



Research article

Transcriptomic analysis of B cells suggests that CD70 and LY9 may be novel features in patients with systemic lupus erythematosus

Qun Liu^{a,b,c,1}, Yiyao Deng^{d,1}, Xiaomin Liu^{b,c}, Ying Zheng^{b,c}, Qinggang Li^{b,c}, Guangyan Cai^b, Zhe Feng^{b,*}, Xiangmei Chen^{a,b,c,2,**}^a School of Medicine, Nankai University, Tianjin, 300071, China^b Department of Nephrology, The First Medical Center, Chinese PLA General Hospital, Medical School of Chinese PLA, Chinese PLA Institute of Nephrology, State Key Laboratory of Kidney Diseases, National Clinical Research Center for Kidney Diseases, Beijing, 100853, China^c Haihe Laboratory of Cell Ecosystem, Tianjin, 300020, China^d Department of Nephrology, Guizhou Provincial People's Hospital, 83, Zhongshan Road, Nanming District, Guiyang, 550002, Guizhou, China

ARTICLE INFO

Keywords:

Systemic lupus erythematosus

Transcriptomic analysis

B cells

CD70

LY9

ABSTRACT

Dysfunction of B-cell subsets is critical in the development of systemic lupus erythematosus (SLE). There is a great diversity of B-lineage cells, and their features and functions in SLE need to be clarified. In this study, we analyzed single-cell RNA sequencing (scRNA-seq) data from peripheral blood mononuclear cells (PBMCs) and bulk transcriptomic data of isolated B-cell subsets from patients with SLE and healthy controls (HCs). We performed scRNA-seq analysis focused on the diversity of B-cell subsets and identified a subset of antigen-presenting B cells in SLE patients that highly expressed ITGAX. A list of marker genes of each B-cell subset in patients with SLE was also identified. Comparison of bulk transcriptomic data of isolated B-cell subpopulations between SLE patients and HCs revealed the upregulated differentially expressed genes (DEGs) for each B-cell subpopulation in SLE. Common genes identified using these two methods were considered to be upregulated marker genes of B cells in SLE. The scRNA-seq data of SLE patients and HCs revealed that CD70 and LY9 were overexpressed in B cells vs. other cell types from SLE patients, and this pattern was validated by RT-qPCR. Because CD70 is the cellular ligand of CD27, previous studies on CD70 have focused mainly on T cells from SLE patients. LY9 appears to have different functions in mice and humans: its expression is decreased in lupus-prone mice but is increased in T cells and some B-cell subpopulations in SLE patients. Here, we describe the overexpression of two costimulatory molecules, CD70 and LY9, which may be a novel feature of B cells in SLE patients.

1. Introduction

Systemic lupus erythematosus (SLE) is a heterogeneous autoimmune disease that can induce systemic organ damage and often leads to a poor prognosis. Genetic susceptibility, environmental factors, sex, innate immune cells and lymphocyte disturbances all contribute to the development of SLE. Multiple factors trigger the loss of immune tolerance of endogenous nuclear material and the

* Corresponding author.

** Corresponding author.

E-mail addresses: zhezhe_4025@126.com (Z. Feng), xmchen301@126.com (X. Chen).¹ These authors contributed equally.² Lead contact.<https://doi.org/10.1016/j.heliyon.2023.e15684>

Received 31 August 2022; Received in revised form 16 April 2023; Accepted 18 April 2023

Available online 23 April 2023

2405-8440/© 2023 Published by Elsevier Ltd.

This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

production of autoantibodies [1]. In the pathogenesis of SLE, the imbalance of B-cell homeostasis and dysregulation of T-cell co-stimulation contributes to aberrant B-cell maturation and autoantibody production [2]. A deep understanding of immune cells is essential for the treatment of SLE and prediction of its progression.

With the development of high-throughput sequencing technologies, transcriptomic analysis has emerged as a method to decode diseases. Both bulk transcriptomic analysis and single-cell RNA sequencing (scRNA-seq) have been used to detect the heterogeneity of SLE. Although a high interferon-stimulated gene (ISG^{hi}) signature is a major feature of SLE, scRNA-seq analysis of peripheral blood from SLE patients revealed that cells with ISG^{hi} signature was predominantly a few transcriptionally defined subpopulations, including a ISG^{hi} DN2 B-cell subset and ISG^{hi} plasma cells, which were expanded in patients with high disease activity [3]. However, some established cell subpopulations failed to be detected by scRNA-seq, mainly due to sparse expression at the single-cell level. To address this issue, a large-scale bulk transcriptome study of peripheral blood from SLE patients was recently conducted, which included data from 27 immune cell types, covering virtually all types of immune cells [4]. The study focused on profiling cell-type-specific transcriptome perturbation patterns during disease onset and progression phases, where transcriptome perturbation of naïve CD4/8+ T cells was dominant in the disease-state pattern, plasmablasts were dominant in the disease-activity pattern and neutrophil-lineage cells were dominant in the shared pattern. These observations suggest that a combination of these two types of transcriptomic analysis of immune cells may provide a better understanding of SLE. B-lineage cells are the most important cells in humoral immunity and are equally essential for cellular immunity [5,6]. Besides their role in antibody production, they are also involved in T-cell activation through antigen presentation, co-stimulation and cytokine production [7]. B cells are considered “professional” antigen-presenting cell (APC), and B-cell-specific MHC II deletion in a lupus-prone mouse model (MRL/lpr) resulted in reduced numbers and impaired function of activated Th cells and substantially ameliorated organ damage [8]. Recent evidence suggests that upon activation of B cells by virus-like antigens, B cells act as predominant APCs and promote the differentiation of naïve T cells to T follicular helper cells (TFHs); this finding provides new insight into the antigen-presenting function of B cells in autoimmune diseases [9]. In addition, studies have revealed that autoreactive CD11c + CD27⁻ IgD⁻ CXCR5⁻ B cells expand in lupus patients and that excessive CD11c + Tbet + B cells promote aberrant TFH differentiation and autoantibody production in lupus, indicating that CD11c + B cells act as a

