of the six genes mentioned above. (F) Gene expression trend plots for the six genes (gene expression changes along the proposed time course from start to termination).

A limitation of this study is the small sample size, where only two patients and two controls were included. Despite this, our data reveal the cellular and molecular landscape of BAV at the single-cell level, providing valuable insights into TAA tissue cell morphology and cellular matrix structure. Furthermore, our findings offer a molecular profile of multiple isoforms of aortic SMCs, fibroblasts, macrophages, and T lymphocytes, indicating a potentially critical role of CD9, FHL1y, HSP90AA1, GAS6, PALLD, and ACTA2 genes in BAV improvement. We believe that our results enhance our understanding of the pathogenesis of BAV and may contribute to the development of new therapeutic approaches.

In this study, the data of the public dataset GSE5180, GSE83675, GSE61128 and GSE26155 were used to obtain genes significantly related to the onset of TAA patients, and then the self-test data were used to obtain significant genes related to BAV patients. By comparing the two, we identified the key immune-related genes of BAV-related TAA, and clarified the cell populations that may be involved in the pathogenesis process and their developmental trajectories. Through this study, we revealed the cellular and molecular landscape of the bicuspid aortic valve at the single-cell level, illustrated the dynamic changes of cells in terms of morphological structure and functional properties during the pathogenesis of BAV, illustrated the dynamic changes of chromatin during the pathogenesis of BAV, and suggested the potential roles of CD9, FHL1y, HSP90AA1, GAS6, PALLD and ACTA2 in BAV. These findings suggest that CD9 and other genes mediate the development regulation of BAV through cell adhesion, motility, metastasis, growth, signal transduction, differentiation and other functions. FHL1y may be involved in inhibiting the proliferation of aortic vascular SMCs and affecting aortic wall remodeling, thereby participating in the development of BAV; HSP90AA1 is stimulated to express during the pathogenesis of BAV, which may participate in the proliferation and metastasis of cells and affect BAV; GAS6 may participate in the occurrence and development of BAV by affecting fibroblasts and apoptosis of vascular smooth muscle cells. PALLD may regulate BAV through regulation of the actin cytoskeleton; ACTA2 encodes smooth muscle α-2 actin, which affects the expansion of the aorta wall and thus BAV development.

In conclusion, we performed a comprehensive analysis of the dynamics of cellular and molecular changes during BAV development. Our study identified SMCs, Fibroblast, and MSCs as the main sources of cells affecting BAV. In addition, we demonstrate that a variety of potential genes such as CD9, FHL1y, HSP90AA1, GAS6, PALLD, and ACTA2 play a role in the development of BAV, which may provide novel research ideas for slowing the progression of BAV and preventing aortic rupture. Overall, our dataset provides a valuable resource for further exploration of the pathogenesis of BAV.

The results of the present study comprised of a comprehensive bioinformatics view of BAV development. Overall, this could shed

light on the mechanisms of BAV, which could lead to improved diagnosis and treatments for patients with BAV. CD9, FHL1y, HSP90AA1, GAS6, PALLD, and ACTA2 are closely related to the occurrence and development of BAV, and are expected to become new biomarkers and therapeutic targets for BAV.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary Material.

Ethics statement

The studies involving humans were approved by the Ethics Committee of Shanghai Changzheng Hospital. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

XW-L: Formal analysis, Investigation, Writing – original draft, Writing – review & editing. PW: Data curation, Methodology, Investigation, Writing – original draft, Writing – review & editing. LZ: Methodology, Supervision, Formal analysis, Writing – review & editing. YZ: Formal analysis, Writing – review & editing. JY-Z: Methodology, Investigation, Writing – original draft. CN-W: Data curation, Supervision, Writing – original draft. JL: Resources, Writing – review & editing. JX. Methodology, Supervision, Funding acquisition, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article.

This work was supported by the National Nature Science Foundation of China (82070255), the Voyage Talent Project of Naval Medical University (2019-YH-12) and the Navigation Talent Project of Naval Medical University (SL03).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcvm.2024.1265378/full#supplementary-material

References

1. Lo Presti F, Guzzardi DG, Bancone C, Fedak PWM, Della Corte A. The science of BAV aortopathy. *Prog Cardiovasc Dis.* (2020) 63(4):465–74. doi: 10.1016/j.pcad.2020.06.009

2. Ma M, Li Z, Mohamed MA, Liu L, Wei X. Aortic root aortopathy in bicuspid aortic valve associated with high genetic risk. *BMC Cardiovasc Disord.* (2021) 21(1):413. doi: 10.1186/s12872-021-02215-y

