

RESEARCH

Open Access



# Screening of four lysosome-related genes in sepsis based on RNA sequencing technology

Guihong Chen<sup>1,2</sup>, Wen Zhang<sup>3</sup>, Chenglin Wang<sup>2</sup>, Muhu Chen<sup>2</sup>, Yingchun Hu<sup>2</sup> and Zheng Wang<sup>1\*</sup>

## Abstract

**Purpose** Screening of lysosome-related genes in sepsis patients to provide direction for lysosome-targeted therapy.

**Methods** Peripheral blood samples were obtained from 22 patients diagnosed with sepsis and 10 normal controls for the purpose of RNA sequencing and subsequent analysis of differential gene expression. Concurrently, lysosome-related genes were acquired from the Gene Ontology database. The intersecting genes between the differential genes and lysosome-related genes were then subjected to PPI, GO and KEGG analyses. Core genes were identified through survival analysis, and their expression trends in different groups were determined using meta-analysis. Single-cell RNA sequencing was used to clarify the cellular localization of core genes.

**Results** The intersection of 1328 sepsis-differential genes with 878 lysosome-related genes yielded 76 genes. PPI analysis showed that intersecting genes were mainly involved in Cellular process, Response to stimulus, Immune system process, Signal transduction, Lysosome. GO and KEGG analysis showed that intersecting genes were mainly involved in leukocyte mediated immunity, cell activation involved in immune response, lytic vacuole, lysosome. Survival analysis screened four genes positively correlated with sepsis prognosis, namely GNLY, GZMB, PRF1 and RASGRP1. The meta-analysis revealed that the expression levels of these four genes were significantly higher in the normal control group compared to the sepsis group, which aligns with the findings from RNA sequencing data. Furthermore, single-cell RNA sequencing demonstrated that T cells and NK cells exhibited high expression levels of GNLY, GZMB, PRF1, and RASGRP1.

**Conclusion** GNLY, GZMB, PRF1, and RASGRP1, which are lysosome-related genes, are closely linked to the prognosis of sepsis and could potentially serve as novel research targets for sepsis, offering valuable insights for the development of lysosome-targeted therapy. The clinical trial registration number is ChiCTR1900021261, and the registration date is February 4, 2019.

**Keywords** RNA sequencing, Lysosome-related genes, Sepsis, Prognosis

## Introduction

Sepsis is a life-threatening organ dysfunction caused by a malfunctioning host response to infection. It is characterized by an acute cytokine storm followed by long-term dysfunction of the immune system in survivors [1]. Due to its high mortality rate and high healthcare costs, sepsis remains one of the most difficult diseases to treat. Overactivation of the immune system and a cascade of inflammation responses in the early stages of sepsis are often accompanied by immunosuppression. The core pathogenesis of sepsis is dysregulation of host innate and

\*Correspondence:

Zheng Wang  
zheng.wang11@mail.xjtu.edu.cn

<sup>1</sup> Department of Hepatobiliary Surgery, The First Affiliated Hospital of Xi'an Jiaotong University, Xi'an, Shaanxi, China

<sup>2</sup> Department of Emergency Medicine, The Affiliated Hospital of Southwest Medical University, Luzhou, Sichuan, China

<sup>3</sup> Department of Endocrinology and Metabolism, The Traditional Chinese Medicine Hospital of Luzhou City, Luzhou, Sichuan, China



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

adaptive immune responses [2]. Early treatment of sepsis includes source control, antimicrobial, and resuscitation [3]. Despite significant advances have been made in the pathogenesis of sepsis, there is currently no clinically effective treatment [4].

Lysosomes are membrane-enclosed vesicle organelle found in all eukaryotic cells and contain two classes of proteins necessary for maintaining structure and function: soluble lysosomal hydrolases, which perform digestive functions, and lysosomal membrane proteins, which have more complex functions. Lysosomes are organelles responsible for degradation, nutrient sensing, and immunity. Recently, more and more studies have shown that lysosomes are also important in the pathological of many diseases [5]. Existing studies have proved that lysosomes are closely related to atherosclerosis [6], neurodegenerative diseases [7], tumor [8] and other diseases. The body's immunity is affected by lysosomal activity in cells such as macrophages and T cells. For innate immunity, pathogens such as bacteria are internalized by phagocytosis and degraded by targeting lysosomes [9, 10]. In addition, several toll-like receptors (TLRs) on the lysosomal membrane are capable of recognizing various micro-organisms and host-derived ligands and eliciting pro-inflammatory signaling [11]. For adaptive immunity, lysosomes produce antigenic peptides for presentation to CD4+ T cells by major histocompatibility complex class II (MHC-II) molecules. In this process, TLR4 signal-induced tubes of phagosomes and lysosomes are important for antigen presentation [12, 13]. It can be seen that lysosomes are widely involved in the process of innate and adaptive immune response, and targeted lysosomal therapy may regulate the immune disorders caused by sepsis, improve the prognosis of sepsis patients, and is expected to become a new treatment direction for sepsis. However, in the field of sepsis treatment, therapeutic research targeting lysosomes is lacking.

Therefore, we intend to use RNA sequencing technology and bioinformatics analysis to screen out lysosome-related genes that contribute to the prognosis of patients with sepsis, so as to provide direction for the treatment of lysosomes.

## Methods

### Volunteer recruitment and blood collection

From January 2019 to December 2020, 22 sepsis patients in the ICU or EICU of the Affiliated Hospital of Southwest Medical University and 10 normal controls in the same period were included. A sample of peripheral blood was collected within 24 h of the patient's admission. The sepsis group was selected as follows: (1) patients with sepsis hospitalized in ICU or EICU; (2) comply with the Sepsis 3.0 standard published in 2016

by the Society of Critical Care Medicine (SCCM) and European Society of Intensive Care Medicine (ESICM) [14]; (3) the patient or his or her legal representative was willing to enter the experiment and signed the informed consent form. The exclusion criteria were: (1) pregnant or lactating women; (2) suffering from mental illness; (3) immunodeficiency or HIV positive. The inclusion criteria for the normal control group were as follows: (1) healthy volunteers with an age > 18 years; (2) no serious heart, liver, kidney, digestive tract, nervous system diseases; (3) the blood pregnancy test of female volunteers of childbearing age was negative; (4) voluntarily participate in the trial and sign the informed consent form. The exclusion criteria were: (1) those with a history of blood sickness, needle sickness, and severe anemia; (2) have any history of serious clinical disease; (3) those who underwent surgery within 4 weeks before the study; (4) those who have a history of substance abuse within 1 year before the study. This study complies with the Declaration of Helsinki and has been approved by the Ethics Committee of the Affiliated Hospital of Southwest Medical University, ethics number: ky2018029, and the clinical trial registration number: ChiCTR1900021261.

### RNA sequencing

RNA sequencing has become a routine method for transcriptomic profiling. Total RNA was extracted from the peripheral blood mixture using the Trizol reagent. The mixture was subjected to centrifugation at 12,000 rpm for 5 min at 4°C, and the resulting supernatant was then transferred to a new EP tube containing 0.3 mL of isoamyl alcohol. After centrifugation at 12,000 rpm for 10 min at 4°C, the upper aqueous phase containing the RNA was transferred to a new tube and mixed with an equal volume of isopropyl alcohol. This mixture was then subjected to centrifugation at 13,600 rpm for 20 min at 4°C. Following the removal of the supernatant, the RNA pellet underwent two washes with 1 mL of 75% ethanol. To dissolve the RNA, 50µL of DEPC-treated water was introduced. The subsequent assessment and quantification of the total RNA were conducted using a Nano Drop and Agilent 2100 bioanalyzer. The sequencing data underwent filtration using SOAPnuke [15]. This involved the removal of reads containing sequencing adapters, reads with an unknown base ('N' base) ratio exceeding 5%. Subsequently, the clean reads were aligned to the reference genome using HISAT2 [16]. Bowtie2 [17] was then utilized to align the clean reads to the reference coding gene set. Finally, the expression level of genes was calculated using RSEM [18].

### Screening for differential expression genes

Due to its intuitive user interface and large annotation database, the iDEP (integrated differential expression and pathway analysis) website is widely used for interactive analysis of RNA sequencing data [19]. Therefore, we used iDEP 9.6 (<http://149.165.154.220/idepg/>) online platform to quality control and filter the data. Subsequently, box-plot and density distribution analysis were performed on the standardized data to clarify the homogeneity and comparability of the RNA sequencing data. Principal component analysis (PCA) was used to find outlier samples and ensure data stability. The DESeq2 method was adopted for statistical analysis, and the differential genes were screened under the conditions of  $|\text{Fold Change}| (\text{FC}) \geq 4$  and False Discovery Rate (FDR)  $< 0.05$ .