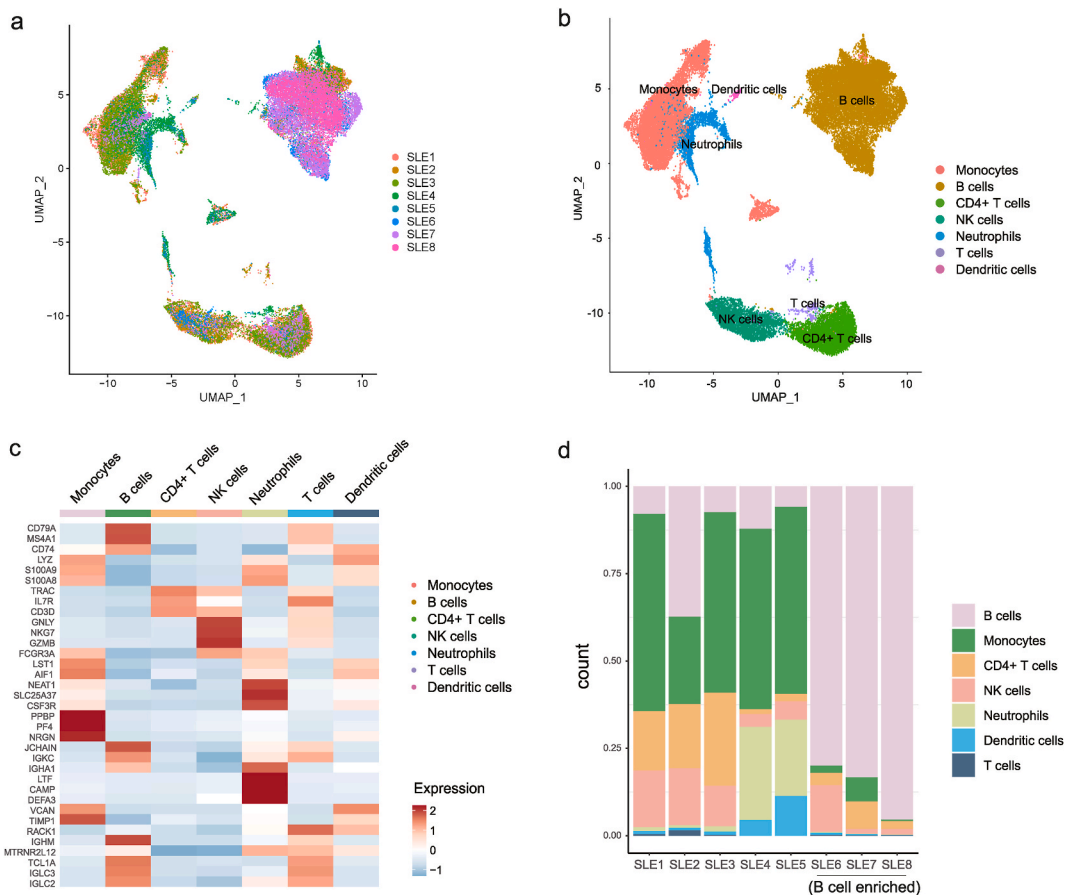


Fig. 1. Integrated analysis of the SLE PBMC scRNA-seq datasets. (a) A UMAP plot representing 8 SLE patients, showing that the B cells from SLE6, SLE7 and SLE8 were specifically enriched. (b) A UMAP plot representing the 7 clusters across 42,732 PBMCs from 8 SLE patients. (c) Heatmap showing the expression of the top 5 DEGs in each cell type in SLE patients. (d) Bar plots showing the proportions of cell types in each sample in SLE patients.

pathogenic subpopulation in lupus [10,11]. In light of the diversity of B-lineage cell features and functions, the roles of B cells in SLE need to be clarified.

In this study, we analyzed scRNA-seq data of peripheral blood mononuclear cells (PBMCs) in combination with bulk transcriptomic data of isolated B-cell subsets from SLE patients to profile B cells in SLE. Using transcriptomic analysis, we identified the features of B-cell subsets and the upregulated marker genes in the B cells of lupus patients. Furthermore, real-time quantitative polymerase chain reaction (RT-qPCR) was used to detect the relative expression level of mRNA in the B cells to verify the upregulated marker genes from SLE patients. Two costimulatory molecules, CD70 and LY9, were confirmed to be overexpressed in the B cells of lupus patients.

2. Results

2.1. Integrated analysis of single-cell transcriptomic datasets of PBMCs from patients with SLE

To characterize the peripheral blood immune profiles of SLE, single-cell transcriptomic data of PBMCs from three GEO datasets representing 8 SLE patients were analyzed. After raw data processing and filtering, 42,732 cells were reserved for further analyses (Fig. S1). Using graph-based clustering, the cells from the 8 SLE patients were partitioned into 7 clusters (Fig. 1a and b). Cell-type annotation with the “SingleR” package of R identified the following 7 cell clusters: B cells, monocytes, CD4⁺ T cells, natural killer (NK) cells, neutrophils, dendritic cells (DCs) and T cells (Fig. 1b). The heatmap of the top 5 marker genes for each cluster revealed a distinct separation of each cluster (Fig. 1c). The cell-type marker gene scores calculated using the “SingleR” package verified the annotation of each cluster (Fig. S1d). There is great heterogeneity in the proportion of each cell lineage among individuals, although the PBMCs of individuals SLE6, SLE7 and SLE8 had been specifically enriched in B cells prior to sequencing (Fig. 1d). Next, to clarify the features of B-lineage cells, we extracted B-cell clusters for downstream analysis.