3. Mozzini C, Girelli D, Cominacini L, Soresi M. An exploratory look at bicuspid aortic valve (Bav) aortopathy: focus on molecular and cellular mechanisms. *Curr Probl Cardiol.* (2021) 46(3):100425. doi: 10.1016/j.cpcardiol.2019.04.005

4. Norton E, Yang B. Managing thoracic aortic aneurysm in patients with bicuspid aortic valve based on aortic root-involvement. *Front Physiol.* (2017) 8:397. doi: 10.3389/fphys.2017.00397

5. Rashed ER, Dembar A, Riasat M, Zaidi AN. Bicuspid aortic valves: an up-to-date review on genetics, natural history, and management. *Curr Cardiol Rep.* (2022) 24(8):1021–30. doi: 10.1007/s11886-022-01716-2

6. Ostberg NP, Zafar MA, Ziganshin BA, Elefteriades JA. The genetics of thoracic aortic aneurysms and dissection: a clinical perspective. *Biomolecules.* (2020) 10(2):182. doi: 10.3390/biom10020182

7. Pinard A, Jones GT, Milewicz DM. Genetics of thoracic and abdominal aortic diseases. *Circ Res.* (2019) 124(4):588–606. doi: 10.1161/CIRCRESAHA.118.312436

8. Quintana RA, Taylor WR. Cellular mechanisms of aortic aneurysm formation. *Circ Res.* (2019) 124(4):607–18. doi: 10.1161/CIRCRESAHA.118.313187

9. Salameh MJ, Black JH 3rd, Ratchford EV. Thoracic aortic aneurysm. *Vasc Med.* (2018) 23(6):573–8. doi: 10.1177/1358863X18807760

10. Senser EM, Misra S, Henkin S. Thoracic aortic aneurysm: a clinical review. *Cardiol Clin.* (2021) 39(4):505–15. doi: 10.1016/j.ccl.2021.06.003

11. Majumdar R, Miller DV, Ballman KV, Unnikrishnan G, McKellar SH, Sarkar G, et al. Elevated expressions of osteopontin and tenascin C in ascending aortic aneurysms are associated with trileaflet aortic valves as compared with bicuspid aortic valves. *Cardiovasc Pathol.* (2007) 16(3):144–50. doi: 10.1016/j.carpath.2006.12.001

12. Folkersen L, Wagsater D, Paloschi V, Jackson V, Petrini J, Kurtovic S, et al. Unraveling divergent gene expression profiles in bicuspid and tricuspid aortic valve patients with thoracic aortic dilatation: the ASAP study. *Mol Med.* (2011) 17(11–12):1365–73. doi: 10.2119/molmed.2011.00286

13. Suur BE, Chemaly M, Liljeqvist ML, Djordjevic D, Stenemo M, Bergman O, et al. Therapeutic potential of the proprotein convertase subtilisin/kexin (PCSK) family in vascular disease. *Arterioscl Throm Vas.* (2022) 13:988561. doi: 10.3389/fphar.2022.988561

14. Li Y, Ren P, Dawson A, Vasquez HG, Ageedi W, Zhang C, et al. Single-cell transcriptome analysis reveals dynamic cell populations and differential gene expression patterns in control and aneurysmal human aortic tissue. *Circulation.* (2020) 142(14):1374–88. doi: 10.1161/CIRCULATIONAHA.120.046528

15. Chumachenko PV, Postnov AY, Ivanova AG, Afanasieva OI, Afanasiev MA, Ekta MB, et al. Thoracic aortic aneurysm and factors affecting aortic dissection. *J Pers Med*. (2020) 10(4):153. doi: 10.3390/jpm10040153

16. Davis FM, Daugherty A, Lu HS. Updates of recent aortic aneurysm research. *Arterioscler Thromb Vasc Biol*. (2019) 39(3):e83–90. doi: 10.1161/ATVBAHA.119.312000

17. Lu H, Daugherty A. Aortic aneurysms. *Arterioscler Thromb Vasc Biol*. (2017) 37(6):e59–65. doi: 10.1161/ATVBAHA.117.309578

18. Wu L. The pathogenesis of thoracic aortic aneurysm from hereditary perspective. *Gene*. (2018) 677:77–82. doi: 10.1016/j.gene.2018.07.047

19. Yap ZJ, Sharif M, Bashir M. Is there an immunogenomic difference between thoracic and abdominal aortic aneurysms? *J Card Surg*. (2021) 36(4):1520–30. doi: 10.1111/jocs.15440

20. Creamer TJ, Bramel EE, MacFarlane EG. Insights on the pathogenesis of aneurysm through the study of hereditary aortopathies. *Genes (Basel)*. (2021) 12(2):183. doi: 10.3390/genes12020183