### Screening for lysosome-related genes

The Gene Ontology (GO) knowledgebase (<http://geneontology.org/>) is a comprehensive resource concerning the functions of genes and gene products (proteins and non-coding RNAs) [20]. Therefore, we used the Gene Ontology database to screen lysosome-related genes. Entered the Gene Ontology database, entered "lysosome" in the search box, selected "Gene Product", limited Organism to "Homo sapiens", and downloaded human lysosome-related genes. Lysosome-related genes closely related to sepsis can be obtained by intersecting differential genes with lysosome-related genes.

### PPI analysis

The functional connections between proteins can often be inferred by associations between the genes that encode them, a group of genes with the same function tends to show similar species coverage, is often located close on the genome, and tends to participate in gene fusion events. The STRING database is a precomputed global resource for the exploration and analysis of these associations [21]. The intersecting genes were submitted to the STRING database, "Homo sapiens" was selected in the species option, "full STRING network" was selected in the network type option, "evidence" was selected in the meaning of network edges, "Text mining, Experiments, Databases" were selected in the active interaction sources, the correlation intensity coefficient was set to 0.15, and points in the network that are not connected were removed.

### GO analysis

To further analyze whether the intersecting genes participate in a specific Biological Process (BP), whether they participate in the formation of Cellular Components (CC), and the Molecular Function (MF) in which they are expressed. We used the Omicshare website (<https://www.omicshare.com/>)

to perform GO annotation and taxonomic enrichment analysis to determine the main functions of intersecting genes and comprehensively describe the properties of genes and gene products. The calculated pvalue was corrected by FDR and adjusted-pvalue  $\leq 0.05$  was used as the threshold, and the GO term that met this condition was defined as significantly enriched in differentially expressed genes.

### KEGG analysis

Kyoto Encyclopedia of Genes and Genomes (KEGG) is a knowledge base for systematic analysis of gene functions, linking genomic information with higher-order functional information, which includes metabolic pathway database, hierarchical classification database, gene database, and genomic database, etc. [22]. The Omicshare website was used for KEGG enrichment analysis of intersecting genes to identify the most pronounced pathways in the entire genomic background, with a significant enrichment criterion of adjusted-pvalue  $\leq 0.05$ .

### Survival analysis

Clinical data serves as the fundamental foundation for scientific investigation. To further investigate the correlation between genes and prognosis, the GSE65682 [23] dataset from the GEO public database (<https://www.ncbi.nlm.nih.gov/geo/>) was employed for survival analysis. This dataset encompasses peripheral blood RNA sequencing data from 479 sepsis patients, along with gene expression values and clinical prognosis information for each individual. Survival analysis was conducted using Graphpad prism 7, with a significance threshold set at a logrank test P-value of less than 0.05.

### Meta-analysis

To enhance the precision of core gene screening and gene expression evaluation across various groups, the peripheral blood RNA sequencing data of sepsis patients from the GEO database were procured for meta-analysis. Specifically, the datasets GSE28750 [24], GSE54514 [25], GSE69528 [26], GSE95233 [27] were utilized. Datasets were homogenized (log<sub>2</sub> logarithmic) and categorized into normal control and sepsis, and subsequent meta-analysis was conducted on individual genes within the same grouping across different datasets using R language packages. Random effects model was selected when  $p \leq 0.05$ , and Fixed effect model when  $p > 0.05$ , for multiple data heterogeneity tests.

### Single-cell RNA sequencing

In order to enhance the understanding of the cellular distribution of core genes, the peripheral blood samples

**Table 1** Demographic and clinical characteristics

Item	Sepsis (n=22)	NC (n=10)	P value
Gender (M/F)	14/8	5/5	-
28-Day Finale (S/D)	12/10	10/0	-
Age (years)	56.09±17.51	53.50±7.663	0.6590
ALT (U/L)	90.02±182.5	20.94±6.443	0.2449
AST (U/L)	146.8±277.8	22.18±4.526	0.1703
DBIL (umol/L)	16.74±14.79	5.24±2.048	0.0214
TBIL (umol/L)	32.30±38.48	16.79±6.368	0.2187
Creatinine (umol/L)	106.5±109.2	63.75±9.259	0.2299
Urea (mmol/L)	10.74±9.923	5.046±1.489	0.0838
Leukocyte (10 <sup>9</sup> /L)	12.88±7.166	6.877±1.848	0.0147
Neutrophil (10 <sup>9</sup> /L)	13.67±13.27	4.128±1.151	0.0319
Monocyte (10 <sup>9</sup> /L)	0.8086±1.084	0.443±0.1832	0.3019
Lymphocyte (10 <sup>9</sup> /L)	1.145±1.565	2.02±0.5712	0.0985

The statistical analysis employed the unpaired t-test to assess the Gender (M: male, F: female), 28-day Finale (S: survivor, D: decedent), Age, ALT, AST, DBIL, TBIL, Creatinine, Urea, Leukocyte count, Neutrophil count, Monocyte count, and Lymphocyte count. The mean ± standard deviation was used to express the results

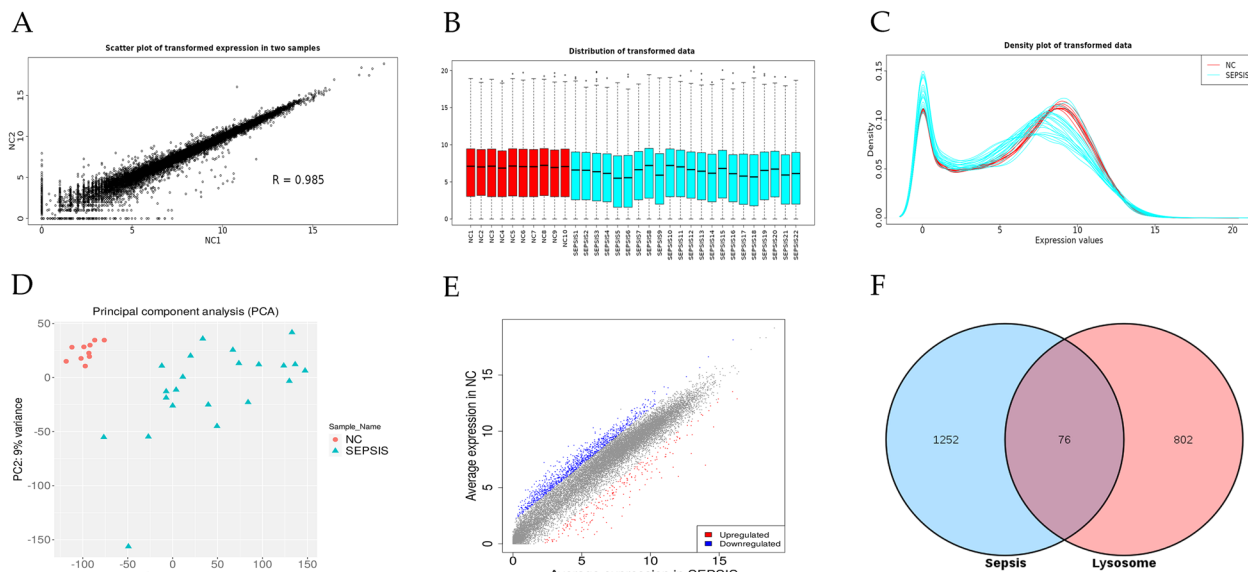
from two healthy individuals, one individual with systemic inflammatory response syndrome (SIRS), and two individuals with sepsis were subjected and analyzed. The Cell Ranger software was employed for quality control measures, ensuring the acquisition of high-quality cell counts, gene counts, and genome alignment. The

utilization of a Unique Molecular Identifier (UMI) and Cell Barcode allowed for the determination of the precise quantity of each transcript molecule within an individual cell. The dimensionality reduction results based on the Mutual Nearest Neighbors (MNN) algorithm were visualized by the t-distributed Stochastic Neighbor Embedding (t-SNE) algorithm, and finally, the optimal cell population classification was obtained. The specific marker gene for each cell population was screened by differentiating the specified cell population from all remaining cell populations using the bimod assay. Based on the HPCA reference dataset [28], cell type annotation was performed using the SingleR package [29].

