Journal of Inflammation Research

Dovepress
Taylor & Francis Group

Open Access Full Text Article

ORIGINAL RESEARCH

Exploring Core Genes Associated with Sepsis and Systemic Inflammatory Response Syndrome Using Single-Cell Sequencing Technology

YuZhou Shen[1], LingHan Leng[2], YingChun Hu[1]

[1]Department of Emergency Medicine, the Affiliated Hospital of Southwest Medical University, Lu Zhou, Sichuan, People's Republic of China; [2]Department of Intensive Care Unit, Chengdu Fifth People's Hospital, Chengdu, Sichuan, People's Republic of China

Correspondence: YingChun Hu, Department of Emergency Medicine, The Affiliated Hospital of Southwest Medical University, 25 Taiping Street, Jiang Yang District, Lu Zhou, Sichuan, People's Republic of China, Tel +86-15228232720, Fax +86-830-3165120, Email huyingchun913@swmu.edu.cn

Purpose: As a crucial aspect of emergency critical medicine, sepsis has been in a difficult stage. As its "preparatory stage", SIRS has attracted the attention of the medical workers all over the world. The frequency of occurrence is on the rise, but there is a lack of certain indicators for the timely detection and recognition of illnesses.

Methods: By virtue of scRNA-seq, this research has analyzed single-cell transcriptome data from samples taken from groups with septic death and systemic inflammatory response syndrome so as to identify the unique markers and patterns in immune response.

Results: By revealing the status of twelve cell clusters of four major cell types in blood samples through UMAP cell clustering and the differences of major cell populations between the dead and SIRS patients, the results have elucidated the components of different cells and their marker genes in two disease states, and the response mechanism beneficial to disease diagnosis in blood samples.

Conclusion: By establishing a theoretical framework centered on cellular and molecular regulation, the study has introduced a novel approach for diagnosing and treating sepsis death group and SIRS patients early, as well as differentiating and preventing these conditions.

Keywords: sepsis, SIRS, single cell sequencing, biomarkers, single cell transcriptome

Introduction

The clinical features of sepsis patients are dis-responsive to external infectious factors and subsequent life-threatening complications.[1] Although there is a significant improvement in treatment for sepsis, including early identification of targets and disease-oriented approaches, its mortality rate is still high,[2] which is caused by the lack of acupuncture for early identification of sepsis and the very effective treatment.[3,4] Therefore, how to enhance the early diagnosis of sepsis is of great importance. To minimize the risk of adverse outcomes, namely progression or death, researchers have proposed the use of "biomarkers" that help diagnose sepsis.[5,6] According to the definition provided by the National Institutes of Health (NIH), biomarkers refer to considered objective "indicators" that can indicate the presence of a disease as well as determine the response to a particular therapeutic agent.[7]

Single-cell sequencing technology has undergone unprecedented changes in the past decade, whose technical data are gradually being used in various fields.[8] The simplified approach to SCS has become easier to use because of the rapid pace of globalization. Compared with the traditional "batch" DNA-seq and RNA-seq methods, the SCS method has the obvious advantage, that is, only it can provide the mixed signal sources of multicellular type and fusion changes of multiple confounding genotype tumor clones.[9]

In the past, the studies on expression profiling of host genes relied on whole blood to characterize diagnostic or prognostic gene features.[10–13] It's a method of collecting transcriptome signals from many different cell types but may not detect features of rare cells or recognize characteristics of cell type-specific disease.[14] In order to address these

Received: 6 November 2024
Accepted: 20 January 2025
Published: 6 February 2025

constraints, this paper has analyzed the immune cell status in the bloodstream of patients with sepsis by utilizing a gene expression profile at the single cell level.

By generating single-cell transcription profiles from septic deaths and systemic inflammatory response syndrome with the same scRNA-seq technique, this study has intended to elucidate the major cell populations by UMAP cell clustering. Founded on the identification of differences in marker genes, it then screened out new marker genes and analyzed the inflammatory cells in groups, so as to obtain information about the cell population and construct a single-cell map of blood. The research is anticipated to offer novel diagnostic concepts and treatment strategies for the early detection, recognition, prevention and management of illnesses.

Materials and Methods

Source of Data

The transcriptome sequencing data from individual cells were collected from blood samples of five patients who were hospitalized in the Emergency Intensive Care Unit at Southwest Medical University Affiliated Hospital in December 2019. The inclusion criteria are as follows. 1. It should meet the definition of sepsis 3.0 and the diagnostic criteria issued by the Society of Critical Care Medicine (SCCM) and the European Society of Intensive Care Medicine (ESICM), which emphasize that sepsis is a life-threatening organ dysfunction caused by a dysregulated host response to infection. Specifically, after infection, the presence of organ dysfunction, quantified by an increase of 2 points or more in the Sequential Organ Failure Assessment (SOFA) score, encompasses respiratory, cardiovascular, hepatic, renal, coagulation and central nervous systems. 2. In the study of sepsis, sample selection typically requires consideration of key factors such as age and gender to ensure the research is both comprehensive and representative: ① Age range. Given the characteristics of patients admitted to this institution, the age range is between 16 and 70 years old; ② Gender balance. To analyze the impact of gender differences on the incidence and prognosis of sepsis, sample selection should strive to maintain a balanced gender ratio. It helped to reveal the various manifestations and outcomes of sepsis in males and females because some studies have suggested that clinical features, treatment responses and prognoses of sepsis may vary by gender; ③ Control group. The healthy control groups should be closely matched with sepsis patient groups in terms of age, gender and baseline health to minimize selection bias. The study should exclude individuals with significant health issues to accurately reflect the impact of sepsis on health. Besides, the Charlson Comorbidity Index, a tool for assessing chronic disease burden and other comorbid conditions, was applied to control potential confounding factors. There was a sufficient sample size to calculate statistical power for detecting expected effect sizes, especially when exploring gender differences or other subgroup analyses. The exclusion criteria for non-eligible patients include non-bacterial infections; failure to meet sepsis definitions, typically based on the latest international consensus such as the "sepsis-3.0" definition; short hospital stays after sepsis diagnosis; specific disease states like malignancies, hematologic disorders or autoimmune diseases; incomplete data, like PCT levels or other critical laboratory results; poor adherence to designated treatment protocols; and other clinical conditions such as pregnant women, patients with end-stage multi-organ failure, and those with poor treatment adherence due to their special circumstances or research risks. All subjects or their legal representatives were willing to participate in the study after signing the informed consent.

Preparation and Sequencing of Single Cell Library

By encapsulating beads with Cell Barcodes and cells in droplets, capturing droplets with cells and lysing cells within the droplets with microfluidic technology, the 10x genomics platform has enabled the connection of mRNAs in cells with Cell Barcodes on the beads to create Single Cell GEMs. The droplets have been employed to carry out the reverse transcription reaction for creating a library of cDNAs, with the sample index on the sequence library used to identify the source of the target sequence.

10x Sequencing Data Preprocessing

As per information from 10x Genomics, visit https //support. 10xgenomics. com/single-cell-gene-expression/software/pipelines/latest/what-is-cell-ranger. The raw sequencing data has been analyzed using CellRanger v2.2.0 software.

Simply put, the "cellranger mkfastq" function has first been used to convert the sequencing raw BCL file to an FASTQ file, which has been processed by virtue of the 'cell ranger count' wrapper function. In addition, the sample has been resized with the '-force-cells¼ 7000' argument. The cell cursor counting function has employed the wrapped STAR software to compare sequences to the reference genome. The output file has contained the gene expression matrix and barcode information for the CellRanger pipeline, which is then utilized for downstream visual analysis.

Characteristics of Cell Clusters

The 10x genomics official software Cell Ranger has been used for quality control of samples, with the integration of the STAR[15] software. Based on the comparison of the reads to the reference genome, the quality control results have been obtained from the original data, encompassing the high-quality cell number, gene number and genome comparison rate, in order to assess the quality of each sample. As for cell clustering, refer to the https://satijalab.org/seurat/. Moreover, the CellRanger-generated "filtered gene bc matrices" has been employed as the input file for Seurat. The quality control criteria are as follows: elimination of cells with fewer than 200 genes, less than 1000 UMIs, log10GenesPerUMI below 0.7, mitochondrial UMI proportion exceeding 10%, and red blood cell gene proportion surpassing 5% to identify high-quality cells. Subsequently, the DoubletFinder software has been utilized for double cell elimination and subsequent analysis. Following normalization, the genes specific to each dataset have been computed for the subsequent clustering analysis.

Initially, the MNN algorithm has been utilized to remove the batch effect by identifying mutual nearest neighbors. Next, the UMAP algorithm has been applied to visualize the single-cell cluster founded on the dimension reduction outcomes from MNN. The clustering algorithm has adopted SNN (spatial clustering algorithm based on sharing the nearest neighbor density) to finally get the optimal cell cluster. The number of cell clusters can be adjusted by setting the resolution parameters. In the later stage, it is decided whether to adjust the clusters according to the cell type identification. Generally, the default is resolution = 0.4.

In order to evaluate the quality of duplicate samples or their individual differences, the different cells between a variety of groups are assessed for further analysis.

To identify the specifically expressed genes in the cluster, this study has adopted the presto test method to test the differences between the designated cell populations and all other cell populations. To identify all maker genes in each cell population, the screening criteria are set as follows: Logfc. threshold had to be above 0.0 and min.pct (the percentage of genes expressed in all cells) had to exceed 0.25.

In order to comprehensively examine the cell clustering and the development of downstream differentiation paths, this paper has utilized the SingleR[16] tool to label the cells with a shared dataset. The FindMarkers function of the Seurat package has been employed to screen all cells in each sample group for differentially expressed genes between groups. Moreover, the default presto difference test from an opensource software library has been adopted to enhance the data analysis speed in bioinformatics.

Enrichment Analysis of Constructed Marker Genes

The hypergeometric distribution test has been utilized to determine whether the function set was significantly enriched in the list of differential protein coding genes, which have been selected from a background list of all protein coding genes. Then, the resulting p value has been adjusted with the Benjamini & Hochberg multiple tests to obtain the value.

KEGG[17] serves as the primary public repository for Pathway information. Pathway analysis (combined with KEGG annotation results) of differential protein encoding genes has been performed by KEGG database, and the significance of differential gene enrichment in each Pathway entry has been calculated by hypergeometric distribution test.

The STRING https://string-db.org/[18] refers to a database of functional correlations between protein predictions. Genomic correlation between protein encoding genes can frequently indicate functional linkages. Genes required for specific functions typically exhibit comparable species coverage, which are closely situated in the genome (especially in prokaryotes) and are prone to cause gene fusion. As a precomputed global resource, the STRING database has been adopted to browse and analyze these dependencies. The analysis has been enhanced by utilizing the STRING database to identify the interaction patterns among different genes and create a network diagram.

Visualization analysis of single-cell localization the cell-line localization of single target gene has been detected by 10x scRNA-seq technique, whose particular procedure has been carried out in accordance with the company's operational guidelines. The collected blood samples have been subjected to high-throughput sequencing after mixing to generate raw data in fastq format. The quality of the raw data has been assessed using the official CellRanger software from 10x genome. The Seurat software package has been utilized for additional data quality control after visiting 10xgenomics.com/single-cell-gene-expression/software/pipelines/latest/what-is-cell-ranger. Gene expression data has been used to conduct linear dimensionality reduction analysis, which has then been visualized in two dimensions using TSNE. FindAllMarkers function has been employed to identify marker genes, while VlnPlot and FeaturePlot function to identify specific genes. SingleR software can determine the correlation between cell expression profiles and reference dataset to create a sepsis-related single-cell library by assigning the most correlated cell type. The main gene identified in the previous research has been encompassed in a single-cell library to investigate the positioning of the specific genes within the cell line.

External Validation

To validate the study, five datasets—GSE6525, GSE12624, GSE28750, GSE63042, and GSE74224—have been downloaded from the GEO database. These datasets consist of the samples from 20, 36, 12, 28, and 22 sepsis patients, as well as 10, 127, 12, 83, and 102 healthy individuals, respectively. After assessing the differences between the two groups, the box plots of hub genes have been generated based on R to verify the differential expression of core genes between sepsis patients and healthy individuals.

Statistical Analysis

The data has been analyzed and visualized according to the r4.1.0 and SPSS 26.0 statistical software. A *t*-test has been employed to calculate the differences in critical gene expression between sepsis and healthy samples. A P-value < 0.05 has been considered statistically significant (*P < 0.05, **P < 0.01, ***P < 0.001).

Results

Single-cell data sequencing and overall technical route in Figure 1.

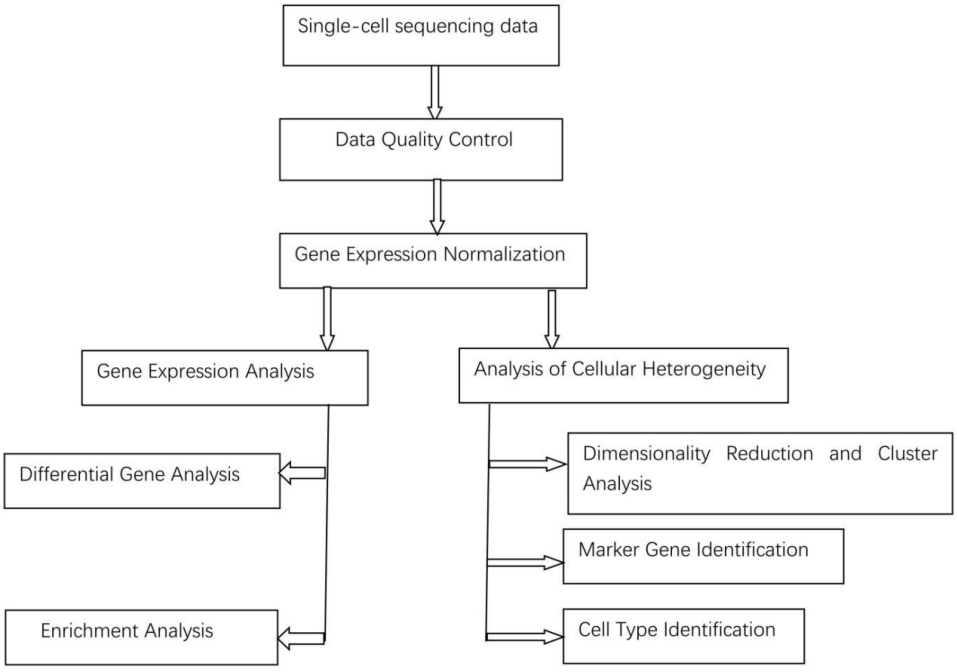

Figure 1 Single cell sequencing analysis process.

Exploration of individual cell transcriptomes and analysis of cellular clustering.

A pie chart has displayed the nGene, nUMI, log10GenesPerUMI, percent_mito, and percent_HB of each cell before and after quality control according to the quality control criteria in Figure 2A. In this study, the blood cells have been divided into 12 main types according to the results of dimension reduction clustering in Figure 2B. The cluster structure between the systemic inflammatory response syndrome group and the sepsis death group has been similar and the cells have been relatively concentrated in Figure 2C. According to the sample proportion the histogram in each cell population has indicated that SIRS mainly concentrated in cluster1, cluster3, cluster7, cluster9, and the sepsis death group mainly concentrated in cluster5, cluster6, cluster10, cluster11, cluster12 in Figure 2D.

Display of cell dimension reduction clustering grouping among samples.

Before clustering quantitative results of single-cell transcriptome, it is generally necessary to extract new dimensions from tens of thousands of genes through dimension reduction and use the data represented by the new dimensions. On the one hand, the data information in the sample can be retained to the maximum extent. On the other hand, there is the redundancy. The data can be effectively reduced, so as to improve the efficiency of subsequent clustering operation.

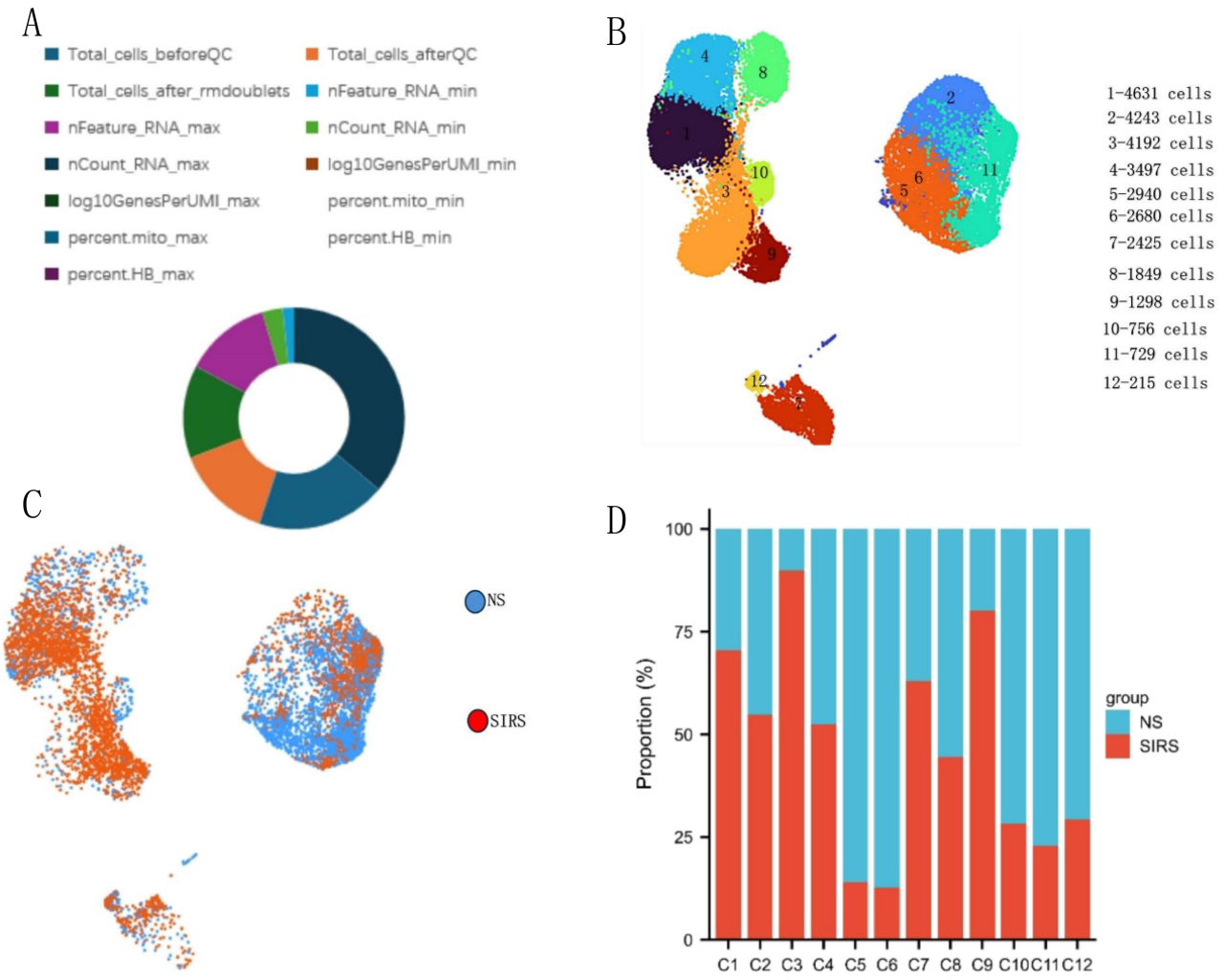

Figure 2 (**A**) Samples are analyzed after QC, and different colors and areas represent different QC contents before and after QC. (**B**) Blood samples are grouped by major cell type. A variety of hues indicate the distinct groups, while the diverse areas signify various cell types and quantities. (**C**) The samples are classified into different color clusters. (**D**) Information on cell number distribution in different cell populations.
Abbreviations: NS-, Non-survival group in sepsis; SIRS, Systemic Inflammatory Response Syndrome.

https://doi.org/10.2147/JIR.S448900

In general, the similar cells have similar gene expression profiles, so they can be clustered together according to the gene expression results of each cell to form a cell population. Based on this, in the present study, the cells have been divided into twelve cell populations by dimensionality reduction clustering, among which the cell types with reference to B_cell, Monocyte, T_cells, NK_cell, etc., as shown in Figure 3A and B. The histogram has been utilized to reveal the proportion of different cell communities in the two groups respectively. Specifically, the highest proportion in the NS group is cluster6; in the SIRS group, the highest proportion is cluster1 and cluster3, as shown in Figure 3C.

Sequencing Analysis of Raw Data

The comparison of gene distribution between the two groups has been examined using a volcano plot, with criteria set at $P < 0.05$ and FC = 1.5 in Figure 4A. The Deps has then been submitted to the STRING https://stringdb.org online platform to build a $P < 0.05$ in Figure 4B.

The data can be effectively reduced, so as to improve the efficiency of subsequent clustering operation.

Identification of marker gene. The designated cell population has been compared to all other cell populations according to the presto test method under the screening conditions in which the log2FC. Threshold was above 0.0 and the min.pct was above 0.25. This has resulted in the identification of all marker genes in each cell population in Figure 5.

Differential Gene Enrichment Analysis

The analysis of functional enrichment in GO for marker genes has indicated that these genes were primarily associated with immune inflammatory response, cell metabolism, cell secretion and cell activation. Predominantly situated on the extracellular region's biofilm, they can facilitate cell cross-linking and information exchange in Figure 6A. KEGG functional pathway analysis, mainly related to infectious disease signal transduction, has been mainly reflected in inflammatory immune pathway. However, few have involved in tumor immunity and cell growth and death pathway in Figure 6B.

Visualization Analysis of Single Cell Localization

In Figure 7A, the x-axis represents the gene name and the y-axis represents the number of cells in the population after analyzing single cell located on https//cloud.oebiotech.com/task/. A larger dot indicates a higher percentage of cells expressing the gene. The color gradient from blue to red signifies the increased gene expression. The genes of LILRB2, PSAP, S100A11, LST1, S100A8, SERPINA1 and IFI30 in the top 20 are predominantly expressed in cell lines three and five, with significantly higher expression compared to other cell lines. In Figure 7B, a cell is represented by a dot in the key, with the color changing from blue to red to show the level of gene expression in the cell. In the graph, the x-axis represents the number of cells, while the y-axis shows the normalized gene expression value. Notably, the levels of LILRB2, PSAP, S100A11, LST1, S100A8, SERPINA1 and IFI30 are much greater compared to other cell lines in Figure 7C.

Survival Curve Analysis

The diagnostic value of marker genes in sepsis death group and SIRS has been explored. LILRB2, PSAP, S100A11, LST1, SERPINA1 and IFI30 had higher survival rate in the over-expression samples, and S100A8 had a higher survival rate in low-expression samples in Figure 8A–G.

ROC Curve Analysis

The x-axis and y-axis of the ROC curve represent the sensitivity and specificity, respectively, with a larger AUC indicating a more accurate diagnostic model. As illustrated in Figure 9A–G, based on the GEO database website, the clinical diagnostic curve analysis has been conducted based on datasets GSE28750, GSE63042 and GSE74224. The results, as shown in the figure above, have demonstrated that the seven genes, likeLILRB2, PSAP, S100A11, exhibit favorable diagnostic sensitivity and specificity, with all AUC values exceeding 0.600. These genes can be utilized to assess the quality of predictive models.

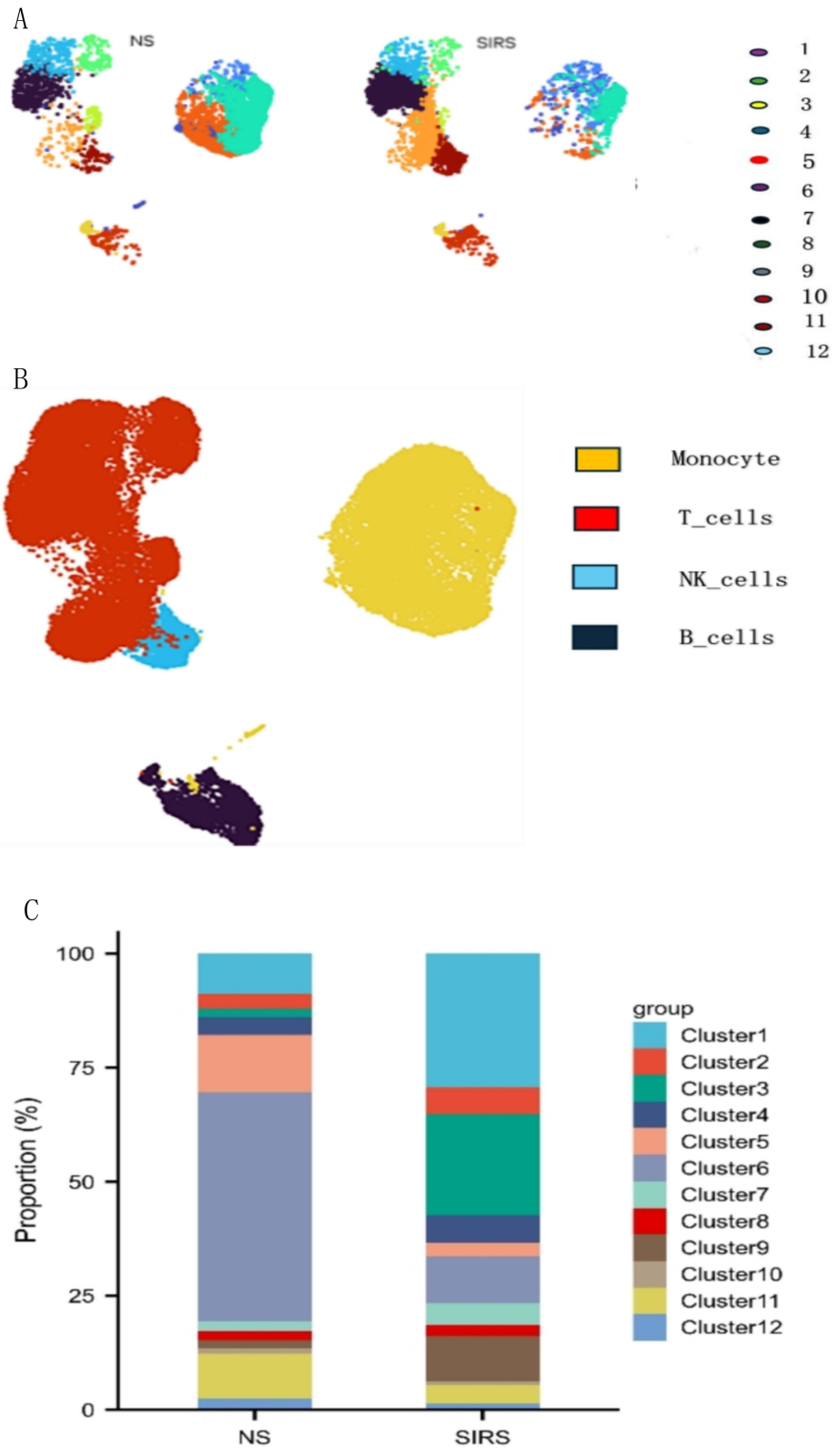

Figure 3 Display of dimension reduction clustering results. Classification dimension reduction results between different samples (**A** and **B**). Histogram statistical analysis (**C**) of dimension reduction results in two samples.

https://doi.org/10.2147/JIR.S448900

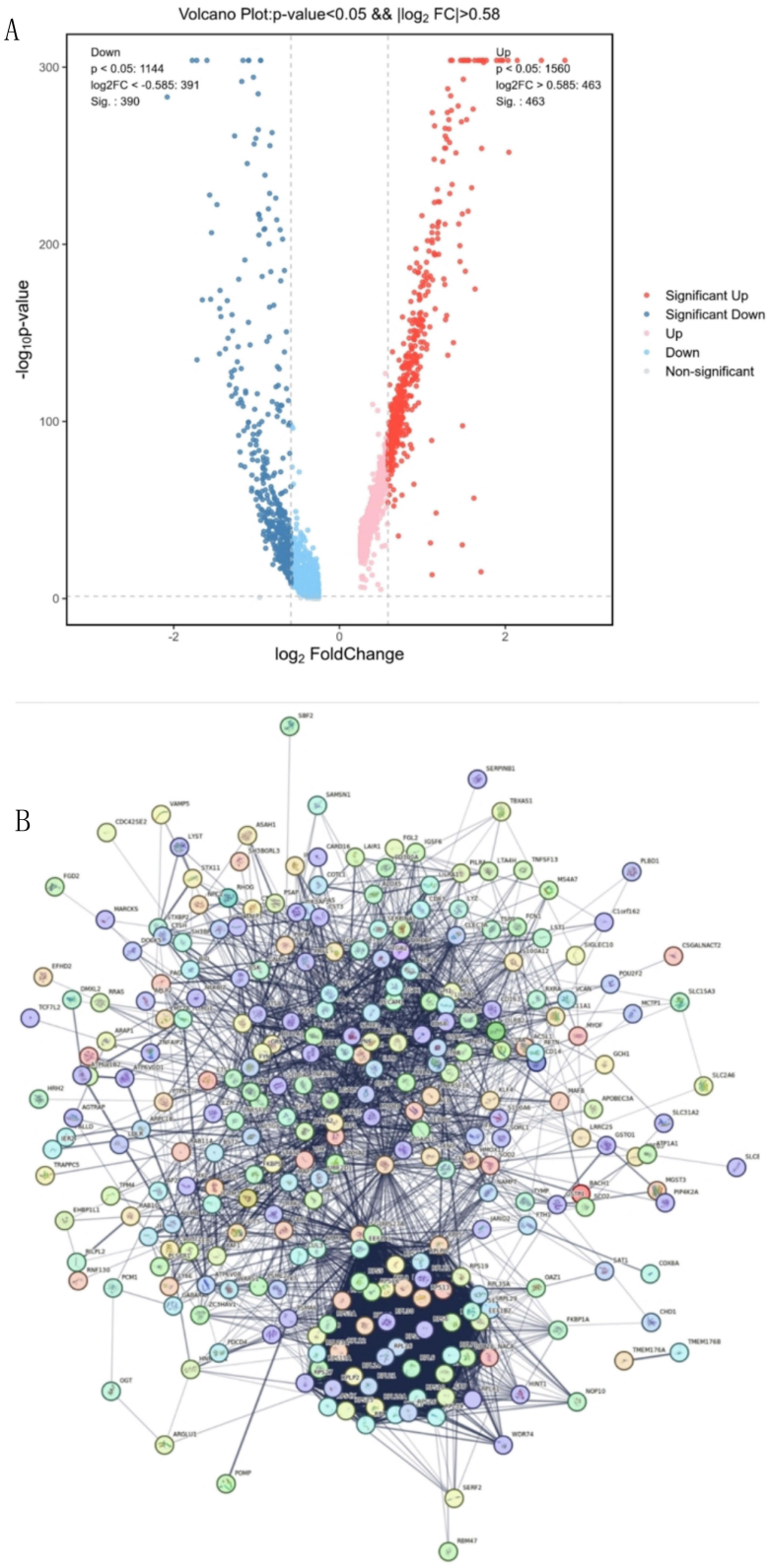

Figure 4 Analysis of the raw data. Volcanic plots have been used to show the change in the number of differential genes between the septic death group and the SIRS group (**A**). The PPI network diagram represents the internal association between common differential genes (**B**).

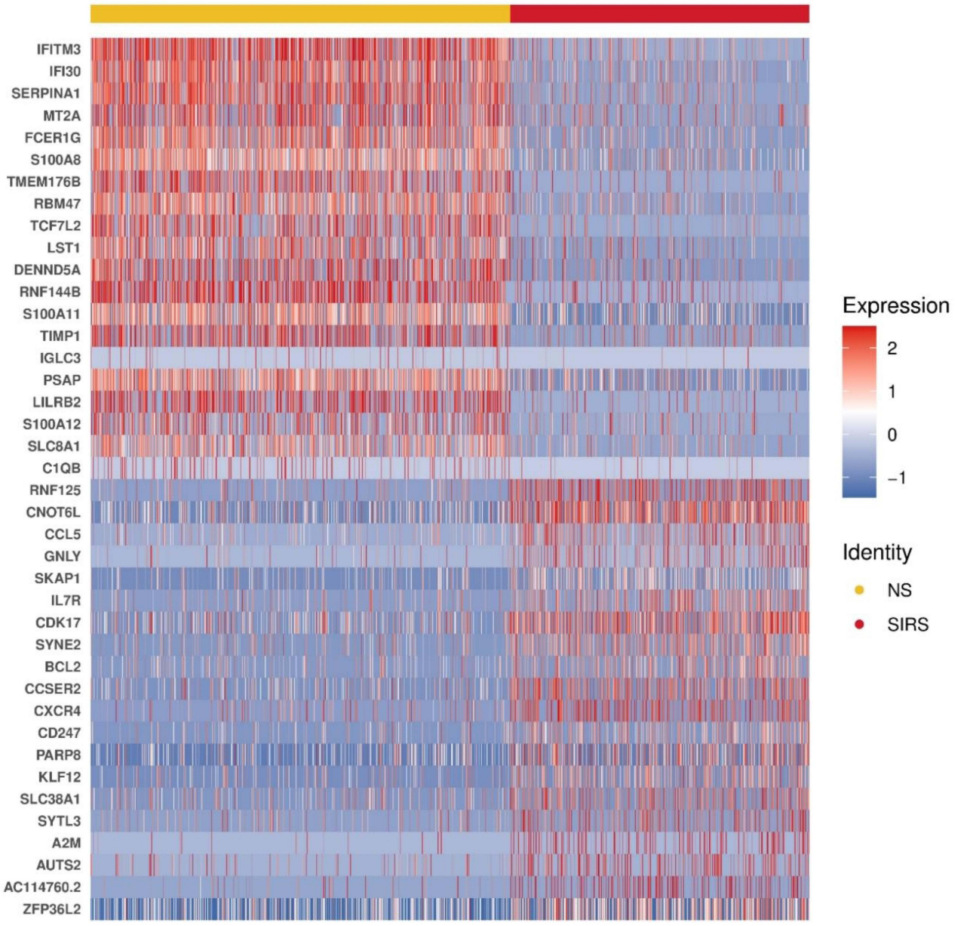

Figure 5 TOP20 gene identification. The abscissa is different grouping information and the ordinate is maker gene. High expression is exhibited by red in the Figure 5, while low expression by blue.
Abbreviations: NS, Non-survival group in sepsis; SIRS, Systemic Inflammatory Response Syndrome.

External Validation of Core Genes

The previous biological analysis results have been validated according to the sepsis datasets of GSE6525, GSE12624, GSE28750, GSE63042 and GSE74224 from the GEO database. It has found that, in the GSE6525 dataset, the expression levels of LILRB2, S100A8 and S100A11 were significantly higher in sepsis patients compared to those in normal subjects ($p < 0.0001$), along with PSAP, S100A11, LST1, SERPINA1, IFI30 and S100A8IFI30. Similarly, in the GSE12624 dataset, the expression levels of S100A8 were significantly elevated in sepsis patients compared to those in normal subjects ($p < 0.0001$), which has confirmed these findings in Figure 10A–G.

To comprehend the immune condition of SIRS or non-surviving sepsis based on clinical status, this study has conducted a retrospective analysis of single-cell PBMC sequencing in two hospitalized patients with gram-negative sepsis, one patient diagnosed with SIR and two healthy volunteers. Two individuals who had been hospitalized as patients successfully recovered and were released from the hospital. But the one patient did not survive. Sepsis survivors and non-survivors were preliminarily distinguished according to clinical parameters such as SOFA and APACHE scores, as well as the SIRS group in Table 1.

Discussion

In emergency intensive care units, sepsis has become a serious threat to human life and health. The number of patient deaths per year cannot be underestimated.[19] Sepsis has brought a heavier medical and financial burden than other patients such as SIRS.[20] Therefore, in the early identification of the disease progression process, blocking the progression of

https://doi.org/10.2147/JIR.S448900

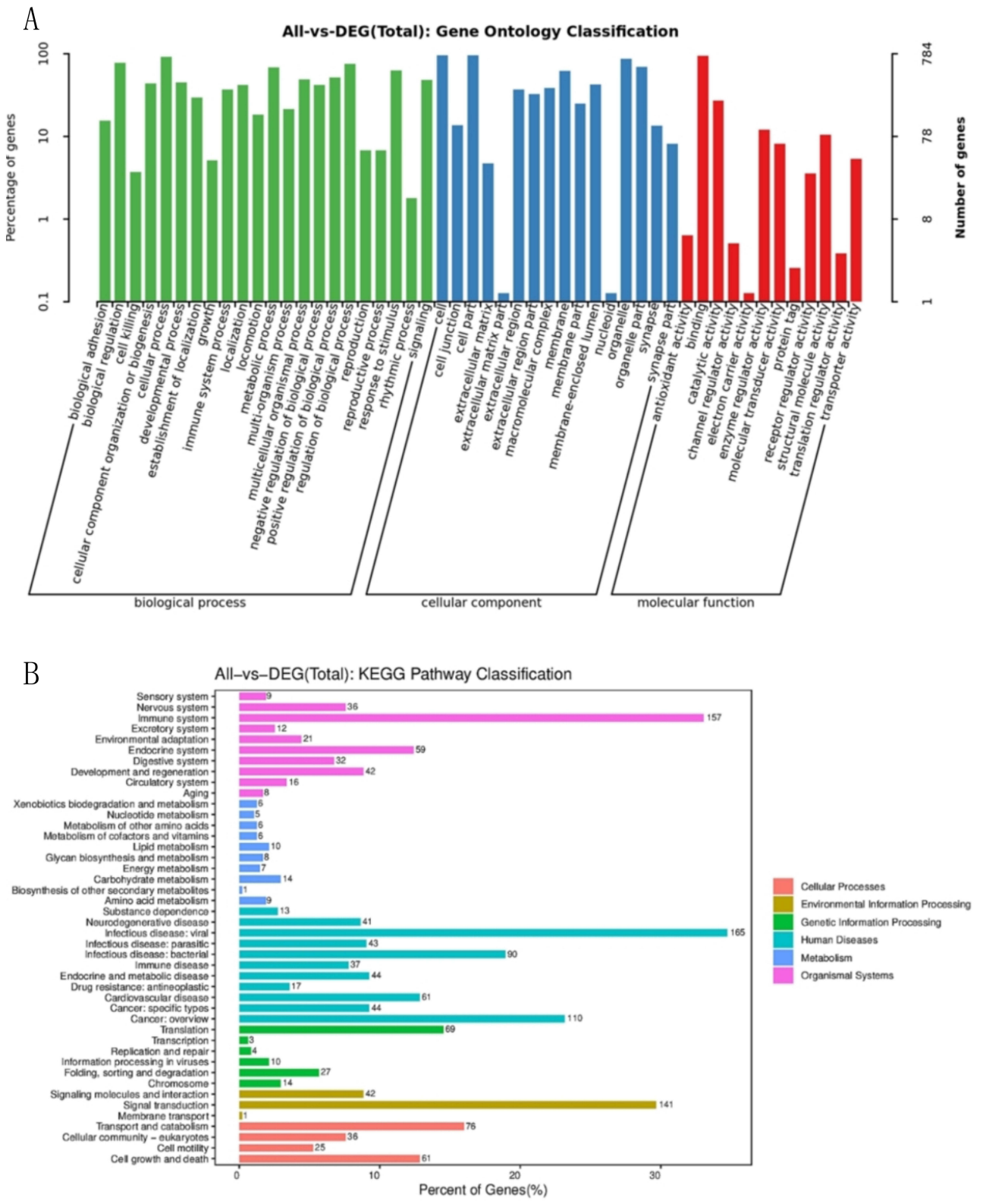

Figure 6 Analysis of enriched genes showing differential expression. The top 30 bar charts from Go enrichment analysis display the three biological processes (**A**). KEGG Analysis of Possible Pathway Loci of Differential Genes (**B**).

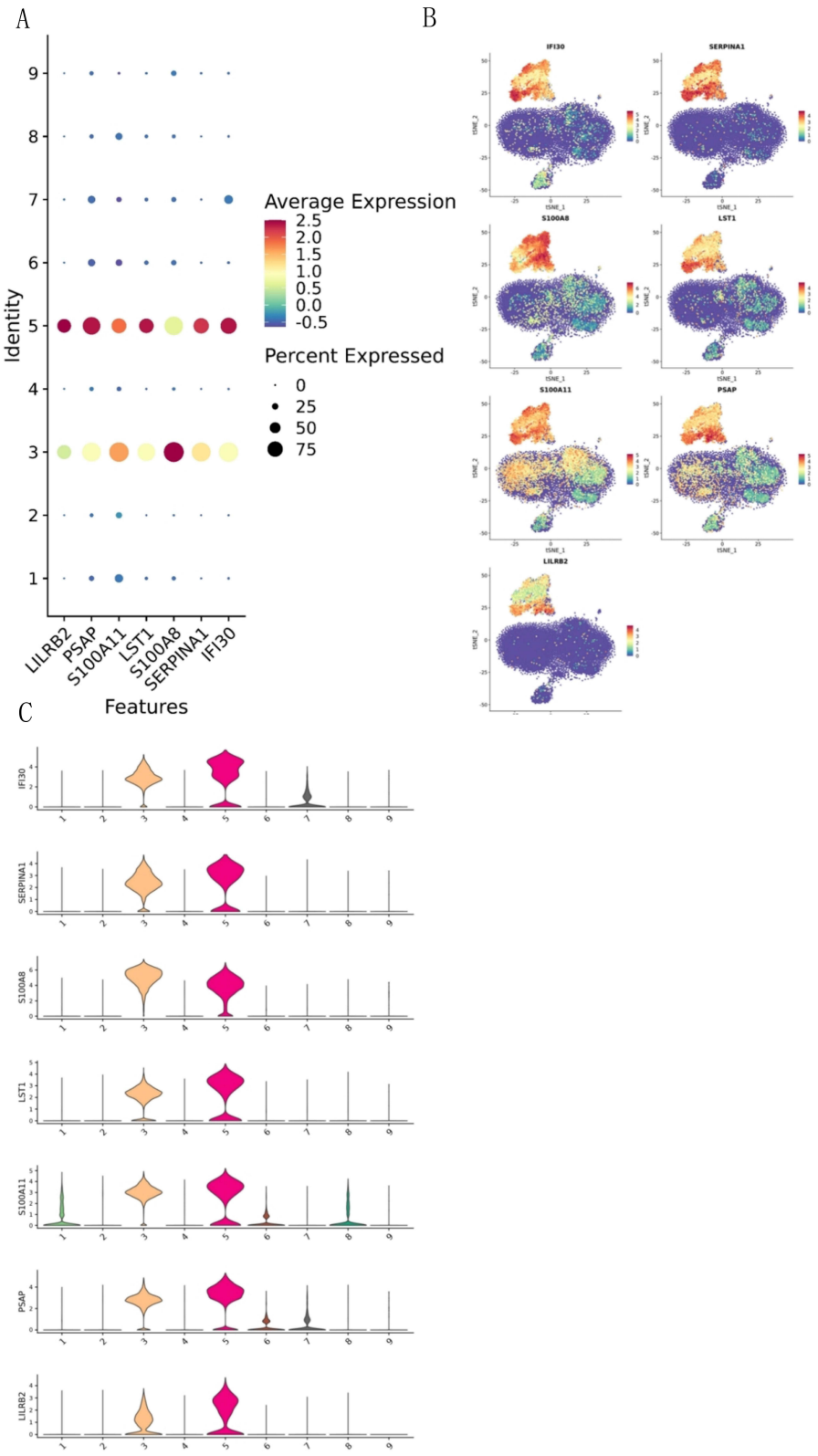

Figure 7 Single cell gene mapping. (**A–C**) The distribution and expression of top20 core genes have been analyzed, with LILRB2, PSAP, S100A11, LST1, S100A8, SERPINA1, IFI30. 7 genes most prominent.

https://doi.org/10.2147/JIR.S448900

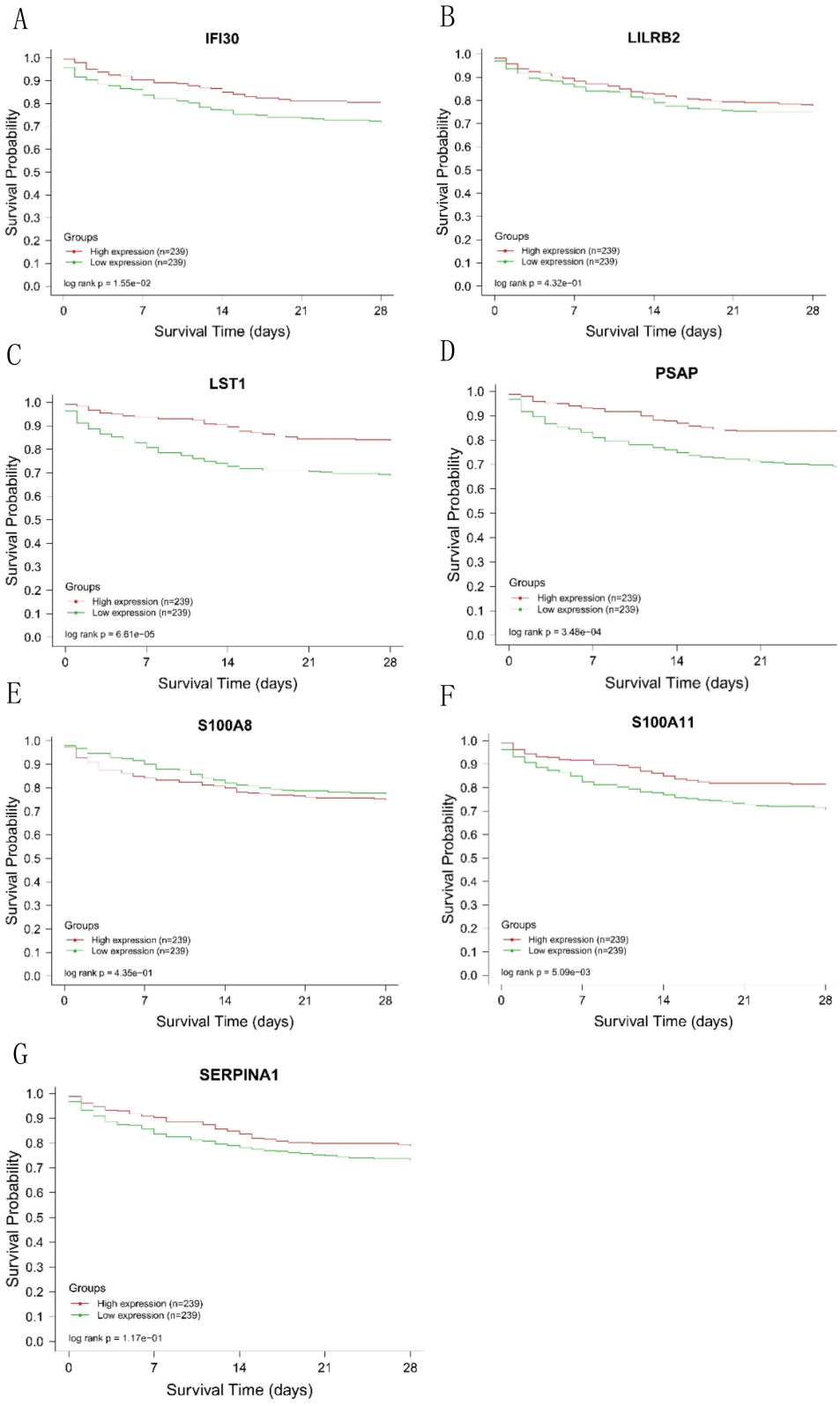

Figure 8 Survival curve analysis. (**A–G**) In the high expression group, the six essential genes, namely LILRB2, PSAP, S100A11, LST1, SERPINA1 and IFI30, have exhibited the increased survival rates. In contrast, S100A8 had a lower survival rate, with statistically significant P<0.05 in the over-expression group.

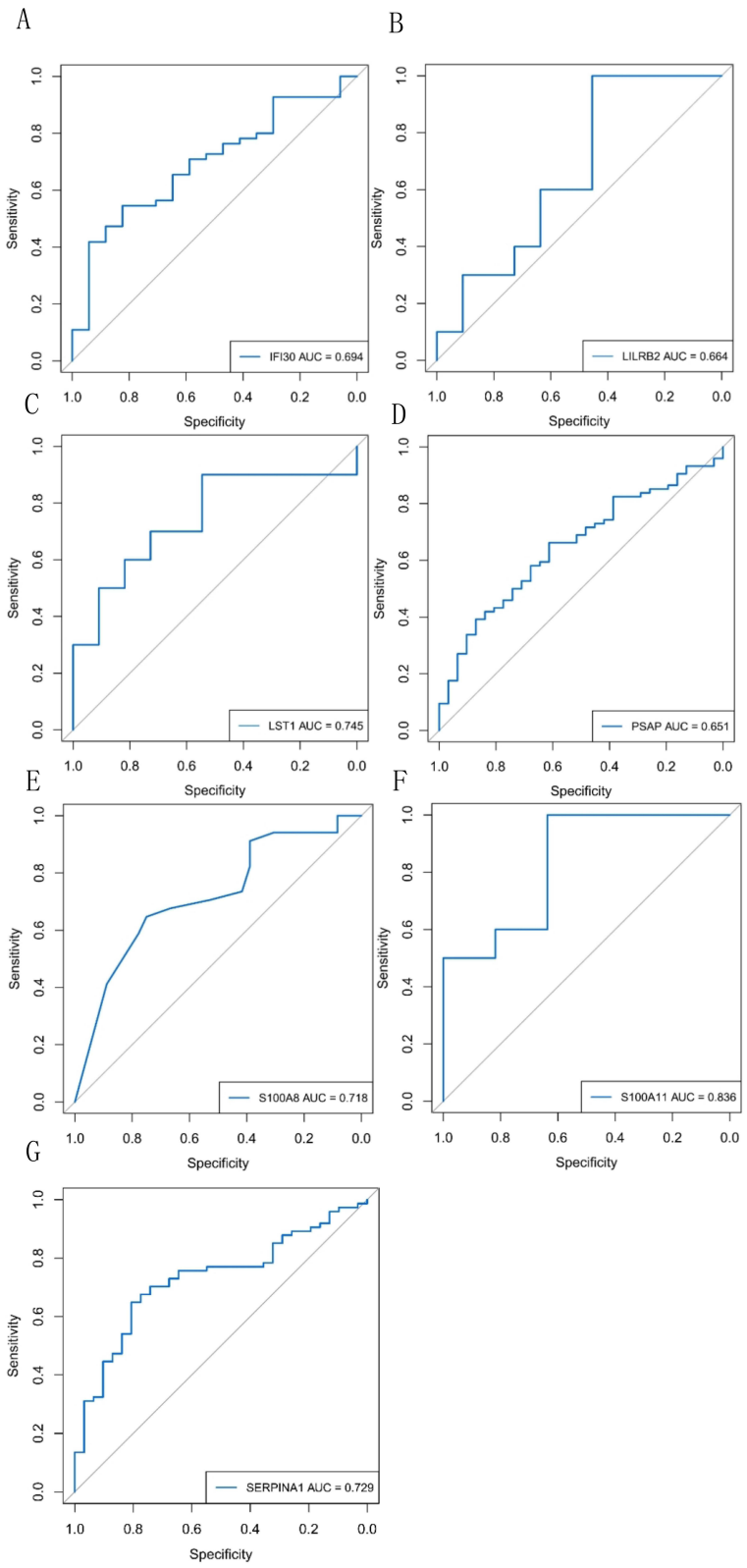

Figure 9 ROC curve analysis. (**A–G**) Under the curve, the seven essential genes, namely LILRB2, PSAP, S100A11, LST1, SERPINA1, IFI30 and S100A8, has separately demonstrated their diagnostic specificity and sensitivity for sepsis and systemic inflammatory response syndrome in large-scale clinical samples. The areas under the curve (AUC) reflecting the discrimination ability of each biomarker are provided in Figure 9A–G.

https://doi.org/10.2147/JIR.S448900

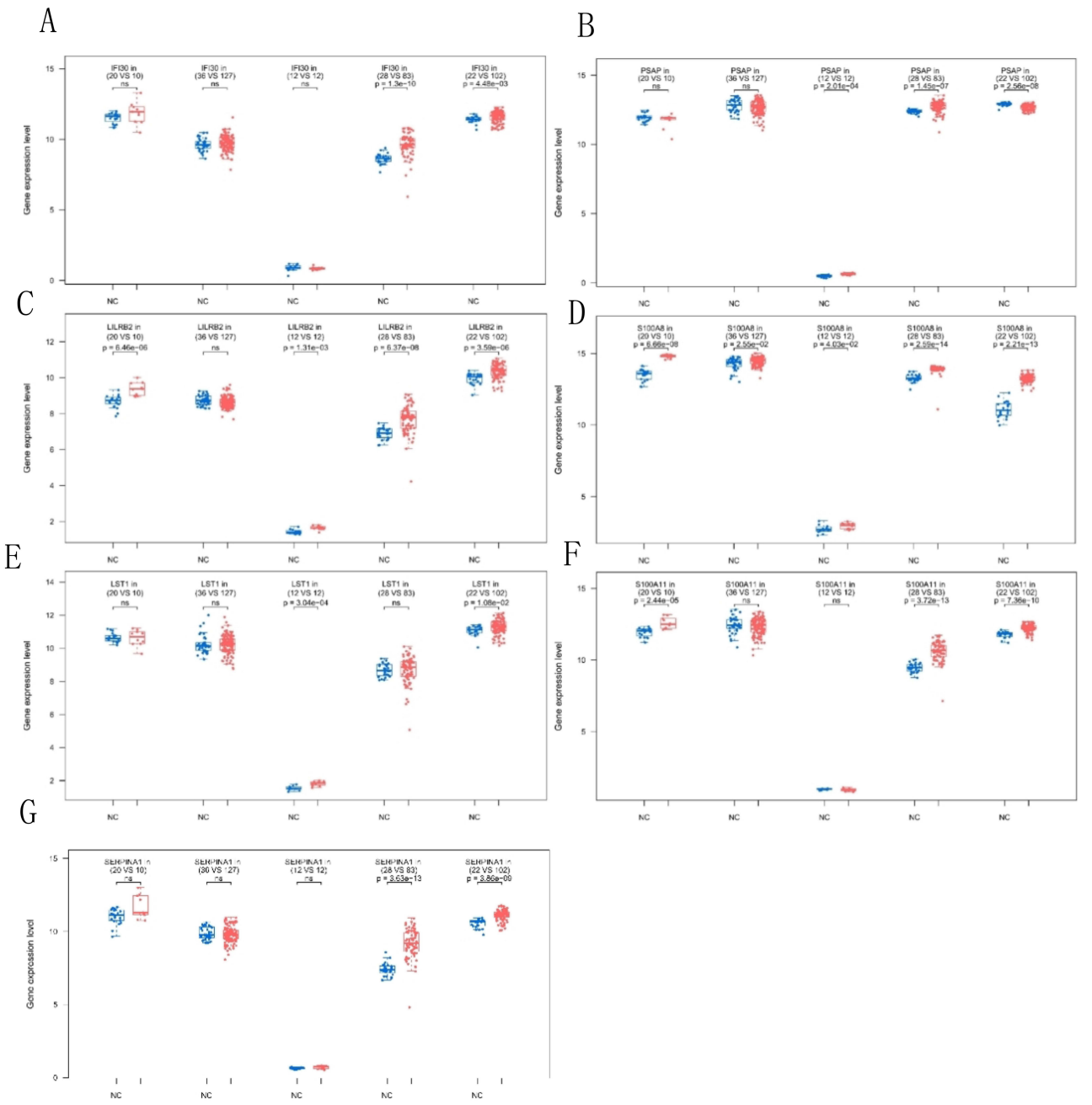

Figure 10 External validation of hub genes. (**A–G**) Hub gene expression between sepsis group and control group, P < 0.001. (**A**) IFI30, (**B**) PSAP, (**C**) LILRB2, (**D**) S100A8, (**E**) LST1, (**F**) S100A11, (**G**) SERPINA1.

disease has become a great tool to reduce the risk of sepsis death. Notwithstanding, the task of finding new diagnostic and prognostic markers has still remained daunting.

Using scRNA-seq, this study has generated the four primary single-cell transcriptome profiles from blood samples of sepsis patients in the NS and SIRS categories. UMAP cell clustering has been adopted to reveal the distinctions between the two primary cell groups. The groups of samples have been analyzed to gather information on cell grouping and create a blood cell map consisting of individual cells. Next, the data on macrophages has been collected and analyzed to investigate how their involvement may impact the onset and progression of diseases. Further, the comparison between death group and SIRS group has helped to clarify their differences and the impacts of gene expression in "inflammatory storm" cells, so as to obtain the highly expressed differential genes in sepsis death group. Furthermore, the enrichment of

Table 1 Displaying the Attributes of Non-Septic Participants and Septic Individuals Upon Sepsis Identification.

	Healthy Control (n=2)		**Sirs (n=1)**	**Sepsis Survivor or Nonsurvivor (n=2)**	
Gender	**Male**	**Female**	**Female**	**Female**	**Female**
Age (years)	46	53	38	54	63
Sepsis etiology	n/a	n/a	n/a	E.coli bacteremia	K.pneumoniae
APACHEII	n/a	n/a	15	28	36
SOFA	n/a	n/a	4	18	25
Time of death (days post enrollment)	n/a	n/a	n/a	>28	4

Abbreviations: Sirs, Systemic Inflammatory Response Syndrome; Non Survivor, None Survival; E.coli, Escherichia Coli; APACHELI, Activated Partial Thromboplastin Time (s); SOFA, Sequential Organ Failure Assessment; N/A, No Observable Differences.

function and KEGG signal pathway have been employed to elucidate the effects of inflammatory changes. It has provided the basis for the follow-up intervention of genes. There are methods for determining cell-type identity in single-cell sequencing clustering: ① Based on common features, such as clustering after preprocessing and identification according to cell-specific gene expression patterns, with the option of cross-reference with feature databases; ② Utilizing marker genes to identify clusters with high expression of specific genes for determination. For instance, CellBIC (Single-Cell Bimodal Clustering) is a newly developed clustering tool that clusters founded on bimodal characteristics of gene expression distributions, which can better preserve the hierarchical structure of cells to enhance the accuracy of cell-type identification.[21] Additionally, CIARA (Cluster-Independent Identification of Rare Alleles), with concentration on identifying marker genes for rare cell types, is capable of directly discerning rare cell types from single-cell data without relying on clustering.[22]

The potential for misclassification can lead to numerous consequences: ① In terms of biological understanding, it can disrupt the research on cell development because of the misinterpretation of cellular functions; ② In disease studies, it may result in misdiagnosis of disease types or even hinder drug development; ③ In data interpretation and subsequent research, it can lead to uncertain conclusions. For instance, in the study of breast cancer, misclassification of stem cell states could cause the erroneous assumptions about tumor origins, which may clinically impact the formulation of treatment strategies.[23]

Finally, with the continuous advancements in single-cell sequencing technology, the new analytical methods and tools have been continually emerging to offer fresh opportunities of enhancing the accuracy of cell-type identification. For instance, the combination of single-cell transcriptomic data with other types of data (such as spatial transcriptomics) can contribute to a more comprehensive understanding of cell functions and interactions, thereby reducing the likelihood of misclassification.[24]

The twelve cell clusters and four major cell types have been identified from the blood cell data center by UMAP cluster analysis. Immune cells are primarily categorized as T cells, B cells, plasma cells, macrophages, monocytes, dendritic cells and mast cells. After dimension reduction and clustering in this research, a noticeable discrepancy in cell community distribution has been observed between the sepsis death group and the SIRS group. Specifically, cluster 6 has been predominantly found in the former, while cluster 1 and cluster 3 in the latter. Therefore, the blood sample tissues between different groups have been explored through dimension reduction and clustering method to "visualize" the blood sample tissues between various groups. Not only repeated samples of the same group, but also the quality or individual differences of repeated samples in the group can be evaluated. In addition, samples of different groups and the cells between groups can be assessed, so as to lay a foundation for further locking the genes of difference.

The marker genes play a role in distinguishing the varying levels of expression changes within the same gene across different cell populations, by utilizing the presto test method to compare the specified cell population with all other cell populations. The screening criteria are as follows: the Log2FC. threshold must exceed 0.25, which ensures each cell population contains all marker genes. In the end, through enrichment analysis method, the functions of marker genes in

the cells and their potential mechanism pathway have been identified, which can provide necessary help for the subsequent identification of specific pathogenesis.

In clinical practice, the discovery of differential genes is primarily translated into clinical intervention strategies or the identification of early sepsis.

Establishment of the Association Between Differential Gene Expression and Sepsis

Differential gene expression refers to the changes in gene expression levels under different conditions, such as between healthy and diseased states. In sepsis, certain genes exhibit significantly different expression compared to those in normal physiological conditions. For instance, genes related to immune response, inflammation regulation and cell apoptosis may show upregulated or downregulated expression in septic patients;[25]

Methodology for Translating Discovery Into Early Detection Tools

This has involved the identification of specific gene combinations through large-scale gene expression analysis, so as to pinpoint gene sets with distinctive expression changes during the early stages of sepsis. These gene combinations serve as potential biomarkers. For instance, by analyzing gene expression profiles of blood samples from sepsis patients and healthy controls, a set of genes highly or lowly expressed in the early stages of sepsis can be screened out, which can be employed for the construction of detection model. If subsequent testing reveals that the expression of these genes in an individual's blood aligns with the characteristic pattern of early-stage sepsis, it can signal the potential for early occurrence of sepsis;[25] ② The dynamic monitoring of gene expression employs techniques such as real-time quantitative PCR to continuously track the expression dynamics of specific genes in suspected patients. Since sepsis is a constantly developing disease process, gene expression also changes as the conditions progress. Monitoring the dynamic expression of genes closely associated with the early pathogenesis of sepsis (such as those related to inflammatory factors) can help detect abnormal gene expression changes before symptoms manifest, so as to achieve early detection.

Pathways to Clinical Intervention

Drug Development Targets

① Functional Changes in Differential Genes

If certain genes, which are associated with adverse disease progression (eg, genes promoting excessive inflammatory responses), have been found to be highly expressed in sepsis, the proteins encoded by these genes can be targeted for drug development. Developing the drugs that inhibit the functions of these proteins improves the condition of septic patients by alleviating the inflammatory response. For instance, for inflammatory-related gene product over-expressed in septic patients, small molecule compounds that can specifically bind to prevent the activity of this product can be identified through computer-aided drug design and high-throughput screening, which can serve as potential therapeutic agents.[25]

② Personalized Treatment Strategies: Guided by Gene Expression Profiling

Sepsis patients are categorized based on the variations in individual genetic expression. Patients exhibiting distinct gene expression profiles may respond differently to various treatment modalities. For instance, certain gene expressions indicate that individuals with heightened inflammatory responses may benefit from more aggressive anti-inflammatory regimens. Conversely, those with severely impaired immunity revealed by their gene expression profiles might require not only anti-infective treatments but also immunomodulatory therapies. This personalized treatment based on gene expression can help reduce adverse reactions by enhancing the efficacy of the treatment.

LILRB2, also known as immunoglobulin-like transcript ILT4, monocyte/macrophage immunoglobulin-like receptor MIR-10, or CD85d, is viewed as a promising target for immunotherapy targeting immunosuppressive molecules and cancer according to references,[26–28] which are mainly expressed[29–31] in myeloid cells. The precise mechanism of sepsis has still remained unclear in research. The distinct focus of this research, differing from this project, has particularly

emphasized tumor immune evasion, neurodegenerative diseases and immunomodulation in recent years. Furthermore, the studies on LILRB2 have increasingly focused on its role in the tumor microenvironment and its interactions with other immune receptors:

① The role of LILRB2 in tumors has garnered significant attention. Research has revealed that LILRB2 expression is associated with the prognosis of various cancers, particularly in breast cancer, where high expression is closely linked to immune evasion mechanisms. LILRB2 can inhibit activity of CD8+ T cells and help tumor cells in evading immune surveillance by enhancing the degradation of HLA-A. This mechanism provides a new target for cancer immunotherapy. Blocking LILRB2 functions may enhance anti-tumor immune responses.[32]

LILRB2's role in the nervous system has also attracted much attention. Studies have found that LILRB2 expression is upregulated in the temporal lobes of people with epilepsy, with its expression levels inversely correlated with seizure frequency. This has revealed that LILRB2 may play a role in the pathogenesis of epilepsy. Further research could uncover its potential functions in neuroplasticity and neurodegenerative diseases.[33]

The relationship between LILRB2 and HIV infection has also been extensively investigated. Studies have displayed that the interaction between LILRB2 and HLA class I molecules may influence the disease progression by affecting HIV-1 viral load. It has underscored the importance of LILRB2 in regulating immune responses and viral control, thus potentially offering new avenues for HIV treatment.[34] In summary, the research focus of LILRB2 lies in its roles in tumor immune evasion, neurodegenerative diseases and viral infections. In the present study, single-cell sequencing has been employed to provide references for the diagnosis and prognosis of sepsis in subsequent clinical applications by identifying biomarkers associated with sepsis.

The PSAP gene is responsible for encoding SapC, a form of GCase activator linked to genetic disorders.[35] PSAP, as a crucial biomarker and functional protein, plays a pivotal role in various biological processes, including lipid metabolism, cell signaling and its association with multiple diseases.

① The role of PSAP in tumor biology has sparked significant interest among researchers. Studies have displayed that the expression levels of PSAP in gastric cancer are closely linked to patient prognosis. The poorer prognosis typically portended by the elevated PSAP expression has made PSAP a potential biomarker for the efficacy assessment of immunotherapy and prognostic evaluation in gastric cancer patients.[36] The study of PSAP in neurodegenerative diseases has been gradually piling up. According to the previous research, variations in the PSAP gene are related with specific types of neurodegenerative diseases, such as subacute leukoencephalopathy. These changes may have effects on survival and functionality of nerve cells by leading to the loss of PSAP function.[37] Furthermore, the application of PSAP in health care and disease management has drawn attention. Studies have displayed that PSAP levels are correlated with the risk of developing various chronic diseases, particularly among the elderly population. By monitoring changes in PSAP, more personalized care plans can be provided for chronic disease patients to enhance their quality of life.[38]

Therefore, the research direction of PSAP involves multiple fields, such as tumor biology, neuroscience, animal breeding and healthcare. These studies have not only deepened our understanding of PSAP's functions, but also provided new insights and methods for the prevention and treatment of related diseases. Through single-cell sequencing, we have gained insight into the expression status in sepsis and its potential roles, thereby compensating for its functional status under extensive inflammatory conditions.

S100A11, a low molecular weight protein ranging from 9–14 KD, belongs to the S100 protein family, which contains a Ca2+ binding EF-hands motif. According to the relevant research, the family members matter both inside and outside the cell, As one of S 100 proteins, S100A11 is expressed in many tissues, but the expression level varies with tissues.[39] Most cancer studies have shown that tumor growth and metastasis are promoted by cancer diseases.[40] Compared to normal tissues, tumor tissues exhibit distinct expression patterns.[41,42] S100A11, as a calcium-binding protein, has induced great attention in recent studies across a variety of tumors. By single-cell technology, researchers have delved into its expressions and functions across different cell types, particularly its role within the tumor microenvironment.

Notably, S100A11 has been discovered to interact closely with cancer-associated fibroblasts (CAFs), which may influence the immune microenvironment of the tumor by affecting T-cell infiltration. Moreover, S100A11 has been demonstrated to have impacts on tumor progression and therapeutic resistance by promoting the recruitment of macrophages in breast cancer.[43] These findings have suggested that S100A11 not only functions within tumor cells, but also affects tumor biological behavior by modulating immune cells within the tumor microenvironment.

② The application of single-cell RNA sequencing technology has enabled researchers to identify specific expression patterns of S100A11 under different disease conditions. For instance, in studies on IgG4-related disease and Kimura's disease, the scholars have observed distinct cellular characteristics among different T follicular helper (Tfh) cell populations in these two conditions, which might be concerned with variations in S100A11 expression.[44]

Single-cell studies have unveiled its intricate roles in various tumors and related diseases. In our research, the leveraged single-cell sequencing technology has been employed to provide strategies for subsequent mechanistic studies by elucidating its distribution within cells. Nevertheless, different from the previous approaches, our focus has solely shifted towards identifying biomarkers specific to sepsis.

①: LST1, a gene located in the human MHC III region,[45,46] is known for its increased expression in inflammatory and autoimmune diseases. It has encompassed the higher levels of LST1 splicing variants in patients with rheumatoid arthritis, as well as induction of LST1 expression by IFN-γ, LPS treatment and Pseudomonas aeruginosa infection in cell lines.[47] Animal experiments have demonstrated that MHC III haplotype with low LST1 expression in rat lymph nodes was associated with decreased arthritis severity.[48] LST1 (Leukocyte-Specific Transcript 1) plays a crucial role in immune regulation and intercellular communication, especially in the context of intestinal inflammation and autoimmune diseases:: LST1 expression significantly increases under inflammatory conditions, which may be related to its function in regulating intestinal immunity and inflammatory responses.[49]

②: In studies of rheumatoid arthritis (RA), the expression levels of LST1 correlates with patients' responses to glucocorticoids. Research has found that before treatment, there is significantly higher LST1 expression in patients who respond well to glucocorticoids, suggesting that LST1 may serve as a biomarker for predicting clinical responses.[50] ①: The crucial interplay of LST1 with other immune-modulating genes in maintaining immune balance has offered new insights into the pathogenesis of autoimmune diseases.[51] In parallel with this research, the utilization of single-cell sequencing technology in this study has revealed its high expression under septic conditions is related with improved survival rates, which not only emphasizes its significance in immune regulation, but also provides potential goals for developing novel therapeutic strategies.

SERPINA1: Located in human chromosome 14q32.1, it encodes an acute-phase glycoprotein expressing α −1 AAT1 that is soluble in water and can diffuse in tissues. Approximately 80% synthesis of AAT1 occurs in hepatocytes, where it functions as a multifunctional protein and plays a crucial role in inhibiting serine proteases in human serum. Its working principle is to use the amino acids methionine and serine found in the active area of AAT1, which is primarily active in the lungs. This helps protect the delicate connective tissues in the lower respiratory tract from the uncontrolled proteolytic AAT1 gene triggered by neutrophils during inflammation, which is activated by inflammation products.[49,52,53] In recent years, the single-cell RNA sequencing technology has enabled us to delve into the expression differences of SERPINA1 across various cell types and its association with diseases:

① The expression levels of SERPINA1 in patients with non-small cell lung cancer (NSCLC) are closely linked to their survival rates. Research has indicated the low expression of SERPINA1 in tumor tissues is associated with poorer prognoses.[54]

② The role of SERPINA1 in chronic obstructive pulmonary disease (COPD) has also garnered attention. A study has revealed there is a significant interaction between single nucleotide polymorphisms (SNPs) in the SERPINA1 genes and smoking, which potentially impact an individual's risk of developing COPD.[55]

Similarly, the single-cell technology has been adopted to identify biomarkers associated with sepsis with the aim of providing valuable insights for subsequent studies on their mechanisms of action. With this technology, we can gain a deeper understanding of SERPINA1's function and its potential applications in diseases to offer fresh perspectives for future therapeutic strategies.

As a crucial controller in various human cancers,[56,57] IFI30 is also considered a predictive biomarker for certain types of cancers.[58,59] S100A8, also known as S100a8/a9, is primarily a heterodimeric Ca2+-binding protein that is secreted by neutrophils, with the capability to trigger long-lasting inflammation and damage to the endothelium.[60–62] IFI30 (Interferon Gamma-Induced Protein 30) matters in tumor immunotherapy, particularly in glioma research. Recent studies have revealed that IFI30 not only plays a crucial role in the immune response against tumors, but is also closely associated with patient prognosis and treatment response:

①: Analysis of data from the Chinese Glioma Genome Atlas (CGGA) and the Cancer Genome Atlas (TCGA) has shown that the high expression of IFI30 in malignant glioma subtypes is concerned with chemotherapy response, which provides a new direction for the treatment of gliomas.[63]

②: The expression of IFI30 in breast cancer has also revealed its potential as a prognostic marker. Studies have suggested that high expression of IFI30 is associated with poor prognosis in patients with breast cancer. Furthermore, the scholars have found that inhibiting the expression of IFI30 can significantly suppress the proliferation, migration and invasive ability of breast cancer cells.[64]

③: The role of IFI30 in immune-mediated pathological research has also received extensive attention. Deficiencies in certain immune responses may also lead to the immunopathology. This notion is called the IMPATH paradox, highlighting that under specific circumstances, immune stimulation or immune reconstitution could be more effective than immunosuppression.[57]

In summary, IFI30, as a vital immune-related target, needs to be investigated in depth for its role in various types of tumors. In contrast, in this study, high expression levels of IFI3 have been found to be associated with higher survival rates, which may be due to its participation in early immune-mediated suppression within the organism, suggesting its potential application in immunotherapy, as further research has revealed.

According to previous studies, S100A8/A9 tends to destroy mitochondrial stability by inducing disturbances in mitochondrial metabolism. Studies have demonstrated that S100A8/A9-specific neutrophils exist in the lung tissues of septic mice. In the previous single-cell studies, the expression and function of S100A8 have garnered increasing attention.

①: Researchers have discovered that S100A8 was significantly upregulated in acute myeloid leukemia (AML) cells through the single-cell transcriptomic analysis, which was associated with chemoresistance. It has indicated that bone marrow stromal cells can promote the expression of S100A8 by secreting interleukin-6 (IL-6), thereby influencing the survival and resistance of AML cells.[65]

②: S100A8 also matters in the immune responses of patients with systemic lupus erythematosus (SLE). Studies have found that plasmacytoid dendritic cells in SLE patients can synthesize S100A8, whose surface expression significantly increases during active disease states. This suggests that S100A8 may contribute to the amplification and perpetuation of inflammation in SLE.[66]

③: In Alzheimer's disease (AD) research, single-cell RNA sequencing has identified a specific subset of S100A8-positive neutrophils rich in the peripheral blood of AD mice. The characteristics exhibited by these cells are concerned with the downregulation of cytokine-mediated signaling pathways, suggesting that S100A8 may be of great significance in the pathological process of AD.[67]

Consequently, S100A8 has demonstrated its role as a biomarker and a potential therapeutic target in single-cell studies across various diseases. This study has employed the single-cell technology to delve into the developmental patterns of

sepsis, which are similar to these efforts. It has offered novel insights for future clinical applications founded on the deeper understanding of the functional changes of S100A8 under septic conditions.

The differential gene pathway enrichment analysis has suggested that most genes mainly focus on the activation reaction of the immune system caused by infectious diseases, which further promotes the development of inflammatory immunity. This study has displayed that S100A8has a definite role in sepsis, like participating in the process of inflammatory reaction, promoting the release of inflammatory factors and inducing the occurrence and development of inflammatory diseases. In some autoimmune inflammatory diseases, the expression of LST1 may increase, which indicates that it may be in the "low valley" stage of the disease when the gene expression is low. Most of the remaining genes have been reported to be related to the development of tumor diseases, and few have been reported in inflammatory diseases or even sepsis.

Based on the visual expression of individual cells, the levels of these seven genes in the macrophage system are notably elevated compared to other cell lines,[68] which is because the mononuclear-macrophage system, an integral part of the human body's innate immune system, is crucial in the progression of sepsis. Therefore, the high expression of these genes in macrophages may predict the prognosis of inflammation, SIRS and even sepsis.

According to the dataset, the high expression levels of LILRB2, PSAP, S100A11, LST1, SERPINA1 and IFI30 have been linked to a notably rising survival rate compared to low expression levels, but with S100A8 of the opposite pattern. When the inflammatory factors produced by mononuclear-macrophages are in the cascade amplification reaction of "waterfall effect", the patient will die within several days. Therefore, it can be inferred that when the S100A8 is in the over-expression state detected in clinical practice, the prognosis of the patient is poor and sepsis death occurs. Hence, during the initial phase of inflammation, monitoring the gene content in the blood can offer essential guidance for patient prognosis and treatment adjustments.

In the previous related studies, there have been corresponding research achievements to describe the changes in transcription status of specific cell types. For example, the sixteen immune cell states have been defined on the basis of analyzing clinical blood samples with single-cell sequencing technology and clustering by gene expression profile, and a unique CD14+ monocyte state was determined. In addition, the effect of characteristic gene expression of the cell line on the occurrence and development of sepsis has been verified from a wide range of perspectives.[69] In our study, 12 cell-line communities were identified by the same dimensionality reduction clustering method, four of which were discussed and analyzed.

In order to understand the inflammatory characteristics of peripheral blood mononuclear cells, the single-cell transcriptome comprehensive analysis method has been used to characterize three different subtypes of cells in different disease states and determine the ten subtypes of highly inflammatory cells, which indicated that the transcription difference in monocytes was the most important marker of infectious diseases. Moreover, the construction of "three-stage" model method of heterogeneity of COVID-19 patients has revealed that the high inflammatory and immunosuppressive characteristics of monocytes were related to.[70] For the purpose of this study, the distribution range and size of the 12 different cell subpopulations under two disease states were determined by single-cell sequencing technique. In the four main cell lines, the macrophage community has affected the disease process of sepsis by releasing a large number of inflammatory factors, which finally predicted the prognosis of sepsis by determining the differences of differential gene expression of macrophage line.

Another research project has put emphasis on enhancing the prospects for sepsis by examining immediate shifts in immune cell populations and their traits in cases of survivals or fatal sepsis, followed by monitoring changes in the expression of certain cell types over time to forecast sepsis results.[71] Conversely, this investigation has identified a marker gene for predicting sepsis prognosis after delving into the analysis and discussion of single-cell sequencing findings.

This research emphasis has lied in the activation and development of human tissues or cells through inflammatory stimuli. Assessing the advancement and likelihood of associated illnesses during substantial alterations in gene expression is critical for mitigating the risk of premature death and enhancing prognostication. By monitoring the alteration in gene expression patterns, diseases can be predicted and diagnosed early, and timely intervention can be allowed to stop progression or deterioration, thereby offering novel insights and theoretical foundation for early disease prediction. In the past, the blood system has been thoroughly studied with mature flow cytometry and transcriptomics methods[72,73] to reveal functional changes in immune cells at one level. Nevertheless, impartial, high-capacity gene expression analyses

with traditional flow cytometry and transcriptomics have failed to identify alterations in various cell subsets. The study of individual cell gene expression has provided a distinct perspective on the genetic characteristics and diversity of immune cells in human blood.[74,75]

However, due to the lack of experimental validation of key genes, there are some limitations of the experimental conclusions. As far as we know, there has been no research discussing the diagnostic and prognostic roles of biomarkers such as LILRB2 and S100A8 in septic patients, which has made our research relatively novel. Consequently, we suggest conducting further studies to corroborate our analysis results.

Although our study has underscored the importance of sepsis-related biomarkers (LILRB2, S100A8, and seven other genes) in potentially identifying sepsis and predicting its outcomes, it is undeniable that the research has several limitations, encompassing a small sample size, the inability to establish relationships between gene expression and disease caused by different bacterial infections, and relatively weak handling of confounding factors. Consequently, the future research will consider more multicenter sample statistics and integrate these potential variable factors into its models to ensure that the reported estimates are not affected by other variables. Secondly, because of lacking in vitro or in vivo experimental conditions, the conclusions of this study are currently limited to theoretical validation using external datasets, resulting in a singularity of findings. This necessitates subsequent corroboration when experimental conditions are sufficiently mature. Therefore, upon completing the multicenter and multiregional sample collection, we will be able to draw more precise conclusions. Finally, we also need to gather larger and more extensive case-control sample studies to further validate our observations.

In this study, GO analysis of the two genes has suggested that the marker genes were mainly concentrated in the cell growth stage and mostly in the cytoplasm. KEGG analysis has displayed that the target, which is mainly related to the immune system, can be reflected in the development of infectious diseases. Furthermore, regulating the macrophage functions has impacted the onset and progression of the illness.

Conclusion

The present study has indicated the status of 12 clusters of four major cell types in the blood sample and explained the differences in major cell populations between the septic death group and the SIRS group. Additional genetic markers can be discovered to illustrate variations in gene expression among multiple groups and create visual representation of cell distribution. Enrichment analysis has been adopted to elucidate the functional expression status and possible signaling pathways of related genes. The survival curve has revealed that the expressions of different genes in various groups have effects on the survival rate. No, the deficiency is that the functional changes of different genes in the relevant groups cannot be further revealed by single-cell sequencing.

Data Sharing Statement

The original data is stored in the China National GeneBank DataBase (CNGBdb) and can be accessed at the following link: http://db.cngb.org/, under the accession number: CNP0002611. You may access it now and it will remain valid indefinitely. Supporting raw data are available upon request from the corresponding author. Source data are provided with this article.

Ethics Approval and Consent to Participate

1. Prior to conducting the experimental research, my research group has obtained informed consent from the relevant participants and signed informed consent forms.
2. The study was conducted in strict accordance with the rules of the Declaration of Helsinki. The study protocol has been approved by the ethics committee of the Affiliated Hospital of Southwest Medical University (Ethical Approval No. ky2018029). The Registration Number was ChiCTR1900021261.

Acknowledgment

We thank OEBIOTECH for instructing scRNA-seq.

Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Funding

This study was funded by Sichuan Provincial Key clinical Specialty Construction Project. Mechanistic Study on the Improvement of Sepsis Outcome by 18α-Glycyrrhetinic Acid via CDK2-CCND1(2024LZXNYDJ038).

Disclosure

Authors state no conflict of interest.

References

1. Cecconi M, Evans L, Levy M, Rhodes A. Sepsis and septic shock. *Lancet*. 2018;392(10141):75–87. PMID: 29937192. doi:10.1016/S0140-6736(18)30696-2
2. Luhr R, Cao Y, Söderquist B, Cajander S. Trends in sepsis mortality over time in randomized sepsis trials: a systematic literature review and meta-analysis of mortality in the control arm, 2002-2016. *Crit Care*. 2019;23(1):241. PMID: 31269976. doi:10.1186/s13054-019-2528-0
3. Wong A, Otles E, Donnelly JP, et al. External validation of a widely implemented proprietary sepsis prediction model in hospitalized patients. *JAMA Intern Med*. 2021;181(8):1065–1070. PMID: 34152373. doi:10.1001/jamainternmed.2021.2626
4. Romagnoli S, Ricci Z, Ronco C. CRRT for sepsis-induced acute kidney injury. *Curr Opin Crit Care*. 2018;24(6):483–492. PMID: 30239411. doi:10.1097/MCC.0000000000000544
5. Sharma D, Farahbakhsh N, Shastri S, Sharma P. Biomarkers for diagnosis of neonatal sepsis: a literature review. *J Matern Fetal Neonatal Med*. 2018;31(12):1646–1659. PMID: 28427289. doi:10.1080/14767058.2017.1322060
6. Jiang S, Yang Z, Shan R, et al.; Reduction of Infection in Neonatal Intensive Care Units using the Evidence-based Practice for Improving Quality Study Group. Neonatal outcomes following culture-negative late-onset sepsis among preterm infants. *Pediatr Infect Dis J*. 2020;39(3):232–238. PMID: 32032307.
7. El-Khazragy N, Mohamed NM, Mostafa MF, et al. miRNAs: novel noninvasive biomarkers as diagnostic and prognostic tools in neonatal sepsis. *Diagne Microbial Infect Dis*. 2023;107(3):116053. PMID: 37659119. doi:10.1016/j.diagmicrobio.2023.116053
8. Lim B, Lin Y, Navin N. Advancing cancer research and medicine with single-cell genomics. *Cancer Cell*. 2020;37(4):456–470. PMID: 32289270. doi:10.1016/j.ccell.2020.03.008
9. Dagogo-Jack I, Shaw AT. Tumour heterogeneity and resistance to cancer therapies. *Nat Rev Clin Oncol*. 2018;15(2):81–94. PMID: 29115304. doi:10.1038/nrclinonc.2017.166
10. Sweeney TE, Khatri P. Benchmarking sepsis gene expression diagnostics using public data. *Crit Care Med*. 2017;45(1):1–10. PMID: 27681387. doi:10.1097/CCM.0000000000002021
11. Sweeney TE, Perumal TM, Henao R, et al. A community approach to mortality prediction in sepsis via gene expression analysis. *Nat Commun*. 2018;9(1):694. PMID: 29449546. doi:10.1038/s41467-018-03078-2
12. Davenport EE, Burnham KL, Radhakrishnan J, et al. Genomic landscape of the individual host response and outcomes in sepsis: a prospective cohort study. *Lancet Respir Med*. 2016;4(4):259–271. PMID: 26917434. doi:10.1016/S2213-2600(16)00046-1
13. Sweeney TE, Shidham A, Wong HR, Khatri P. A comprehensive time-course-based multicohort analysis of sepsis and sterile inflammation reveals a robust diagnostic gene set. *Sci Transl Med*. 2015;7(287):287ra71. PMID: 25972003. doi:10.1126/scitranslmed.aaa5993
14. Reyes M, Vickers D, Billman K, et al. Multiplexed enrichment and genomic profiling of peripheral blood cells reveal subset-specific immune signatures. *Sci Adv*. 2019;5(1):eaau9223. PMID: 30746468. doi:10.1126/sciadv.aau9223
15. Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013;29(1):15–21. PMID: 23104886. doi:10.1093/bioinformatics/bts635
16. Aran D, Looney AP, Liu L, et al. Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. *Nat Immunol*. 2019;20(2):163–172. PMID: 30643263. doi:10.1038/s41590-018-0276-y
17. Kanehisa M, Araki M, Goto S, et al. KEGG for linking genomes to life and the environment. *Nucleic Acids Res*. 2008;36(Database issue):D480–4. PMID: 18077471; PMCID. doi:10.1093/nar/gkm882
18. Szklarczyk D, Gable AL, Lyon D, et al. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res*. 2019;47(D1):D607–D613. PMID: 30476243. doi:10.1093/nar/gky1131
19. Rudd KE, Johnson SC, Agesa KM, et al. Global, regional, and national sepsis incidence and mortality, 1990-2017: analysis for the global burden of disease study. *Lancet*. 2020;395(10219):200–211. PMID: 31954465. doi:10.1016/S0140-6736(19)32989-7
20. Buchman TG, Simpson SQ, Sciarretta KL, et al. Sepsis among medicare beneficiaries: 1. the burdens of sepsis, 2012-2018. *Crit Care Med*. 2020;48(3):276–288. PMID: 32058366. doi:10.1097/CCM.0000000000004224
21. Kim J, Stanescu DE, Won KJ. CellBIC: bimodality-based top-down clustering of single-cell RNA sequencing data reveals hierarchical structure of the cell type. *Nucleic Acids Res*. 2018;46(21):e124. PMID: 30102368. doi:10.1093/nar/gky698
22. Lubatti G, Stock M, Iturbide A, et al. CIARA: a cluster-independent algorithm for identifying markers of rare cell types from single-cell sequencing data. *Development*. 2023;150(11):dev201264. PMID: 37294170. doi:10.1242/dev.201264

23. Chen W, Morabito SJ, Kessenbrock K, Enver T, Meyer KB, Teschendorff AE. Single-cell landscape in mammary epithelium reveals bipotent-like cells associated with breast cancer risk and outcome. *Commun Biol*. 2019;2(1):306. PMID: 31428694. doi:10.1038/s42003-019-0554-8
24. Fletcher RB, Das D, Ngai J. Creating lineage trajectory maps via integration of single-cell RNA-sequencing and lineage tracing: integrating transgenic lineage tracing and single-cell RNA-sequencing is a robust approach for mapping developmental lineage trajectories and cell fate changes. *Bioessays*. 2018;40(8):e1800056. PMID: 29944188. doi:10.1002/bies.201800056
25. Liu X, Mu Y. Lymphocyte to C-reactive protein ratio as an early biomarker to distinguish sepsis from pneumonia in neonates. *J Inflamm Res*. 2023;16:3509–3517. PMID: 37608883. doi:10.2147/JIR.S424897
26. Li DP, Huang L, Kan RR, et al. LILRB2/PirB mediates macrophage recruitment in fibrogenesis of nonalcoholic steatohepatitis. *Nat Commun*. 2023;14(1):4436. PMID: 37481670. doi:10.1038/s41467-023-40183-3
27. Deng M, Lu Z, Zheng J, et al. A motif in LILRB2 critical for Angptl2 binding and activation. *Blood*. 2014;124(6):924–935. PMID: 24899623. doi:10.1182/blood-2014-01-549162
28. Wang Y, Johnson KCC, Gatti-Mays ME, Li Z. Emerging strategies in targeting tumor-resident myeloid cells for cancer immunotherapy. *J Hematol Oncol*. 2022;15(1):118. PMID: 36031601.
29. Chen HM, van der Touw W, Wang YS, et al. Blocking immunoinhibitory receptor LILRB2 reprograms tumor-associated myeloid cells and promotes antitumor immunity. *J Clin Invest*. 2018;128(12):5647–5662. PMID: 30352428. doi:10.1172/JCI97570
30. Zhang Z, Yuan Y, Hu L, et al. ANGPTL8 accelerates liver fibrosis mediated by HFD-induced inflammatory activity via LILRB2/ERK signaling pathways. *J Adv Res*. 2023;47:41–56. PMID: 36031141. doi:10.1016/j.jare.2022.08.006
31. Zhao P, Xu Y, Jiang LL, et al. LILRB2-mediated TREM2 signaling inhibition suppresses microglia functions. *Mol Neurodegener*. 2022;17(1):44. PMID: 35717259. doi:10.1186/s13024-022-00550-y
32. Jiang Z, Huang Q, Chang Y, et al. LILRB2 promotes immune escape in breast cancer cells via enhanced HLA-A degradation. *Cell Oncol Dordr*. 2024;47(5):1679–1696. PMID: 38656573. doi:10.1007/s13402-024-00947-5
33. Yue J, Li W, Liang C, et al. Activation of LILRB2 signal pathway in temporal lobe epilepsy patients and in a pilocarpine induced epilepsy model. *Exp Neurol*. 2016;285(Pt A):51–60. PMID: 27637803. doi:10.1016/j.expneurol.2016.09.006
34. Bashirova AA, Martin-Gayo E, Jones DC, et al. LILRB2 interaction with HLA class I correlates with control of HIV-1 infection. *PLoS Genet*. 2014;10(3):e1004196. PMID: 24603468. doi:10.1371/journal.pgen.1004196
35. Pavan E, Peruzzo P, Cattarossi S, et al. Deficiency of glucocerebrosidase activity beyond gaucher disease: PSAP and LIMP-2 dysfunctions. *Int J mol Sci*. 2024;25(12):6615. PMID: 38928321. doi:10.3390/ijms25126615
36. Wen Z, Yang C, Zou D, et al. Pan-cancer analysis of PSAP identifies its expression and clinical relevance in gastric cancer. *Pathol Res Pract*. 2022;238:154027. PMID: 36084426. doi:10.1016/j.prp.2022.154027
37. Kolnikova M, Jungova P, Skopkova M, et al. Late infantile metachromatic leukodystrophy due to novel pathogenic variants in the PSAP gene. *J Mol Neurosci*. 2019;67(4):559–563. PMID: 30632081. doi:10.1007/s12031-019-1259-7
38. Meng LC, Huang ST, Chen HM, Hashmi AZ, Hsiao FY, Chen LK. Health care utilization and potentially preventable adverse outcomes of high-need, high-cost middle-aged and older adults: needs for integrated care models with life-course approach. *Arch Gerontol Geriatr*. 2023;109:104956. PMID: 36804698. doi:10.1016/j.archger.2023.104956
39. Zhang L, Zhu T, Miao H, Liang B. The calcium binding protein S100A11 and its roles in diseases. *Front Cell Dev Biol*. 2021;9:693262. PMCID: PMC8226020. doi:10.3389/fcell.2021.693262
40. Bresnick AR, Weber DJ, Zimmer DB. S100 proteins in cancer. *Nat Rev Cancer*. 2015;15(2):96–109. PMID: 25614008. doi:10.1038/nrc3893
41. Meding S, Balluff B, Elsner M, et al. Tissue-based proteomics reveals FXYD3, S100A11 and GSTM3 as novel markers for regional lymph node metastasis in colon cancer. *J Pathol*. 2012;228(4):459–470. PMID: 22430872. doi:10.1002/path.4021
42. Zhang MX, Gan W, Jing CY, et al. S100A11 promotes cell proliferation via P38/MAPK signaling pathway in intrahepatic cholangiocarcinoma. *Mol, Carcinog*. 2019;58(1):19–30. PMID: 30182496. doi:10.1002/mc.22903
43. Han D, Guo C, Cheng H, et al. Downregulation of S100A11 promotes T cell infiltration by regulating cancer-associated fibroblasts in prostate cancer. *Int Immunopharmacol*. 2024;128:111323. PMID: 38286714. doi:10.1016/j.intimp.2023.111323
44. Lee S, Cho Y, Li Y, et al. Cancer-cell derived S100A11 promotes macrophage recruitment in ER+ breast cancer. *Oncoimmunology*. 2024;13(1):2429186. PMID: 39587886. doi:10.1080/2162402X.2024.2429186
45. Munemura R, Maehara T, Murakami Y, et al. Distinct disease-specific Tfh cell populations in 2 different fibrotic diseases: igG4-related disease and Kimura disease. *J Allergy Clin Immunol*. 2022;150(2):440–455.e17. PMCID: PMC10369367. doi:10.1016/j.jaci.2022.03.034
46. Holzinger I, de Baey A, Messer G, Kick G, Zwierzina H, Weiss EH. Cloning and genomic characterization of LST1: a new gene in the human TNF region. *Immunogenetics*. 1995;42(5):315–322. PMID: 7590964. doi:10.1007/BF00179392
47. Nalabolu SR, Shukla H, Nallur G, Parimoo S, Weissman SM. Genes in a 220-kb region spanning the TNF cluster in human MHC. *Genomics*. 1996;31(2):215–222. PMID: 8824804. doi:10.1006/geno.1996.0034
48. Mulcahy H, O'Rourke KP, Adams C, Molloy MG, O'Gara F. LST1 and NCR3 expression in autoimmune inflammation and in response to IFN-gamma, LPS and microbial infection. *Immunogenetics*. 2006;57(12):893–903. PMID: 16362817. doi:10.1007/s00251-005-0057-2
49. Yau AC, Tuncel J, Haag S, et al. Conserved 33-kb haplotype in the MHC class III region regulates chronic arthritis. *Proc Natl Acad Sci USA*. 2016;113(26):E3716–24. PMID: 27303036. doi:10.1073/pnas.1600567113
50. Heidemann J, Kebschull M, Tepasse PR, Bettenworth D. Regulated expression of leukocyte-specific transcript (LST) 1 in human intestinal inflammation. *Inflamm Res*. 2014;63(7):513–517. PMID: 24682411. doi:10.1007/s00011-014-0732-6
51. Fritsch-Stork RD, Silva-Cardoso SC, Groot Koerkamp MJ, Broen JC, Lafeber FF, Bijlsma JW. Expression of ERAP2 and LST1 is increased before start of therapy in rheumatoid arthritis patients with good clinical response to glucocorticoids. *Clin Exp Rheumatol*. 2016;34(4):685–689. Epub 2016 Jun 22. PMID: 27384923.
52. Pini L, Tiberio L, Venkatesan N, et al. The role of bronchial epithelial cells in the pathogenesis of COPD in Z-alpha-1 antitrypsin deficiency. *Respir Res*. 2014;15(1):112. PMID: 25218041. doi:10.1186/s12931-014-0112-3
53. Bergin DA, Reeves EP, Hurley K, et al. The circulating proteinase inhibitor α-1 antitrypsin regulates neutrophil degranulation and autoimmunity. *Sci Transl Med*. 2014;6(217):217ra1. PMID: 24382893. doi:10.1126/scitranslmed.3007116
54. Baraldo S, Turato G, Lunardi F, et al. Immune activation in α1-antitrypsin-deficiency emphysema. Beyond the protease-antiprotease paradigm. *Am J Respir Crit Care Med*. 2015;191(4):402–409. PMID: 25412116. doi:10.1164/rccm.201403-0529OC

55. Ercetin E, Richtmann S, Delgado BM, et al. Clinical significance of SERPINA1 gene and its encoded alpha1-antitrypsin protein in NSCLC. *Cancers*. 2019;11(9):1306. PMID: 31487965. doi:10.3390/cancers11091306
56. Deng X, Yuan CH, Chang D. Interactions between single nucleotide polymorphism of SERPINA1 gene and smoking in association with COPD: a case-control study. *Int J Chron Obstruct Pulmon Dis*. 2017;12:259–265. PMID: 28138235. doi:10.2147/COPD.S116313
57. Fan Y, Wang X, Li Y. IFI30 expression predicts patient prognosis in breast cancer and dictates breast cancer cells proliferation via regulating autophagy. *Int J Med Sci*. 2021;18(14):3342–3352. PMID: 34400904. doi:10.7150/ijms.62870
58. Phipps-Yonas H, Cui H, Sebastiao N, et al. Low GILT expression is associated with poor patient survival in diffuse large B-cell lymphoma. *Front Immunol*. 2013;4:425. PMID: 24409177. doi:10.3389/fimmu.2013.00425
59. Xiang YJ, Fu QY, Ma ZB, et al. Screening for candidate genes related to breast cancer with cDNA microarray analysis. *Chronic Dis Transl Med*. 2015;1(2):65–72. PMID: 29062989. doi:10.1016/j.cdtm.2015.02.001
60. Liu X, Song C, Yang S, et al. IFI30 expression is an independent unfavourable prognostic factor in glioma. *J Cell Mol Med*. 2020;24 (21):12433–12443. PMID: 32969157. doi:10.1111/jcmm.15758
61. Zhang H, Wu D, Wang Y, et al. METTL3-mediated N6-methyladenosine exacerbates ferroptosis via m6A-IGF2BP2-dependent mitochondrial metabolic reprogramming in sepsis-induced acute lung injury. *Clin Transl Med*. 2023;13(9):e1389. PMID: 37715457. doi:10.1002/ctm2.1389
62. Li Y, Chen B, Yang X, et al. S100a8/a9 signaling causes mitochondrial dysfunction and cardiomyocyte death in response to ischemic/reperfusion injury. *Circulation*. 2019;140(9):751–764. PMID: 31220942. doi:10.1161/CIRCULATIONAHA.118.039262
63. Wang Y, Shi Y, Shao Y, Lu X, Zhang H, Miao C. S100A8/A9hi neutrophils induce mitochondrial dysfunction and PANoptosis in endothelial cells via mitochondrial complex I deficiency during sepsis. *Cell Death Dis*. 2024;15(6):462. PMID: 38942784. doi:10.1038/s41419-024-06849-6
64. Zhu C, Chen X, Guan G, et al. IFI30 is a novel immune-related target with predicting value of prognosis and treatment response in glioblastoma. *Onco Targets Ther*. 2020;13:1129–1143. PMID: 32103982. doi:10.2147/OTT.S237162
65. Ehl S, Thimme R. Immune-mediated pathology as a consequence of impaired immune reactions: the IMPATH paradox. *Eur J Immunol*. 2022;52 (9):1386–1389. PMID: 35934899. doi:10.1002/eji.202250069
66. Böttcher M, Panagiotidis K, Bruns H, et al. Bone marrow stroma cells promote induction of a chemoresistant and prognostic unfavorable S100A8/A9high AML cell subset. *Blood Adv*. 2022;6(21):5685–5697. PMID: 35390134. doi:10.1182/bloodadvances.2021005938
67. Lood C, Stenström M, Tydén H, et al. Protein synthesis of the pro-inflammatory S100A8/A9 complex in plasmacytoid dendritic cells and cell surface S100A8/A9 on leukocyte subpopulations in systemic lupus erythematosus. *Arthritis Res Ther*. 2011;13(2):R60. PMID: 21492422. doi:10.1186/ar3314
68. Zhang X, He G, Hu Y, et al. Single cell transcriptome analysis identified a unique neutrophil type associated with Alzheimer's disease. *Immun Ageing*. 2024;21(1):42. PMID: 38918830. doi:10.1186/s12979-024-00448-x
69. Aparicio-Vergara M, Tencerova M, Morgantini C, Barreby E, Aouadi M. Isolation of Kupffer cells and hepatocytes from a single mouse liver. *Methods mol Biol*. 2017;1639:161–171. PMID: 28752456.
70. Reyes M, Filbin MR, Bhattacharyya RP, et al. An immune-cell signature of bacterial sepsis. *Nat Med*. 2020;26(3):333–340. PMID: 32066974. doi:10.1038/s41591-020-0752-4
71. Liu N, Jiang C, Cai P, et al. Single-cell analysis of COVID-19, sepsis, and HIV infection reveals hyperinflammatory and immunosuppressive signatures in monocytes. *Cell Rep*. 2021;37(1):109793. PMID: 34587478. doi:10.1016/j.celrep.2021.109793
72. Qiu X, Li J, Bonenfant J, et al. Dynamic changes in human single-cell transcriptional signatures during fatal sepsis. *J Leukoc Biol*. 2021;110 (6):1253–1268. PMID: 34558746. doi:10.1002/JLB.5MA0721-825R
73. Qi F, Wei Y, Wu M, et al. Immunotyping of peripheral blood lymphocytes by flow cytometry reveals Th cell as a potential prognostic biomarker for extranodal NK/T-cell lymphoma. *Ann Hematol*. 2024;103(5):1643–1653. PMID: 38191717. doi:10.1007/s00277-023-05605-8
74. Cuyutupa VR, Moser D, Diedrich V, et al. Blood transcriptomics mirror regulatory mechanisms during hibernation-a comparative analysis of the Djungarian hamster with other mammalian species. *Pflugers Arch*. 2023;475(10):1149–1160. PMID: 37542567. doi:10.1007/s00424-023-02842-8
75. Hu J, Guo Q, Liu C, et al. Immune cell profiling of preeclamptic pregnant and postpartum women by single-cell RNA sequencing. *Int Rev Immunol*. 2024;43(1):1–12. PMID: 36369864. doi:10.1080/08830185.2022.2144291