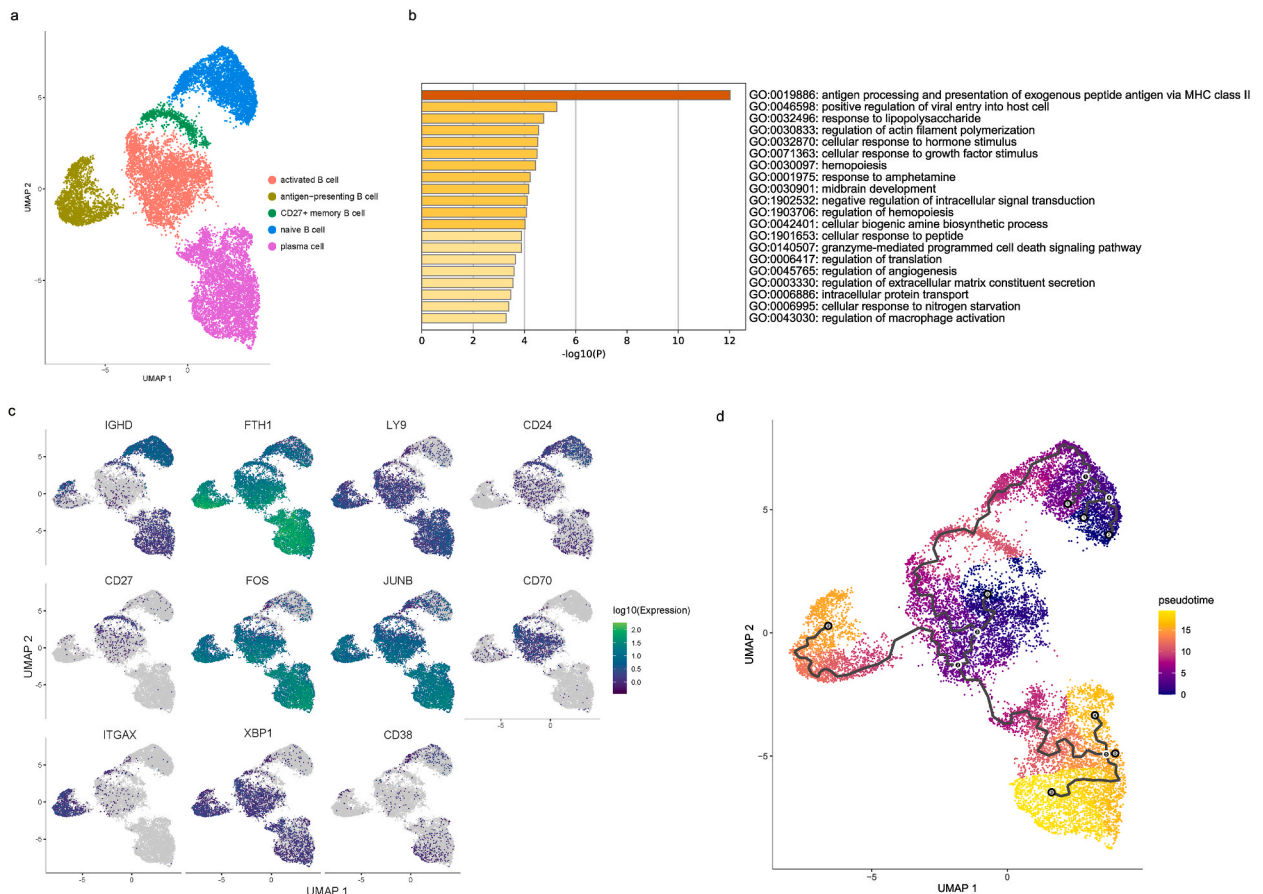


Fig. 2. Reclustering and analysis of the B cells of SLE patients. (a) A UMAP plot representing the 5 subclusters of B cells. (b) DEG enrichment analysis of antigen-presenting B-cell subclusters by Metascape. (c) UMAP plots showing the expression of some feature genes in B-cell subclusters. (d) Monocle pseudotime trajectory analysis revealing the arrangement of different subclusters.

2.2. Differentiated features of different B-cell subclusters

Based on their differentially expressed genes (DEGs), B cells were divided into five subclusters: activated B cells, antigen-presenting B cells, CD27⁺ memory B cells, naïve B cells and plasma cells (Fig. 2a and Fig. S2a). The top 5 marker genes for antigen-presenting B cells were CD74, MT-CO1, RPL12, EIF1, and HLA-B (Table S1). As shown in Fig. 2b, DEG pathway enrichment analysis of the antigen-presenting B-cell subset showed that the top GO term was antigen processing and presentation of the exogenous peptide antigen via MHC class II (GO: 0019886), which confirmed the annotation of antigen-presenting B cells.

We found that some feature genes, including some canonical markers of B cells, such as IGHD, CD24 and CD27, clearly delineated B-cell subclusters (Fig. 2c). Previous studies have shown that CD24, CD27 and IgD can be used to illustrate the maturation process of B cells [12,13]. Consistent with previous studies, our results revealed that IGHD and CD24 were mainly expressed on naïve B cells and that CD27 was mainly expressed on CD27⁺ memory B cells (Fig. 2c). Interestingly, ITGAX, the gene encoding CD11c, was a marker found on autoimmune-associated B cells and was mostly expressed by antigen-presenting B cells (Fig. 2c).

Moreover, pseudotime trajectory analysis revealed that the different clusters were aligned in line with their expected differentiation pathways. Naïve B cells were distributed at the start of the trajectory, followed by activated B cells and CD27⁺ memory B cells. Antigen-presenting B cells and plasma cells were distributed at the end of the trajectory (Fig. 2d and Fig. S2b).

2.3. Identification of feature genes by combining the single-cell and bulk transcriptomic data of B cells from SLE patients

Bulk transcriptomic data of naïve B cells, memory B cells and plasmablasts of healthy donors and SLE patients from GSE156751 were analyzed. The isolated B-cell subpopulations were classified based on surface markers. All the data were found to be suitable for comparison (Figs. S3a and S3b), and significant DEGs were found between SLE patients and healthy donors (Fig. S3c). A Venn diagram was generated from the marker genes of the B-cell subset from the scRNA-seq data and the upregulated DEGs in B cells from the bulk transcriptomic data, and some upregulated genes, including GPR137B, NKG7, CD70 and ITGAX, were found in both the bulk and single-cell datasets (Fig. 3 and Table S2).

In the single-cell sequencing data, all of these common genes in SLE were expressed in antigen-presenting B cells with the exception of CD70, which was expressed in activated B cells, suggesting that antigen-presenting B cells may play an important pathogenic role in SLE (Table S1). However, the bulk transcriptomic analysis showed that CD70 and ITGAX were overexpressed in naïve B cells, whereas GPR137B and NKG7 were overexpressed in memory B cells. These results suggest that cell classification based on transcriptomic profiles may reflect the cell features better than that based on surface markers.

2.4. B cells activated and overexpressed CD70 and LY9 in SLE patients compared to healthy controls (HCs)

ScRNA-seq data of HCs were obtained from two GEO datasets representing three healthy people. After data processing, the cells were partitioned into 7 clusters. The “SingleR” R package was used for unbiased automatic cell-type annotation, and the heatmap of the top 5 marker genes revealed distinct separation of each cluster (Fig. 4a and Fig. S4a). The cell-type marker gene scores, which were obtained using the “SingleR” package, verified the annotation of each cluster (Fig. S4b). The expression of canonical B-cell marker genes and the common upregulated genes, i.e., genes identified as upregulated in SLE in both bulk and single-cell transcriptomic datasets, was analyzed (Fig. 4b). Among these genes, four genes, FTH1, LY9, GPR137B and CD70, were significantly upregulated in B cells compared with other cell types in patients with SLE but not in HCs. However, IGHD and CD24, markers of unswitched B cells and

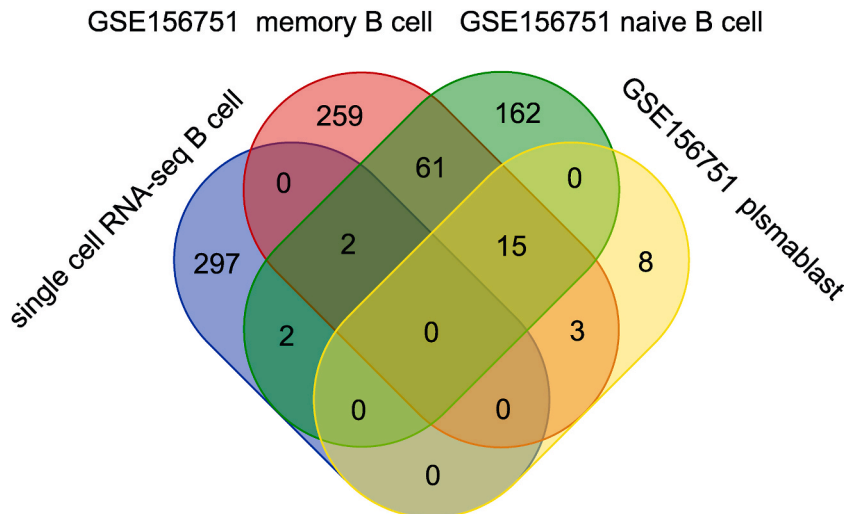


Fig. 3. Venn diagram revealing the common upregulated genes between the bulk and single-cell transcriptomic data of B cells from SLE patients.

resting B cells, were underexpressed in SLE patients. The protein–protein interaction (PPI) network showed the connections among the overexpressed genes (Fig. S4c).

To verify the overexpression of the four genes significantly upregulated in B cells in patients with SLE, RT–qPCR was performed, and the results confirmed the upregulation of CD70 and LY9 in SLE patients ($p < 0.05$). In addition, there was a trend for higher expression of FTH1 and GPR137B in SLE patients than in HCs, but the difference was not statistically significant, which probably due to the low sample size (Fig. 4c).

3. Discussion

SLE is a heterogeneous autoimmune disease in which all types of immune cells, and especially B-cell subpopulations, may contribute to its pathogenesis [14]. Dysfunction of pathogenic and regulatory B-cell subsets is critical in the development of SLE [15]. With the development of high-throughput sequencing technologies, many researchers have used transcriptomic analysis to characterize the B-cell profile of SLE patients. Akita, Yasaka [16] classified B cells into naïve B cells, memory B cells and plasmablasts according to surface markers and identified the unique biological characteristics of SLE plasmablasts. Since different cell subsets play different roles in the pathogenesis of SLE, it is better to classify the cells according to their biological features. However, cell classification and isolation are usually based on surface markers. ScRNA-seq has emerged with the development of single-cell genomics technologies and shown an unbiased ability to identify immune cell features and functions [17]. In our current study, we focused on the diversity of B cells and identified five subsets of B cells in SLE patients. We further analyzed the bulk transcriptomic data of classified B cells and identified some common upregulated genes by the both bulk transcriptomic analysis and scRNA-seq analysis, including GPR137B, NKG7, CD70 and ITGAX. CD11c, encoded by ITGAX, is a marker of autoimmune-associated B cells, and CD11c/ITGAX-expressing B cells have been found to be expanded in lupus patients [11,18]. CD11c + B cells have been a focus of studies on B cells in lupus over the past years. Transcriptional and functional analysis showed that CD11c^{hi} B cells were a distinct subpopulation of B cells and that CD11c^{hi} double negative B cells in SLE might be precursors of autoantibody producing cells [11,18]. In the presence of IL-21, CD11c^{hi}T-bet + B cells can differentiate into autoantibody-secreting cells [11,19]. Additionally, CD11c + B cells have been found to have antigen-presenting function, and excessive accumulation of CD11c + Tbet + B cells was found to dysregulate TFH cell differentiation in the context of SLE based on the potent antigen-presenting function of these cells [10,20]. Consistent with previous studies, our current study identified a subset of antigen-presenting B cells that highly express ITGAX in SLE patients. Moreover, most of the common upregulated genes in our study, particularly ITGAX, were marker genes of antigen-presenting

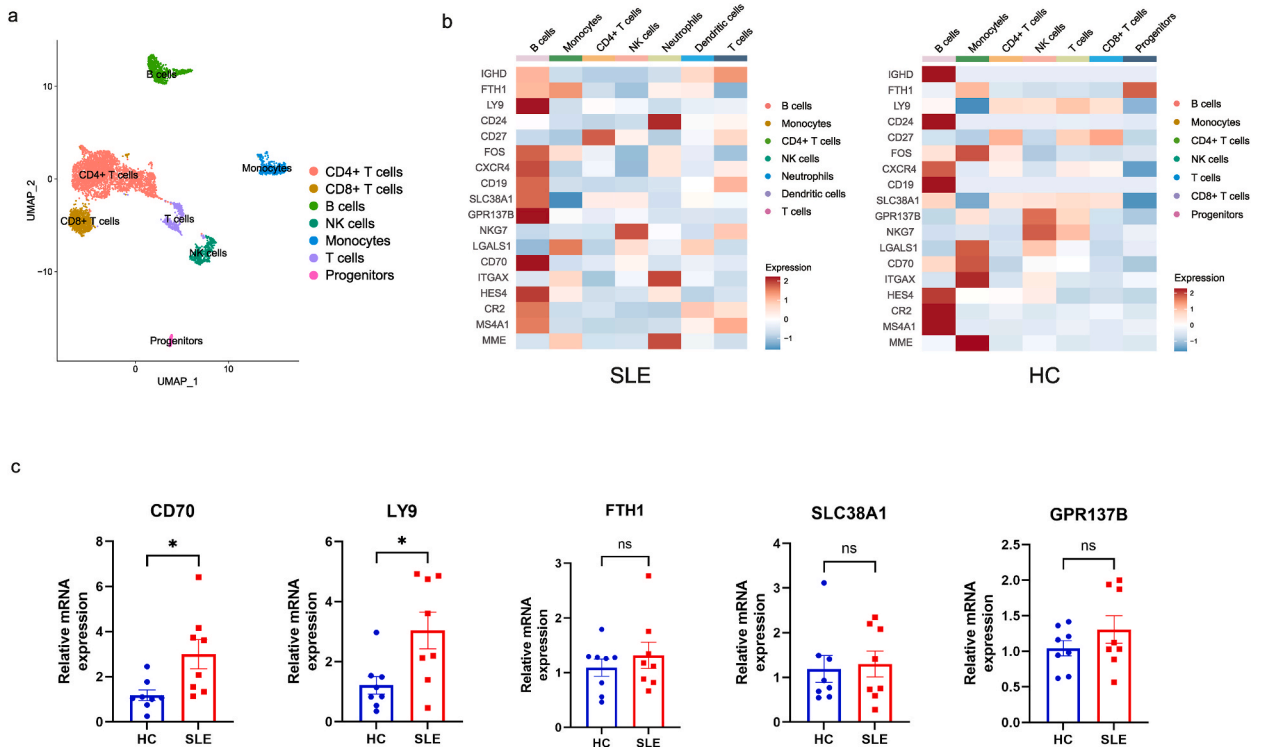


Fig. 4. Overexpressed genes in SLE patients compared to HCs. (a) A UMAP plot representing the 7 clusters of PBMCs from 3 healthy donors. (b) Heatmap showing the expression of some canonical B-cell marker genes and common upregulated genes, i.e., upregulated genes in SLE identified by both bulk transcriptomic analysis and single-cell sequencing, in SLE patients (left) and HCs (right). (c) The mRNA levels of genes in the HC and SLE groups that are overexpressed in SLE ($n = 8$). The data are expressed as the means \pm SEMs. * $P < 0.05$.

B cells, suggesting that antigen-presenting B cells may play an important pathogenic role in SLE. However, how the ITGAX-expressing antigen-presenting B cells contribute to SLE needs further investigation. Most of these genes were upregulated in antigen-presenting B cells according to the scRNA-seq data, but when the cells were classified based on surface markers, these genes were found to be overexpressed in naïve B cells or memory B cells because the antigen-presenting B cells could not be isolated by surface markers. In our study, we verified the unbiased ability of scRNA-seq to reveal the features of immune cells and identified some common genes upregulated in the B cells of SLE patients using two types of transcriptomics-based technology.

Compared with their expression in B cells from healthy donors, IGHD and CD24 were underexpressed in B cells from SLE patients, while FTH1, LY9, GPR137B and CD70 were overexpressed in SLE. Since IGHD and CD24 are markers of unswitched B cells and resting B cells, respectively, their low expression revealed B-cell activation in SLE. Next, we verified gene overexpression by RT-qPCR and confirmed the overexpression of CD70 and LY9. In addition, we analyzed the correlations of age, SLEDAI-2K and some serological indexes with CD70 and LY9 expression. However, no significant correlations were found, which may have been due to the low sample size and the heterogeneity of lupus. As the cellular ligand of CD27, CD70 is a costimulatory molecule transiently expressed on activated T cells, B cells, and DCs [21,22]. And CD70 is sometimes constitutively expressed on the surface of many B lymphocyte malignancies [23]. In a previous study, 6–30% of the B cells in the peripheral blood of healthy donors were found to be CD70⁺ B cells, and the vast majority of these also expressed CD27, indicating that CD70 is also a marker of mature B cells. Based on previous studies, although CD70 appears to be unessential for lymphocyte development, it plays unique roles in lymphocyte activation, germinal center formation, and effector cell differentiation [24]. In a previous study, CD70 expression in B cells was found to be strongly correlated with their ability to secrete immunoglobulin in response to T-cell-derived differentiation signals [25]. In a mouse model, constitutive transgene expression of CD70 in B lymphocytes was shown to lead to expansion and differentiation of T cells into IFN- γ producing effector lymphocytes [26]. In many studies of SLE patients, T cells have been reported to overexpress CD70, and this increased expression might contribute to B-cell costimulation and immunoglobulin overproduction [27–29]. Although most studies in SLE patients have been focused on CD70 overexpression in T cells, our present study confirmed CD70 overexpression in B cells. CD70 overexpression in B cells indicates the hyperactivation of B cells in SLE, but the role of CD70⁺ B-cell interaction with CD27 in the pathogenesis of SLE remains to be clarified.

The signaling lymphocytic activation molecule family (SLAMF) represents a complex family of surface coreceptors that are expressed on cells of hematopoietic origin and play important roles in immunomodulation [30]. LY9, a member of SLAMF, also known as CD229 or SLAMF3, is mainly expressed on the surfaces of B and T cells and acts as a cosignaling molecule that regulates lymphocyte homeostasis and activation. Defects in the LY9 gene in mice lead to autoantibody production, a process that initiates SLE syndrome [31]. Similarly, LY9, along with other BCR signaling regulators, may serve as a potential causative gene for B-cell hyperactivation in lupus-prone mice by reducing its mRNA synthesis [32], indicating that LY9 might be a negative regulator of B-cell hyperactivation and autoantibody production in mice. However, altered LY9 expression was also found in the lymphocytes of autoimmune disease patients, in contrast to the findings in mice. Previous studies have revealed increased LY9 expression on the surfaces of naïve T cells in SLE patients [33]. Upregulated LY9 and LY108 expression in T cells from SLE patients was found to serve as a major costimulus for Th17 differentiation and IL-17 production [34]. Using single-cell mass cytometry to identify the expression of all SLAMF receptors on SLE PBMCs, Humbel, Bellanger [35] found that naïve B cells coexpressing SLAMF1 and LY9 and switch memory B cells coexpressing SLAMF1, LY9, SLAMF5 and SLAMF6 were significantly increased in SLE patients. However, the expression of LY9 in total B cells showed no significant difference between SLE patients and HCs. An increase in LY9⁺ B cells has been reported in rheumatoid arthritis synovium, indicating B-cell activation [36]. In our study, we first verified increased LY9 mRNA expression in the B cells of SLE patients. Our results indicate that LY9 in B cells might have different functions between humans and mice: although LY9 mRNA expression has been found to be decreased in lupus-prone mice, it was increased in B cells of SLE patients in our study. The function of LY9 in human B cells needs further investigation.

In conclusion, we described the features of B cells in SLE patients by both scRNA-seq and bulk transcriptomic analysis. We verified the unbiased ability of scRNA-seq to explore the features of immune cells and confirmed some common upregulated genes. The markers of unswitched B cells and resting B cells were underexpressed, revealing B-cell activation in SLE. In addition, CD70 and LY9, which were previously reported to be upregulated in T cells of SLE patients, were confirmed in our study to also be overexpressed in B cells; CD70 overexpression and LY9 overexpression may be novel features of B cells in SLE patients. Both CD70 and LY9 are costimulatory molecules, and their functions of antigen presentation and costimulation in B cells and T cells need further investigation.

4. Limitations of the study

The transcriptomics data in this study are obtained from different datasets in GEO database, which may have introduced bias. Considering the multiple data sources and the heterogeneity of SLE patients and HCs, we analyzed scRNA-seq data of SLE and HCs separately. The lack of integration is a limitation because it prevented us from directly comparing gene expression, as determined from the scRNA-seq data, between the two groups at the beginning. However, after finding the genes of interest, we compared their expression levels among the different cell types in each of the two groups separately and used RT-qPCR to verify relative gene expression using CD19⁺ B cells from 8 SLE patients and 8 healthy donors.

We identified a group of antigen-presenting B cells that highly expressed ITGAX in SLE; however, we didn't assess the functions of this B-cell subset. As this was a basic transcriptome-based study, no protein confirmation of the markers was performed. Our study described the transcriptomic features of B cells in SLE patients and verified the overexpression of CD70 and LY9 by RT-qPCR. However, the functions of CD70 and LY9 proteins in B cells of SLE patients remain to be investigated. In addition, the results may be biased due to the small sample size.

Author contribution statement

Qun Liu, Yiyao Deng: Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper.

Xiaomin Liu, Ying Zheng, Qinggang Li, Guangyan Cai: Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data.

Zhe Feng, Xiangmei Chen: Conceived and designed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data.

Funding statement

Xiangmei Chen was supported by National Natural Science Foundation of China { 32435895 }, Haihe Laboratory of Cell Ecosystem Innovation Fund { 22HHXBSS00002 }, National Natural Science Foundation of China { 32141005 }.

Yiyao Deng was supported by Basic Research Plan of Guizhou Province in 2023 (Natural Science Project) { QKH-ZK[2023] General 218 }, Talent Fund of Guizhou provincial People's Hospital { [2022]-1 }.

Data availability statement

Data associated with this study has been deposited at Both single-cell and bulk transcriptomic data were obtained from the GEO database. For the SLE group, three scRNA-seq datasets, GSE142016, GSE162577 and GSE163121, were enrolled based on sample type (only the expression profiles of peripheral immune cells were considered). For the HC group, two scRNA-seq datasets, GSE142016 and GSE162577, were enrolled.

For bulk transcriptomic data, the gene expression profiles of naïve B cells, memory B cells and plasmablasts were obtained from GSE156751. GSE156751 was based on GPL17077 Agilent-039494 SurePrint G3 Human GE v2 8 × 60K Microarray 039381, containing 12 SLE and 12 healthy control samples.

Star methods

Patients and sample preparation

The study was approved by the Ethics Committee of the Chinese People's Liberation Army General Hospital (No. S2019-095-01). Blood samples were collected from patients who met the criteria established by the American College of Rheumatology for a diagnosis of SLE. Blood samples from age- and sex-matched HCs without any autoimmune or inflammatory diseases were also obtained. A total of 8 SLE patients and 8 HCs were recruited from December 2021 to April 2022. Clinical information of SLE patients and healthy donors is provided in [Supplemental Table 3](#). Fresh blood samples were collected, and PBMCs were isolated by Ficoll density gradient centrifugation. CD19⁺ B cells were isolated from PBMCs using CD19 MicroBeads according to manufacturer's instructions. Then, RNA was extracted from CD19⁺ B cells using TRIzol (Invitrogen, Carlsbad, CA, USA) and stored at −80 °C for subsequent analysis.

Data source

Both single-cell and bulk transcriptomic data were obtained from the GEO database. For the SLE group, three scRNA-seq datasets, GSE142016, GSE162577 and GSE163121, were retrieved based on the sample type (only the expression profiles of peripheral immune cells were considered). For SLE6, SLE7 and SLE8 from GSE163121, B cells were isolated from PBMCs by the EasySep Human Pan-B Cell Enrichment Kit and then subjected to scRNA-seq. For the HC group, two scRNA-seq datasets, GSE163121 and GSE162577, were retrieved.

For the bulk transcriptomic data, the gene expression profiles of naïve B cells, memory B cells and plasmablasts were obtained from GSE156751. GSE156751 was based on GPL17077 Agilent-039494 SurePrint G3 Human GE v2 8 × 60K Microarray 039381 and contained 4 SLE and 4 HC samples.

ScRNA-seq data quality control

Considering that the SLE and HC samples came from different datasets and that integrated analyses may introduce bias, the SLE group and HC group were analyzed separately. We used R (version 4.1.3) and the Seurat R package (version 4.1.0) for follow up analyses [37]. The MergeSeurat function was used to integrate multiple datasets. Cells with <200 and >2000 genes, a mitochondrial gene percentage of >10% and a hemoglobin gene percentage of >3% were filtered according to the median number of genes and the percentages of mitochondrial genes and hemoglobin genes. (Fig. S1A). The raw counts of each sample were normalized using NormalizeData function in Seurat. After normalization, all highly variable genes in each cell were identified using FindVariableFeatures function in Seurat (Fig. S1B). Subsequently, principal component (PC) analysis with variable genes was used as the input, and significant PCs were identified based on the jackstraw function. Twenty PCs were selected as the input for uniform manifold approximation and projection (UMAP).

Batch effect correction and cell type identification

Since the data came from different samples, batch correction was performed using “Harmony” package (version 1.0) [38] to avoid batch effect interfering with downstream analysis (Fig. S1C). Cells were clustered by the FindClusters function, and marker genes of each corresponding cluster were identified by the function FindAllMarkers. Genes with a log-fold change threshold >0.25 and false discovery rate (FDR) < 0.01 were identified as marker genes or significant DEGs.

The package “SingleR” is a computational tool for unbiased cell annotation that assigns a cellular identity to single-cell transcriptomes by evaluating the global similarity of gene expression between each cell and the reference datasets [39]. In our study, two reference datasets of purified immune cells, Blueprint/ENCODE and Monaco Immune Cell Data, were used. The Spearman correlation between the single-cell data and reference datasets was calculated, and cell labels were assigned to each single cell from the reference datasets. Cell types were eventually identified by manually checking the cell signature marker genes and cell labels from “SingleR”. After clustering and cell type identification of total PBMCs, B cells were extracted for further analysis. Due to the inconsistency and ambiguity remaining with the SingleR automatic assignment, B-cell subclusters were identified based on the CellMarker database or published articles.

Pseudotime analysis of B cells

To investigate the potential differentiated trajectory of each B-cell subcluster, we performed pseudotime trajectory analysis by using the Monocle package (version 3.0). Based on biological knowledge, we chose naïve B cells as the starting point of the trajectory. Then, RNA velocity was calculated and visualized on the first two PCs.

Comparison with the bulk gene expression dataset

Bulk transcriptomic data of naïve B cells, memory B cells and plasmablasts of healthy donors and SLE patients were obtained from GSE156751. The GEO2R (<https://www.ncbi.nlm.nih.gov/geo/geo2r/>) online tool was used to analyze the DEGs between the two groups. The adjusted p value (p_{adj}) was calculated by the Benjamini and Hochberg method. The following cutoff criteria were used: p_{adj} value < 0.05 and $|\log FC| > 2$. A Venn diagram was generated to identify the common upregulated genes in B cells, i.e., the upregulated genes shared between the scRNA-seq and bulk gene expression data from B cells.

PPI network construction and gene pathway enrichment analysis

The STRING database (version 11.5) was used to detect PPIs of the common upregulated genes. Gene pathway enrichment analysis was performed by Metascape [40].

RT-qPCR analysis of candidate genes

RNA that had been extracted with TRIzol and stored at -80°C was thawed and cDNA was then synthesized using the ProtoScript II First Strand cDNA Synthesis Kit (New England Biolab) according to manufacturer’s instructions. RT-qPCR was performed using SYBR Select Master Mix (Life Technologies) and an RT-qPCR detection system (Bio-Rad, Hercules, CA, USA). The following primers were synthesized: CD70 forward GCTTTGGTCCCATTGGTCG, CD70 reverse CGTCCCACCAAGTGACTC, FTH1 forward AAGCTGCA-GAACCAACGAGG, FTH1 reverse AGTCACACAAATGGGGGTCATT, GPR137B forward CTCACCGTCTACACCG, GPR137B reverse GCCGCCAGGAAGTCTTTGA, LY9 forward CTTCTGCTCATGGGACTAAGA, and LY9 reverse GCCCAGGTAGCTTTTGACCA. The expression of each target gene was normalized to GAPDH, and relative expression differences were calculated by the $2^{-\Delta\Delta\text{CT}}$ method.

Statistical analysis

RT-qPCR data were analyzed and plotted using GraphPad Prism 9.0 software. Statistical analysis was performed using unpaired Student’s t-test. The means, SEMs, and levels of statistical significance are shown in figures. $P < 0.05$ was considered to indicate a statistically significant difference.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.heliyon.2023.e15684>.

References

- [1] G.C. Tsokos, Autoimmunity and organ damage in systemic lupus erythematosus, *Nat. Immunol.* 21 (6) (2020) 605–614.
- [2] D.Y.H. Yap, T.M. Chan, B cell abnormalities in systemic lupus erythematosus and lupus nephritis-role in pathogenesis and effect of immunosuppressive treatments, *Int. J. Mol. Sci.* 20 (24) (2019).
- [3] D. Nehar-Belaïd, et al., Mapping systemic lupus erythematosus heterogeneity at the single-cell level, *Nat. Immunol.* 21 (9) (2020) 1094–1106.
- [4] M. Nakano, et al., Distinct transcriptome architectures underlying lupus establishment and exacerbation, *Cell* 185 (18) (2022) 3375–3389.e21.
- [5] T.W. LeBien, T.F. Tedder, B lymphocytes: how they develop and function, *Blood* 112 (5) (2008) 1570–1580.
- [6] W. Hoffman, F.G. Lakkis, G. Chalasani, B cells, antibodies, and more, *Clin. J. Am. Soc. Nephrol.* 11 (1) (2016) 137–154.
- [7] D. Ghosh, et al., New insights into B cells as antigen presenting cells, *Curr. Opin. Immunol.* 70 (2021) 129–137.
- [8] J.R. Giles, et al., B cell-specific MHC class II deletion reveals multiple nonredundant roles for B cell antigen presentation in murine lupus, *J. Immunol.* 195 (6) (2015) 2571–2579.
- [9] Z. Hua, B. Hou, The role of B cell antigen presentation in the initiation of CD4⁺ T cell response, *Immunol. Rev.* 296 (1) (2020) 24–35.
- [10] W. Zhang, et al., Excessive CD11c(+)Tbet(+) B cells promote aberrant T(FH) differentiation and affinity-based germinal center selection in lupus, *Proc. Natl. Acad. Sci. U. S. A.* 116 (37) (2019) 18550–18560.
- [11] S.A. Jenks, et al., Distinct effector B cells induced by unregulated toll-like receptor 7 contribute to pathogenic responses in systemic lupus erythematosus, *Immunity* 49 (4) (2018) 725–739.e6.
- [12] A.M. Vale, H.W. Schroeder Jr., Clinical consequences of defects in B-cell development, *J. Allergy Clin. Immunol.* 125 (4) (2010) 778–787.
- [13] R. Feng, et al., Comparison of the deep immune profiling of B cell subsets between healthy adults and Sjögren's syndrome, *Ann. Med.* 54 (1) (2022) 472–483.
- [14] V.R. Moulton, et al., Pathogenesis of human systemic lupus erythematosus: a cellular perspective, *Trends Mol. Med.* 23 (7) (2017) 615–635.
- [15] K. Ma, et al., Multiple functions of B cells in the pathogenesis of systemic lupus erythematosus, *Int. J. Mol. Sci.* 20 (23) (2019) 6021.
- [16] K. Akita, et al., Interferon α enhances B cell activation associated with FOXM1 induction: potential novel therapeutic strategy for targeting the plasmablasts of systemic lupus erythematosus, *Front. Immunol.* 11 (2020), 498703.
- [17] M.J.T. Stubbington, et al., Single-cell transcriptomics to explore the immune system in health and disease, *Science* 358 (6359) (2017) 58–63.
- [18] R.W. Maul, et al., Transcriptome and IgH repertoire analyses show that CD11c(hi) B cells are a distinct population with similarity to B cells arising in autoimmunity and infection, *Front. Immunol.* 12 (2021), 649458.
- [19] S. Wang, et al., IL-21 drives expansion and plasma cell differentiation of autoreactive CD11c(hi)Tbet(+) B cells in SLE, *Nat. Commun.* 9 (1) (2018) 1758.
- [20] A.V. Rubtsov, CD11c-Expressing B cells are located at the T cell/B cell border in spleen and are potent APCs, *J. Immunol.* 195 (1) (2015) 71–79.
- [21] T. Kobata, et al., CD27-CD70 interactions regulate B-cell activation by T cells, *Proc. Natl. Acad. Sci. U.S.A.* 92 (24) (1995) 11249–11253.
- [22] M.A. Nolte, et al., Timing and tuning of CD27-CD70 interactions: the impact of signal strength in setting the balance between adaptive responses and immunopathology, *Immunol. Rev.* 229 (1) (2009) 216–231.
- [23] B.K. Han, N.J. Olsen, A. Bottaro, The CD27-CD70 pathway and pathogenesis of autoimmune disease, *Semin. Arthritis Rheum.* 45 (4) (2016) 496–501.
- [24] H. Abolhassani, Specific immune response and cytokine production in CD70 deficiency, *Front. Pediatr.* 9 (2021), 615724.
- [25] S.M. Lens, et al., Phenotype and function of human B cells expressing CD70 (CD27 ligand), *Eur. J. Immunol.* 26 (12) (1996) 2964–2971.
- [26] R. Arens, et al., Constitutive CD27/CD70 interaction induces expansion of effector-type T cells and results in IFN γ -mediated B cell depletion, *Immunity* 15 (5) (2001) 801–812.
- [27] K. Oelke, et al., Overexpression of CD70 and overstimulation of IgG synthesis by lupus T cells and T cells treated with DNA methylation inhibitors, *Arthritis Rheum.* 50 (6) (2004) 1850–1860.
- [28] M. Zhao, et al., Epigenetics and SLE: RFX1 downregulation causes CD11a and CD70 overexpression by altering epigenetic modifications in lupus CD4⁺ T cells, *J. Autoimmun.* 35 (1) (2010) 58–69.
- [29] W. Liao, et al., Down-regulation of MBD4 contributes to hypomethylation and overexpression of CD70 in CD4⁺ T cells in systemic lupus erythematosus, *Clin. Epigenet.* 9 (2017) 104.
- [30] D. Comte, et al., Signaling lymphocyte activation molecule family in systemic lupus erythematosus, *Clin. Immunol.* 204 (2019) 57–63.
- [31] J. De Salort, et al., Ly9 (CD229) cell-surface receptor is crucial for the development of spontaneous autoantibody production to nuclear antigens, *Front. Immunol.* 4 (2013) 225.
- [32] J.Y. Ju, Z.W. Xu, Potential genetic basis of B cell hyperactivation in murine lupus models, *Lupus* 30 (9) (2021) 1438–1448.
- [33] M.P. Karampetsou, et al., Expression patterns of signaling lymphocyte activation molecule family members in peripheral blood mononuclear cell subsets in patients with systemic lupus erythematosus, *PLoS One* 12 (10) (2017), e0186073.
- [34] M. Chatterjee, et al., Increased expression of SLAM receptors SLAMF3 and SLAMF6 in systemic lupus erythematosus T lymphocytes promotes Th17 differentiation, *J. Immunol.* 188 (3) (2012) 1206–1212.
- [35] M. Humbel, et al., SLAMF receptor expression identifies an immune signature that characterizes systemic lupus erythematosus, *Front. Immunol.* 13 (2022), 843059.
- [36] N. Meednu, et al., Dynamic spectrum of ectopic lymphoid B cell activation and hypermutation in the RA synovium characterized by NR4A nuclear receptor expression, *Cell Rep.* 39 (5) (2022), 110766.
- [37] T. Stuart, et al., Comprehensive integration of single-cell data, *Cell* 177 (7) (2019) 1888–1902.e21.
- [38] I. Korsunsky, et al., Fast, sensitive and accurate integration of single-cell data with Harmony, *Nat. Methods* 16 (12) (2019) 1289–1296.
- [39] D. Aran, et al., Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage, *Nat. Immunol.* 20 (2) (2019) 163–172.
- [40] Y. Zhou, et al., Metascape provides a biologist-oriented resource for the analysis of systems-level datasets, *Nat. Commun.* 10 (1) (2019) 1523.