## Results

### Demographic and clinical characteristics

A total of 22 individuals diagnosed with sepsis were included as participants in this study, comprising 14 males and 8 females. Additionally, there were 10 individuals in the normal control group, consisting of 5 males and 5 females. Among the sepsis group, 12 individuals survived, while 10 individuals died within a 28-day period. Conversely, all 10 individuals in the control group survived without any fatalities. The statistical analysis employed the unpaired t-test to assess the Age, Alanine transaminase (ALT), Aspartate aminotransferase (AST), Direct bilirubin (DBIL), Total bilirubin (TBIL), Creatinine, Urea, Leukocyte count, Neutrophil count,



**Fig. 1** Screening for sepsis-differential genes and lysosome-related genes. **A** Scatter plot with an R value of 0.985 indicates a strong positive correlation and data stability between the two variables. **B-C** Box plot, and density distribution plot demonstrated homogeneity and comparability of the two data sets. **D** PCA successfully distinguishes the normal control group from the sepsis group without any outlier samples. **E** The volcano diagram visually represents 221 upregulated genes in the sepsis group, indicated in red, while blue color represents 1107 downregulated genes in the sepsis group. **F** Venn diagram showed the intersection of 1328 sepsis-differential genes with 878 lysosome-related genes, and a total of 76 intersecting genes were obtained



Monocyte count, and Lymphocyte count in both groups. The mean  $\pm$  standard deviation was used to express the results, which are presented in Table 1.

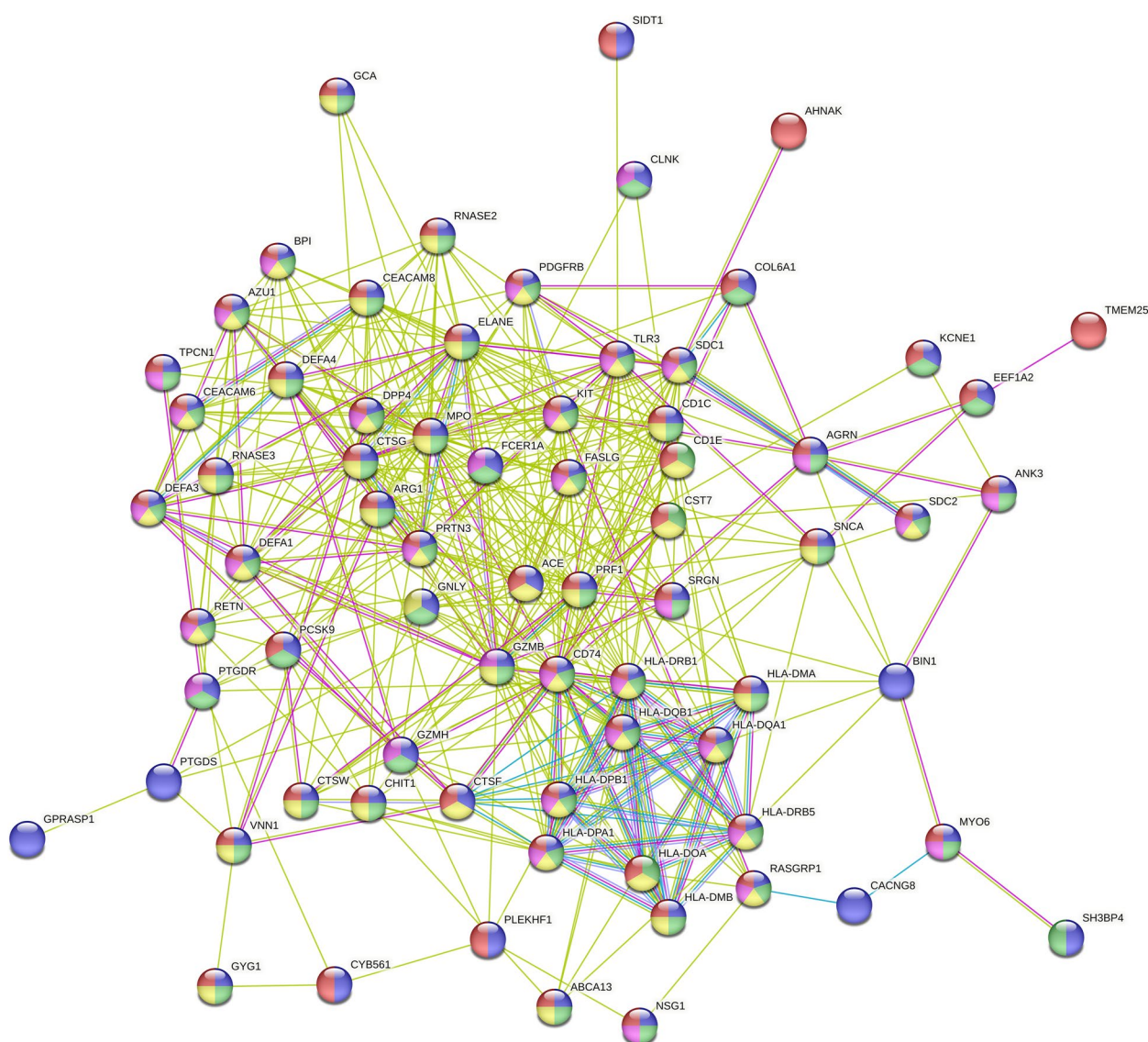
### Screening for differential expression genes

Two representative groups were randomly chosen to generate scatterplots for the purpose of quality control in RNA sequencing data. The obtained Pearson product-moment correlation coefficient ( $R=0.985$ ) indicated a highly robust positive correlation between the two variables, thereby confirming the stability of the data (Fig. 1A). Furthermore, the box plot and density distribution plot demonstrated

the homogeneity and comparability of the two datasets (Fig. 1B-C). Additionally, the PCA successfully distinguished the normal group from the sepsis group, with no presence of outlier samples (Fig. 1D). A total of 1328 differential genes were obtained through screening for genes with  $|FC| \geq 4$  and  $FDR < 0.05$ . Among these, 221 genes were upregulated in the sepsis group, while 1107 genes were downregulated in the sepsis group (Fig. 1E).

### Screening for lysosome-related genes

A total of 878 genes associated with lysosomes were obtained from the Gene Ontology database for further



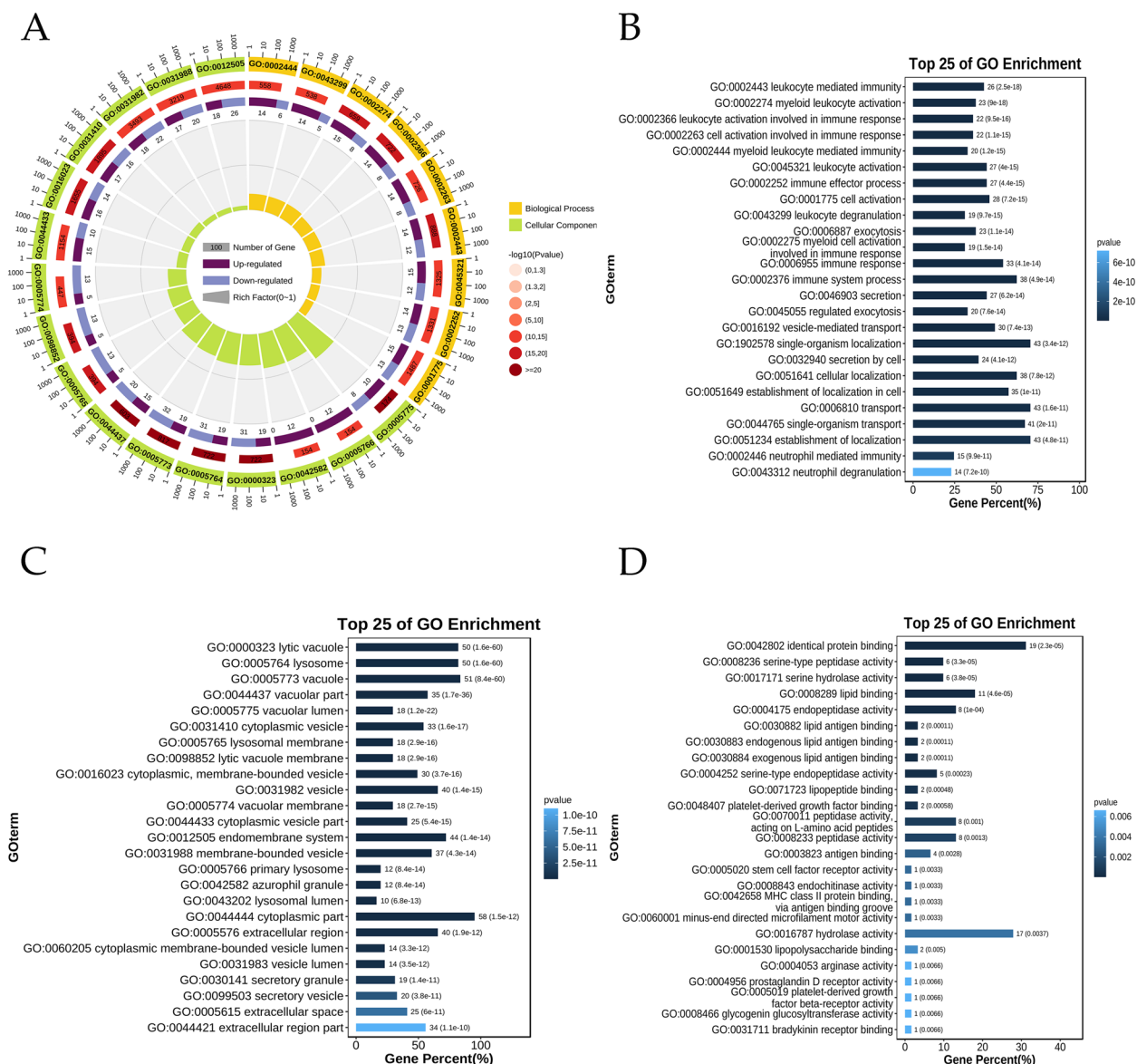
**Fig. 2** PPI analysis. The PPI network comprises 75 nodes and 450 connections, the dots in the network were color-coded to represent different biological processes, with blue indicating Cellular process, green indicating Response to stimulus, yellow indicating Immune system process, purple indicating Signal transduction, and red indicating Lysosome. The color of the lines connecting the nodes signified the type of interaction evidence, with green denoting Text mining, red denoting Experiments, and blue denoting Databases

analysis. 878 lysosome-related genes were found to intersect with 1328 genes that exhibit differential expression in sepsis, resulting in the identification of 76 lysosome-related genes that were closely associated with sepsis (Fig. 1F).

**PPI analysis**

The PPI network comprises 75 nodes and 450 connections, with GNLY, GZMB, PRF1, TLR3, KIT, and other proteins occupying central positions within the network.

These proteins have the potential to serve as core targets that could impact the prognosis of sepsis. The dots in the network were color-coded to represent different biological processes, with blue indicating Cellular process, green indicating Response to stimulus, yellow indicating Immune system process, purple indicating Signal transduction, and red indicating Lysosome. The color of the lines connecting the nodes signified the type of interaction evidence, with green denoting Text mining, red denoting Experiments, and blue denoting Databases (Fig. 2).



**Fig. 3** GO analysis. **A** The analysis of GO revealed that among the top 25 items, 9 were related to BPs, 16 were related to CCs, and none were related to MFs. **B** The intersecting genes primarily participated in various BPs, such as leukocyte mediated immunity, cell activation involved in immune response, immune effector process, leukocyte degranulation, and vesicle-mediated transport. **C** The CCs associated with the intersection genes mainly consisted of lytic vacuole, lysosome, cytoplasmic vesicle, lysosomal membrane, and lysosomal lumen. **D** The BPs with the intersection genes mainly involved in identical protein binding, serine-type peptidase activity, serine hydrolase activity, lipid binding, endopeptidase activity

### GO analysis

The analysis of GO revealed that among the top 25 items, 9 were related to BPs, 16 were related to CCs, and none were related to MFs (Fig. 3A). Additionally, the taxonomic enrichment analysis indicated that the intersecting genes primarily participated in various BPs, such as leukocyte mediated immunity, cell activation involved in immune response, immune effector process, leukocyte degranulation, and vesicle-mediated transport (Fig. 3B). The CCs associated with the intersection genes mainly consisted of lytic vacuole, lysosome, cytoplasmic vesicle, lysosomal membrane, and lysosomal lumen (Fig. 3C). Furthermore, the MFs with the intersection genes mainly involved in identical protein binding, serine-type peptidase activity, serine hydrolase activity, lipid binding, endopeptidase activity (Fig. 3D).

### KEGG analysis

The analysis conducted by KEGG identified the 25 most noteworthy items, comprising 1 Metabolism, 4 Environmental Information Processing, 2 Cellular Processes, 5 Organismal Systems, and 13 Human Diseases (Fig. 4A). Among the pathways with the highest significance for the intersection gene were Graft-versus-host disease, Type I diabetes mellitus, Apoptosis, Allograft rejection, Autoimmune thyroid disease, and Lysosome (Fig. 4B).

### Survival analysis

Based on a survival analysis of the GSE65682 dataset in the GEO database, it was observed that patients with high expression levels of GNLY, GZMB, PRF1, and RASGRP1

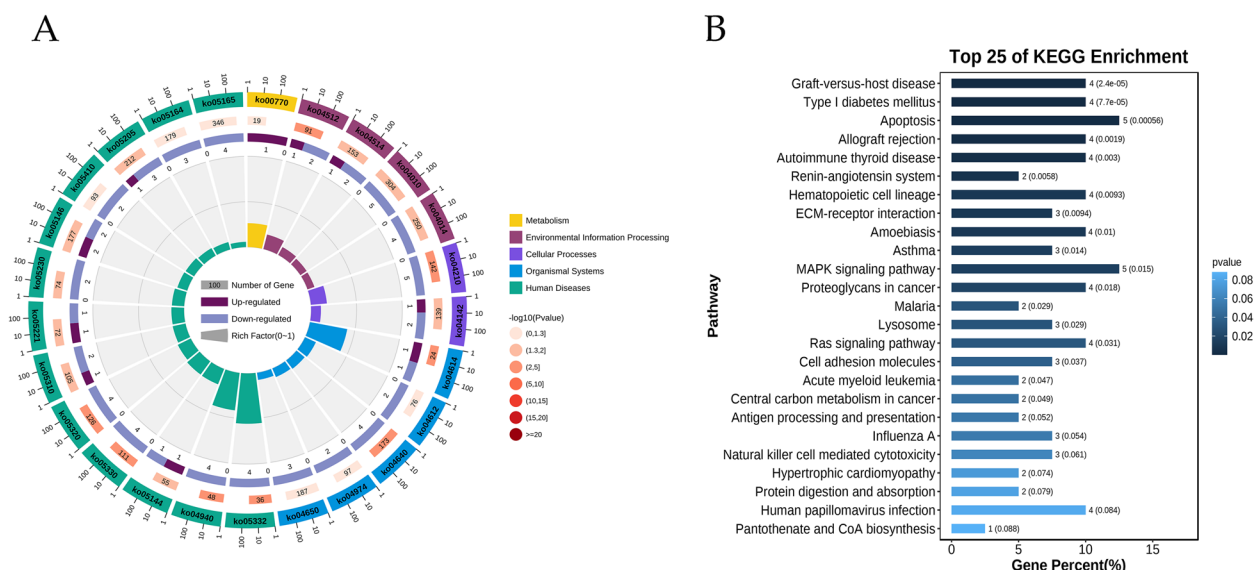
exhibited a higher 28-day survival rate compared to those with low expression levels ( $P < 0.05$ ). This finding suggests a positive correlation between these genes and the prognosis of sepsis patients, with their high expression levels potentially serving as a novel research focus for sepsis. Consequently, these findings provide valuable insights for the development of lysosome-targeting treatments for sepsis (Fig. 5A-D).

### Meta-analysis

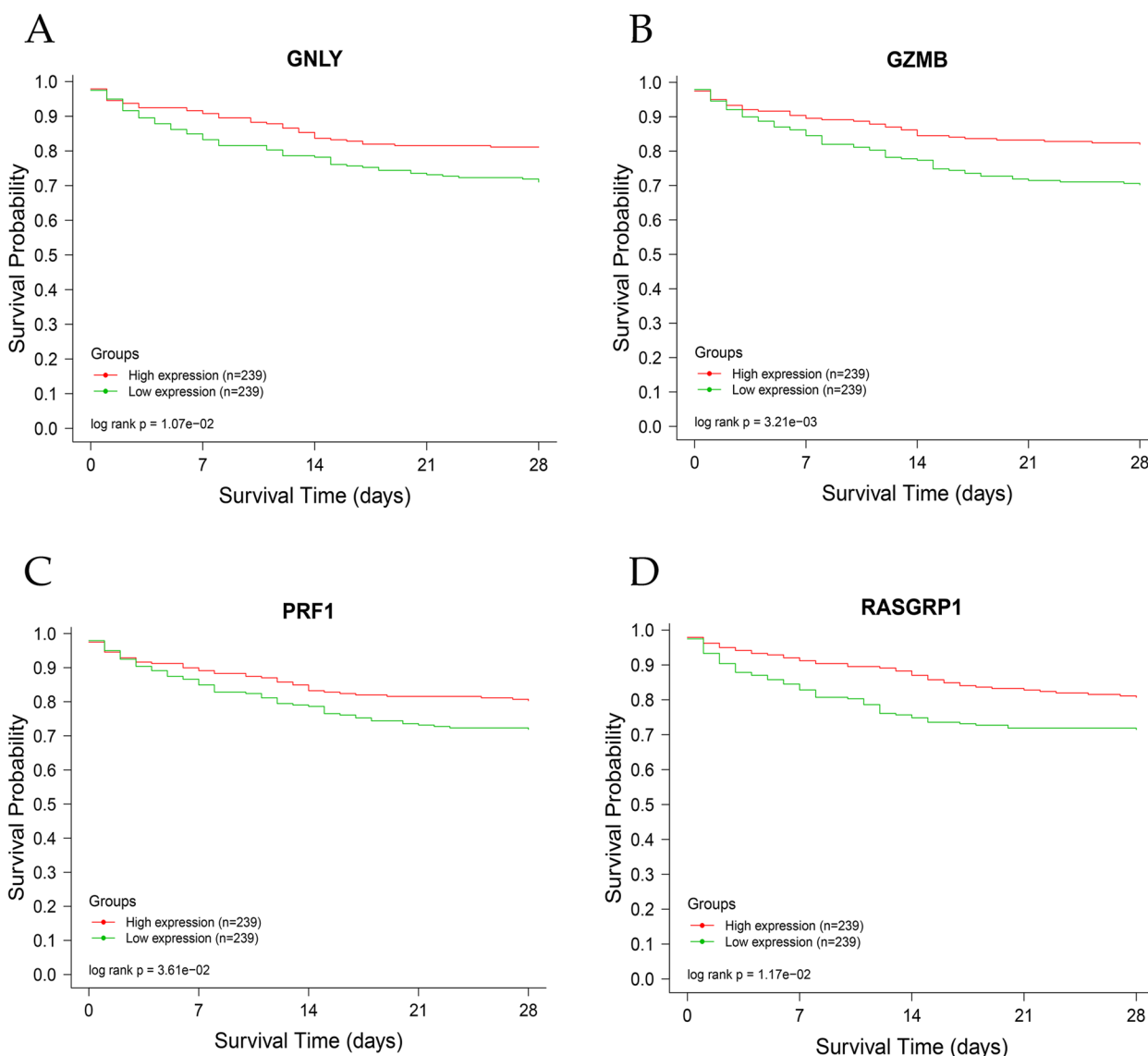
At the transcriptome level, a meta-analysis was conducted on core genes using RNA sequencing datasets GSE28750, GSE54514, GSE69528, and GSE95233. The analysis revealed that GNLY, GZMB, PRF1, and RASGRP1 exhibited high expression levels in the normal control group, while their expression was low in the sepsis group (Fig. 6A-D). Further information regarding the GEO dataset of the included studies can be found in Table 2. Based on the peripheral blood RNA sequencing data of the normal control group ( $n=10$ ) and the sepsis group ( $n=22$ ), the expression of core genes was plotted pod plots, and the results showed that GNLY, GZMB, PRF1 and RASGRP1 were highly expressed in the normal control group and low in the sepsis group, and the difference was statistically significant (Fig. 7).

### Single-cell RNA sequencing

The distribution of high-quality cells for quality control of each sample ranged from 4050 to 10,191. After



**Fig. 4** KEGG analysis. **A** The analysis conducted by KEGG identified the 25 most noteworthy items, comprising 1 Metabolism, 4 Environmental Information Processing, 2 Cellular Processes, 5 Organismal Systems, and 13 Human Diseases. **B** Among the pathways with the highest significance for the intersection gene were Graft-versus-host disease, Type I diabetes mellitus, Apoptosis, Allograft rejection, Autoimmune thyroid disease, and Lysosome



**Fig. 5** Survival analysis. **A-D** Survival analysis based on the GSE65682 dataset observed that patients with high expression levels of GNLY, GZMB, PRF1, and RASGRP1 exhibited a higher 28-day survival rate compared to those with low expression levels ( $P < 0.05$ )

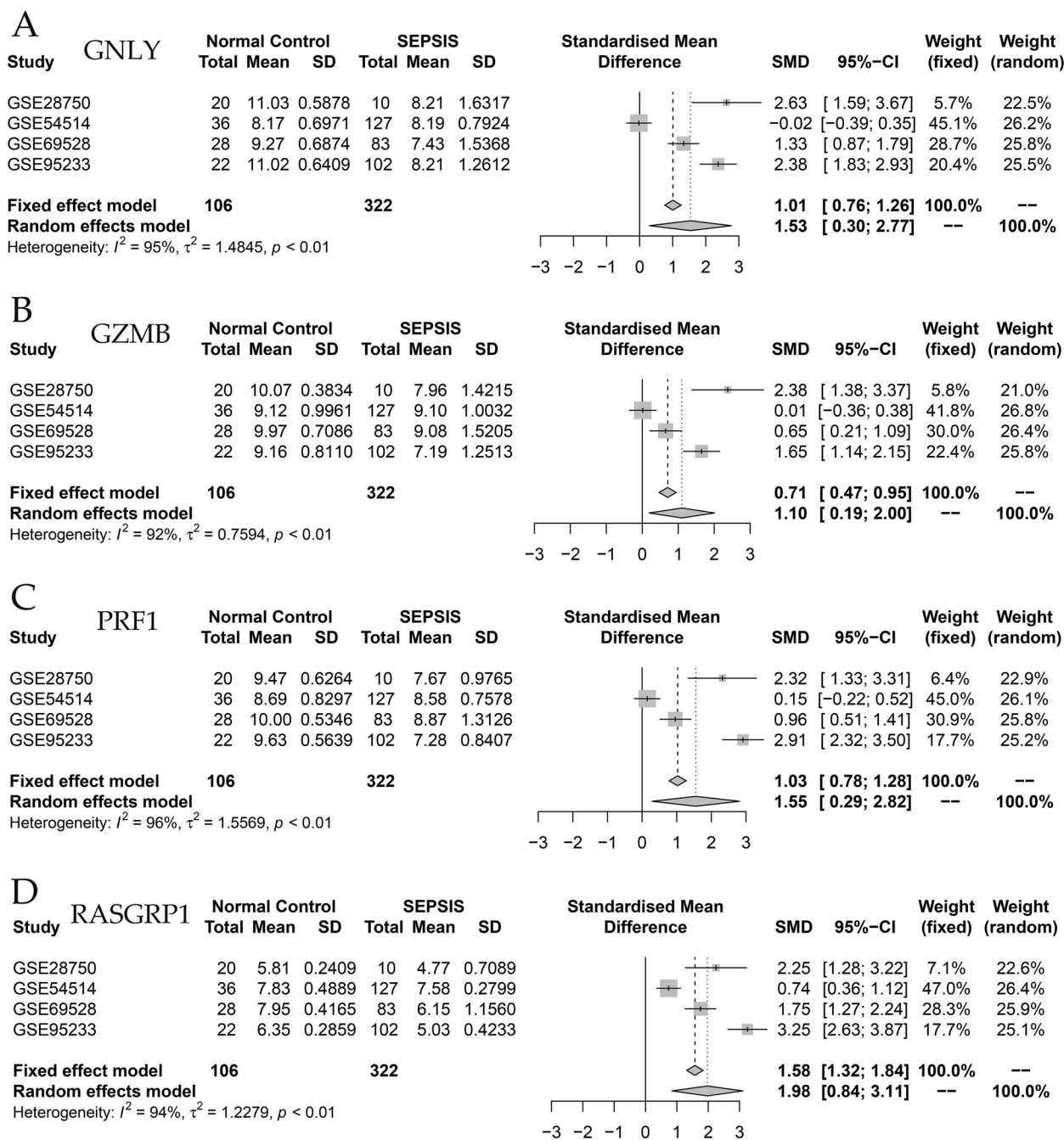
excluding two-cells, multicellular and apoptotic cells, the final number of cells obtained ranged from 3108 to 8509. The average UMI number in each cell ranged from 519 to 8529, and the average gene number in each cell ranged from 343 to 2337. Following dimensionality reduction, the cells were classified into 9 groups (Fig. 8A), with B cells, NK cells, T cells, Platelets, and Macrophages serving as reference cell types. Among the identified cell lines, T cells were represented by 1, 2, 6, and 8, Macrophages by 3 and 5, NK cells by 4, B cells by 7, and Platelets by 9 (Fig. 8B). Analysis of single-cell RNA sequencing data revealed that GNLY, GZMB, and PRF1 were predominantly expressed in cell lines 2 and

4, while RASGRP1 was primarily localized in cell lines 1, 2, 4, 6, and 8 (Fig. 8C). Consequently, GNLY, GZMB, PRF1, and RASGRP1 exhibited high expression levels in T cells and NK cells (Fig. 8D-E).

**Discussion**

Deng M. et al. conducted a study which demonstrated that hepatocytes release high mobility group box 1 (HMGB1), which then binds to lipopolysaccharide (LPS) and is internalized into the lysosomes of macrophages and endothelial cells through the receptor for advanced glycation end-products (RAGE). Subsequently, HMGB1 causes permeabilization of the phospholipid





**Fig. 6** Meta-analysis. **A-D** Meta-analysis based on the GSE28750, GSE54514, GSE69528, GSE95233 datasets revealed that GNLy, GZMB, PRF1, and RASGRP1 exhibited high expression levels in the normal control group, while their expression was low in the sepsis group

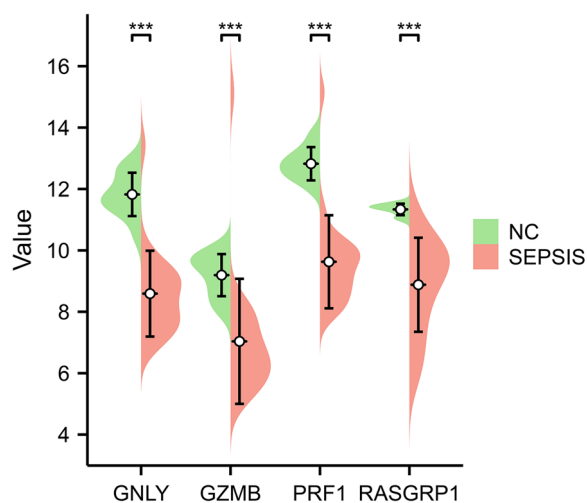
bilayer within the acidic environment of the lysosomes. This leads to the leakage of LPS into the cytoplasm and the activation of caspase-11. Caspase-11, being an LPS receptor, plays a role in caspase-11-dependent pyroptosis, which ultimately contributes to the lethality of endotoxemia [30, 31]. Hence, the close association between lysosomes and LPS-mediated sepsis, which can impact

the pathophysiological progression of sepsis, is evident. Numerous chemicals exhibit specific affinity for lysosomal proteins, thereby governing cellular behavior [32]. Consequently, this investigation employed RNA sequencing and bioinformatics techniques to identify four lysosome-related genes that significantly influence the prognosis of sepsis patients. These findings offer valuable

**Table 2** The GSE datasets for Survival analysis and Meta-analysis

GSE datasets	Organism	Platform	Number of samples
GSE65682	Whole Blood of Human	GPL13667	802
GSE28750	Whole Blood of Human	GPL570	41
GSE54514	Whole Blood of Human	GPL6947	163
GSE69528	Whole Blood of Human	GPL10558	138
GSE95233	Whole Blood of Human	GPL570	124

The GSE65682 dataset of the GEO database was used for survival analysis, GSE28750, GSE54514, GSE69528, and GSE95233 were used for meta-analysis. Statistics of Organism, Platform, and Number of samples for each dataset were also used



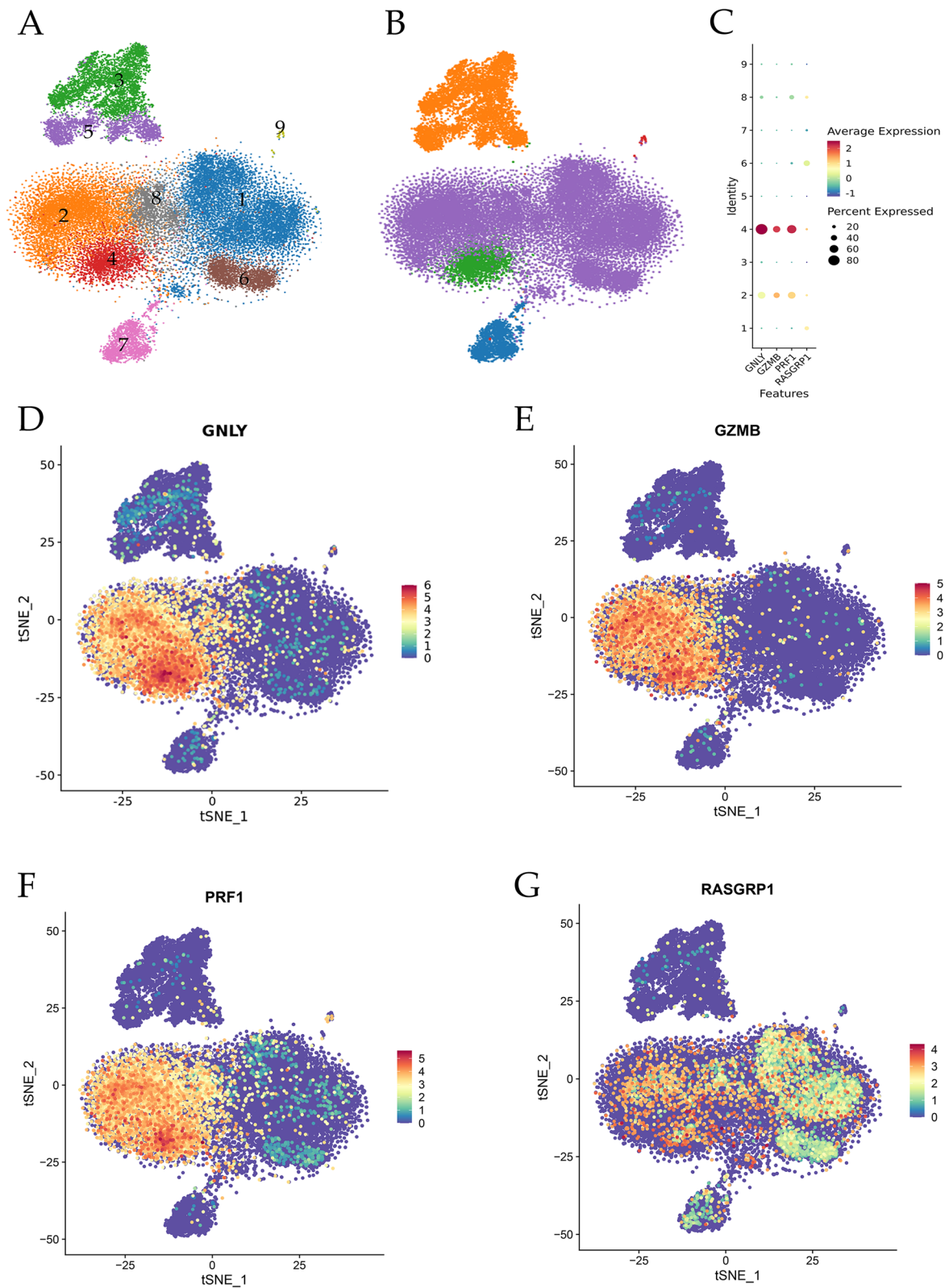
**Fig. 7** Core gene expression analysis. The pod plot illustrated the expression patterns of core genes as determined through peripheral blood RNA sequencing data obtained from both the normal control group ( $n=10$ ) and the sepsis group ( $n=22$ ). Notably, the analysis revealed that GNLY, GZMB, PRF1, and RASGRP1 exhibited significantly higher expression levels in the normal control group compared to the sepsis group ( $P<0.05$ ). It is worth mentioning that the significance levels were denoted as  $*P<0.05$ ,  $**P<0.01$ , and  $***P<0.001$

insights for the development of targeted therapies aimed at lysosomal intervention.

Granulysin (GNLY) is a cytotoxic granule that is co-secreted with granzyme and perforin from the granules of human cytotoxic T lymphocytes (CTLs) and natural killer (NK) cells. It has been documented to possess diverse antimicrobial properties [33, 34]. GNLY is predominantly expressed in human CTLs and NK cells and has been implicated in a plethora of pathological conditions, including infections, cancer, transplantation, and dermatological disorders [35]. Granzyme B (GZMB) is an integral constituent of cytolytic granules found within natural killer (NK) cells and serves as a crucial cytotoxic molecule employed by T cells to eliminate cells infected by pathogens or transformed tumor cells [36, 37]. CD8+ T cells have the ability to induce apoptosis in target cells through the release of GZMB, which, in turn, may lead to tissue damage and restructuring [38]. Emerging evidence indicates its involvement in the lysosome-mediated demise of the cytotoxic lymphocyte itself [39]. Perforin (PRF1) is an essential pore-forming protein involved in lymphocyte cytotoxicity [40]. Mutations in the PRF1 gene have been identified as a causative factor for the development of Hemophagocytic Lymphohistiocytosis (HLH) [41]. HLH is a severe immunodeficiency and multi-organ disorder that can result in a potentially lethal hyperinflammatory state, characterized by fever, hepatosplenomegaly, and distinctive laboratory abnormalities [42]. RAS guanyl-releasing protein 1 (RASGRP1) plays a crucial role as a guanine nucleotide exchange factor and a vital regulator of T cell receptor signaling in the immune system [43]. In murine models, the absence of RASGRP1 leads to impaired development of T lymphocytes [44]. In humans, deficiencies in RASGRP1 can result in a primary immunodeficiency (PID) syndrome, characterized by lymphopenia in CD4+ T cells and the development of Epstein-Barr virus (EBV)-associated B cell lymphoma [45]. By utilizing RNA sequencing and conducting bioinformatics analysis, our study revealed that the genes GNLY, GZMB, PRF1, and RASGRP1 exhibited predominant expression within the lysosomes of T cells and NK cells. These genes were found to actively participate in crucial biological processes, including the

(See figure on next page.)

**Fig. 8** Single-cell RNA sequencing. **A** The total sequencing diagram showed that, following dimensionality reduction, the cells were categorized into nine distinct groups. **B** Among the identified cell lines, T cells were represented by 1, 2, 6, and 8, Macrophages by 3 and 5, NK cells by 4, B cells by 7, and Platelets by 9. **C** In the dotplot, the gene name was plotted on the abscissa, while the cell group number was plotted on the ordinate. The size of the dot corresponded to the proportion of cells expressing the gene within that particular cell group. Additionally, a color gradient ranging from blue to red indicated varying levels of gene expression, with red indicating higher expression. The findings of this study indicate that GNLY, GZMB, and PRF1 were predominantly localized within cell lines 2 and 4, while RASGRP1 exhibited predominant localization within cell lines 1, 2, 4, 6, and 8. **D-G**, A dot in a feature plot represented an individual cell, with the color gradient ranging from blue to red, signifying higher gene expression levels within the cell. These results further demonstrate that GNLY, GZMB, PRF1, and RASGRP1 exhibit high expression levels, specifically within T cells and NK cells



**Fig. 8** (See legend on previous page.)

immune system process, response to stimulus, cellular process, and signal transduction. Notably, their expression levels were significantly higher in the normal control group compared to the sepsis group. Furthermore, a positive correlation was observed between the expression of these genes and the prognosis of sepsis patients. Previous research has also indicated a strong association between these four genes and the immune response mediated by lymphocytes. Hence, it is postulated that these four genes could serve as innovative targets for precise lysosomal therapy, which plays a significant role in lysosomal-mediated inflammation and immune responses. Consequently, they may modulate the expression of lysosomal proteins, influencing the pathophysiological progression of sepsis and ultimately impacting the diverse clinical outcomes observed in patients.

In conclusion, the genes *GNLY*, *GZMB*, *PRF1*, and *RASGRP1* exhibited significant upregulation in the normal control group and downregulation in the sepsis group, thereby displaying a positive correlation with the prognosis of sepsis patients. This observation implies that their heightened expression potentially contributes to the survival of sepsis patients. Additionally, given their association with lysosomal functions, these four genes may serve as promising targets for lysosomal therapy, thereby offering a novel avenue for the clinical management of sepsis.

## Conclusion

*GNLY*, *GZMB*, *PRF1*, and *RASGRP1* are lysosome-related genes that are closely related to the prognosis of sepsis, and may be used as new research targets for sepsis and provide direction for lysosome-targeted therapy.

## Authors' contributions

Z.W. designed the study. G.C. and W.Z. collected the blood samples. C.W., M.C. and Y.H. analyzed the data. G.C. wrote the manuscript.

## Funding

This study was supported by the Southwest Medical University, grant number: [2018] 6 (2017-ZRQN-009).

## Availability of data and materials

The datasets generated during the current study are available in the China National GeneBank DataBase (CNCBdb, <https://db.cngb.org/>) repository, [ACCESSION NUMBER: CNP0002611].

The GEO database analyzed during this study are available in the GEO (<https://www.ncbi.nlm.nih.gov/geo/>) repository, [ACCESSION NUMBERS: GSE65682, GSE28750, GSE54514, GSE69528, GSE95233].

## Declarations

### Ethics approval and consent to participate

The Affiliated Hospital of Southwest Medical University Ethics Committee has approved this study protocol, ethics number: ky2018029. An informed

consent form was signed by all participants included in the study. The registered clinical trial number is ChiCTR1900021261.

### Consent for publication

Not applicable.

### Competing interests

The authors declare no competing interests.

Received: 15 July 2023 Accepted: 20 November 2023

Published online: 06 December 2023

## References

- Miljković Đ, Stanisavljević S, Jensen IJ, Griffith TS, Badovinac VP. Sepsis and multiple sclerosis: Causative links and outcomes. *Immunol Lett*. 2021;238:40–6.
- Wen X, Xie B, Yuan S, Zhang J. The "Self-Sacrifice" of Immune Cells in Sepsis. *Front Immunol*. 2022;13:833479.
- Leong K, Gaglani B, Khanna AK, McCurdy MT. Novel Diagnostics and Therapeutics in Sepsis. *Biomedicines*. 2021;9(3):311.
- López-Mestanza C, Andaluz-Ojeda D, Gómez-López JR, Bermejo-Martín JF. Clinical factors influencing mortality risk in hospital-acquired sepsis. *J Hosp Infect*. 2018;8(2):194–201.
- Zhang Z, Yue P, Lu T, Wang Y, Wei Y, Wei X. Role of lysosomes in physiological activities, diseases, and therapy. *J Hematol Oncol*. 2021;14(1):79.
- Reynolds T. Cholesteryl ester storage disease: a rare and possibly treatable cause of premature vascular disease and cirrhosis. *J Clin Pathol*. 2013;66(11):918–23.
- Menzies FM, Fleming A, Rubinsztein DC. Compromised autophagy and neurodegenerative diseases. *Nat Rev Neurosci*. 2015;16(6):345–57.
- Maiuri MC, Kroemer G. Therapeutic modulation of autophagy: which disease comes first? *Cell Death Differ*. 2019;26(4):680–9.
- Gutierrez MG, Master SS, Singh SB, Taylor GA, Colombo MI, Deretic V. Autophagy is a defense mechanism inhibiting BCG and Mycobacterium tuberculosis survival in infected macrophages. *Cell*. 2004;119(6):753–66.
- Nakagawa I, Amano A, Mizushima N, Yamamoto A, Yamaguchi H, Kamimoto T, Nara A, Funao J, Nakata M, Tsuda K, et al. Autophagy defends cells against invading group A Streptococcus. *Science*. 2004;306(5698):1037–40.
- Balka KR, De Nardo D. Understanding early TLR signaling through the Myddosome. *J Leukoc Biol*. 2019;105(2):339–51.
- Saric A, Hipolito VE, Kay JG, Canton J, Antonescu CN, Botelho RJ. mTOR controls lysosome tubulation and antigen presentation in macrophages and dendritic cells. *Mol Biol Cell*. 2016;27(2):321–33.
- Mantegazza AR, Zajac AL, Twelvetrees A, Holzbaur EL, Amigorena S, Marks MS. TLR-dependent phagosome tubulation in dendritic cells promotes phagosome cross-talk to optimize MHC-II antigen presentation. *Proc Natl Acad Sci USA*. 2014;111(43):15508–13.
- Guo FM, Qiu HB. Definition and diagnosis of sepsis 3.0. *Zhonghua nei ke za zhi*. 2016;55(6):420–2.
- Li R, Li Y, Kristiansen K, Wang J. SOAP: short oligonucleotide alignment program. *Bioinformatics*. 2008;24(5):713–4.
- Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. *Nat Methods*. 2015;12(4):357–60.
- Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 2012;9(4):357–9.
- Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics*. 2011;12:323.
- Ge X. iDEP Web Application for RNA-Seq Data Analysis. *Methods Mol Biol (Clifton, NJ)*. 2021;2284:417–43.
- Aleksander SA, Balhoff J, Carbon S, Cherry JM, Drabkin HJ, Ebert D, Feuerhann M, Gaudet P, Harris NL, Hill DP, et al. The Gene Ontology knowledgebase in 2023. *Genetics*. 2023;224(1):iyad031.
- von Mering C, Huynen M, Jaeggi D, Schmidt S, Bork P, Snel B. STRING: a database of predicted functional associations between proteins. *Nucleic Acids Res*. 2003;31(1):258–61.



22. Ogata H, Goto S, Sato K, Fujibuchi W, Bono H, Kanehisa M. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* 1999;27(1):29–34.
23. Scicluna BP, Klein Klouwenberg PM, van Vught LA, Wiewel MA, Ong DS, Zwinderman AH, Franitza M, Toliat MR, Nürnberg P, Hoogendijk AJ, et al. A molecular biomarker to diagnose community-acquired pneumonia on intensive care unit admission. *Am J Respir Crit Care Med.* 2015;192(7):826–35.
24. Sutherland A, Thomas M, Brandon RA, Brandon RB, Lipman J, Tang B, McLean A, Pascoe R, Price G, Nguyen T, et al. Development and validation of a novel molecular biomarker diagnostic test for the early detection of sepsis. *Crit Care.* 2011;15(3):R149.
25. Parnell GP, Tang BM, Nalos M, Armstrong NJ, Huang SJ, Booth DR, McLean AS. Identifying key regulatory genes in the whole blood of septic patients to monitor underlying immune dysfunctions. *Shock.* 2013;40(3):166–74.
26. Pankla R, Buddhisa S, Berry M, Blankenship DM, Bancroft GJ, Banchereau J, Lertmemongkolchai G, Chaussabel D. Genomic transcriptional profiling identifies a candidate blood biomarker signature for the diagnosis of septicemic melioidosis. *Genome Biol.* 2009;10(11):R127.
27. Tabone O, Mommert M, Jourdan C, Cerrato E, Legrand M, Lepape A, Allaouchiche B, Rimmelé T, Pachot A, Monneret G, et al. Endogenous Retroviruses Transcriptional Modulation After Severe Infection. *Trauma and Burn Frontiers in immunology.* 2018;9:3091.
28. Mabbott NA, Baillie JK, Brown H, Freeman TC, Hume DA. An expression atlas of human primary cells: inference of gene function from coexpression networks. *BMC Genomics.* 2013;14:632.
29. Aran D, Looney AP, Liu L, Wu E, Fong V, Hsu A, Chak S, Naikawadi RP, Wolters PJ, Abate AR, et al. Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. *Nat Immunol.* 2019;20(2):163–72.
30. Deng M, Tang Y, Li W, Wang X, Zhang R, Zhang X, Zhao X, Liu J, Tang C, Liu Z, et al. The Endotoxin Delivery Protein HMGB1 Mediates Caspase-11-Dependent Lethality in Sepsis. *Immunity.* 2018;49(4):740–753.e747.
31. Li Y, Suo L, Fu Z, Li G, Zhang J. Pivotal role of endothelial cell autophagy in sepsis. *Life Sci.* 2021;276:119413.
32. Pance K, Gramespacher JA, Byrnes JR, Salangsang F, Serrano JC, Cotton AD, Steri V, Wells JA. Modular cytokine receptor-targeting chimeras for targeted degradation of cell surface and extracellular proteins. *Nat Biotechnol.* 2023;41(2):273–81.
33. Ermis E, Celik SK, Solak N, Genc GC, Dursun A. The role of GNLY gene polymorphisms in psoriasis pathogenesis. *An Bras Dermatol.* 2019;94(2):198–203.
34. Khalid HN, Elghobashy YAE, Elsayed AN. GNLY gene polymorphism: A potential role in understanding psoriasis pathogenesis. *J Cosmet Dermatol.* 2022;21(10):4805–9.
35. Lettau M, Dietz M, Dohmen K, Leippe M, Kabelitz D, Janssen O. Granulysin species segregate to different lysosome-related effector vesicles (LREV) and get mobilized by either classical or non-classical degranulation. *Mol Immunol.* 2019;107:44–53.
36. Oboshi W, Watanabe T, Hayashi K, Nakamura T, Yukimasa N. QPY/RAH haplotypes of the GZMB gene are associated with natural killer cell cytotoxicity. *Immunogenetics.* 2018;70(1):29–36.
37. Du W, Leigh ND, Bian G, Alqassim E, O'Neill RE, Mei L, Qiu J, Liu H, McCarthy PL, Cao X. Granzyme B Contributes to the Optimal Graft-Versus-Tumor Effect Mediated by Conventional CD4+ T Cells. *J Immunol Res Ther.* 2016;1(1):22–8.
38. Kim WD, Chi HS, Choe KH, Kim WS, Hogg JC, Sin DD. The Role of Granzyme B Containing Cells in the Progression of Chronic Obstructive Pulmonary Disease. *Tuberc Respir Dis.* 2020;83(Suppl 1):S25–s33.
39. Mangan MS, Bird CH, Kaiserman D, Matthews AY, Hitchen C, Steer DL, Thompson PE, Bird PI. A Novel Serpin Regulatory Mechanism: SerpinB9 Is reversibly inhibited by vicinal disulfide bond formation in the reactive center loop. *J Biol Chem.* 2016;291(7):3626–38.
40. Tesi B, Chiang SC, El-Ghoneimy D, Hussein AA, Langenskiöld C, Wali R, Fadoo Z, Silva JP, Lecumberri R, Unal S, et al. Spectrum of Atypical clinical presentations in patients with Biallelic PRF1 missense mutations. *Pediatr Blood Cancer.* 2015;62(12):2094–100.
41. Zhu GH, Zhang LP, Li ZG, Wei A, Yang Y, Tian Y, Ma HH, Wang D, Zhao XX, Zhao YZ, et al. Associations between PRF1 Ala91Val polymorphism and risk of hemophagocytic lymphohistiocytosis: a meta-analysis based on 1366 subjects. *World J Pediatr.* 2020;16(6):598–606.
42. Caldito NG, Lorenzo J, Wang CX. Familial CNS-Isolated Hemophagocytic Lymphohistiocytosis Due to a Novel PRF1 Mutation Triggered by SARS-CoV2. *Ann Indian Acad Neurol.* 2022;25(6):1170–3.
43. Salzer E, Cagdas D, Hons M, Mace EM, Garncarz W, Petronczki Ö, Platzer R, Pfajfer L, Bilic I, Ban SA, et al. RASGRP1 deficiency causes immunodeficiency with impaired cytoskeletal dynamics. *Nat Immunol.* 2016;17(12):1352–60.
44. Baars MJD, Douma T, Simeonov DR, Myers DR, Kulhanek K, Banerjee S, Zwakenberg S, Baltissen MP, Amini M, de Roock S, et al. Dysregulated RASGRP1 expression through RUNX1 mediated transcription promotes autoimmunity. *Eur J Immunol.* 2021;51(2):471–82.
45. Somekh I, Marquardt B, Liu Y, Rohlf M, Hollizeck S, Karakucuk M, Unal E, Yilmaz E, Patrioglu T, Cansever M, et al. Novel Mutations in RASGRP1 are Associated with Immunodeficiency, Immune Dysregulation, and EBV-Induced Lymphoma. *J Clin Immunol.* 2018;38(6):699–710.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