21. Jia H, Kang L, Ma Z, Lu S, Huang B, Wang C, et al. MicroRNAs involve in bicuspid aortic aneurysm: pathogenesis and biomarkers. *J Cardiothorac Surg*. (2021) 16(1):230. doi: 10.1186/s13019-021-01613-9

22. Brosseau C, Colas L, Magnan A, Brouard S. CD9 tetraspanin: a new pathway for the regulation of inflammation? *Front Immunol*. (2018) 9:2316. doi: 10.3389/fimmu.2018.02316

23. Lou Y, Miao J, Li F, Ding J, Wang L. Maternal smoking during pregnancy aggravated muscle phenotype in FHL1$^{-/y}$ offspring mice similar to congenital clubfoot through P2RX7-mediated pyroptosis. *Toxicol Lett*. (2021) 345:54–60. doi: 10.1016/j.toxlet.2021.04.014

24. Weng J, Liao M, Zou S, Bao J, Zhou J, Qu L, et al. Downregulation of FHL1 expression in thoracic aortic dissection: implications in aortic wall remodeling and pathogenesis of thoracic aortic dissection. *Ann Vasc Surg*. (2011) 25(2):240–7. doi: 10.1016/j.avsg.2010.10.001

25. Liu H, Zhang Z, Huang Y, Wei W, Ning S, Li J, et al. Plasma HSP90AA1 predicts the risk of breast cancer onset and distant metastasis. *Front Cell Dev Biol*. (2021) 9:639596. doi: 10.3389/fcell.2021.639596

26. Xiao X, Wang W, Li Y, Yang D, Li X, Shen C, et al. HSP90AA1-mediated autophagy promotes drug resistance in osteosarcoma. *J Exp Clin Cancer Res*. (2018) 37(1):201. doi: 10.1186/s13046-018-0880-6

27. Zuehlke AD, Beebe K, Neckers L, Prince T. Regulation and function of the human HSP90AA1 gene. *Gene*. (2015) 570(1):8–16. doi: 10.1016/j.gene.2015.06.018

28. Ekman C, Site DF, Gottsäter A, Lindblad B, Dahlbäck B. Plasma concentrations of growth arrest specific protein 6 and the soluble form of its tyrosine kinase receptor Axl as markers of large abdominal aortic aneurysms. *Clin Biochem*. (2010) 43(1–2):110–4. doi: 10.1016/j.clinbiochem.2009.07.025

29. Liu X, Xu S, Li Y, Chen Q, Zhang Y, Peng L. Identification of CALU and PALLD as potential biomarkers associated with immune infiltration in heart failure. *Front Cardiovasc Med*. (2021) 8:774755. doi: 10.3389/fcvm.2021.774755

30. Regalado ES, Guo DC, Prakash S, Bensend TA, Flynn K, Estrera A, et al. Aortic disease presentation and outcome associated with ACTA2 mutations. *Circ Cardiovasc Genet*. (2015) 8(3):457–64. doi: 10.1161/CIRCGENETICS.114.000943

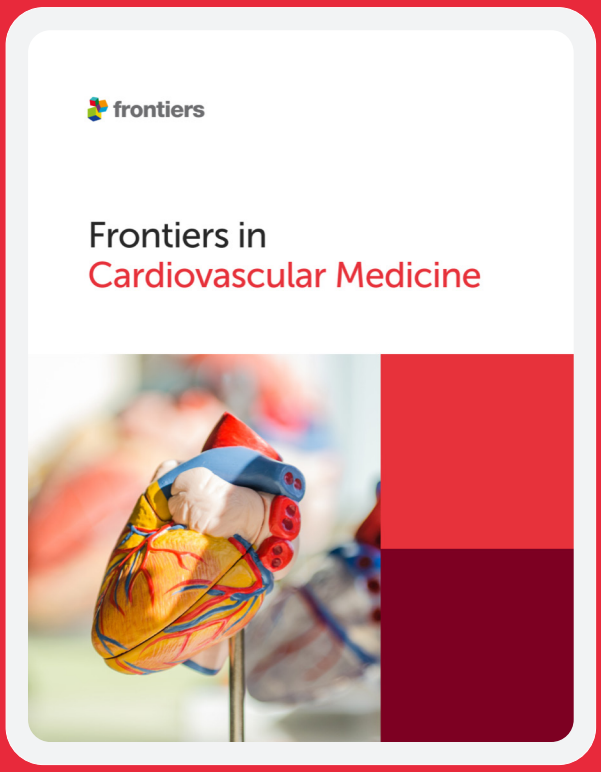
Frontiers in
Cardiovascular Medicine

Innovations and improvements in cardiovascular treatment and practice

Focuses on research that challenges the status quo of cardiovascular care, or facilitates the translation of advances into new therapies and diagnostic tools.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact