Original Article

Emergency and Critical Care Medicine

OPEN

Transcriptome analysis for identifying hub genes and prognosis biomarkers of mRNA/lncRNA in septic shock

Chao Gong[a], Wenzhong Zhang[b], Xin Lu[a], Shiyuan Yu[a], Zengzheng Ge[a], Mubing Qin[a], Huadong Zhu[a], Yi Li[a,*]

Abstract

Background: Septic shock is a life-threatening disease with high mortality rates, and the relevant hub genes and biomarkers are poorly understood. We aimed to identify hub genes and prognostic biomarkers of mRNAs/lncRNAs in septic shock to rapidly and accurately diagnose infection, identify patients at a high risk of developing septic shock, and predict prognosis.

Methods: Gene expression profiles of 279 patients with septic shock and 100 healthy controls were analyzed using bioinformatics methods. We screened for differentially expressed genes (DEGs), identified hub genes, and investigated the correlations between mRNA/lncRNA expression and disease severity/prognosis. Protein level validation was performed using blood proteomic data from an independent cohort study.

Results: The protein–protein interaction network constructed using upregulated DEGs contained 102 nodes and 222 edges, with LTF, MMP8, MMP9, CEACAM8, CTSG, LCN2, and PRTN3 identified as hub genes. There was a possible association between LCN2 mRNA upregulation and increased severity of septic shock (odds ratio: 1.518; 95% confidence interval: 0.999–2.305; $P = 0.050$), approaching statistical significance, and BCL2A1 mRNA upregulation correlated with higher mortality risk (odds ratio: 1.178; 95% confidence interval: 1.035–1.341; $P = 0.013$). No significant prognostic correlation was observed for lncRNAs. The validation cohort confirmed significant upregulation of MMP9, CTSG, LCN2, LTF, and MMP8 proteins in patients with septic shock, with MMP9, LCN2, CTSG, and LTF exhibiting strong diagnostic performance (area under the curve >0.8).

Conclusion: Seven hub genes related to septic shock were identified, including MMP9, LCN2, CTSG, and LTF, which could potentially function as biotargets and biomarkers for the diagnosis and prognostic prediction of septic shock.

Key words: Biomarker, lncRNA, mRNA, Septic shock, Transcriptome

Introduction

Sepsis is a life-threatening organ dysfunction resulting from dysregulated host responses to infection[1,2] and has become the main cause of death in intensive care unit patients, with millions of deaths worldwide every year.[3] Septic shock is a subset of sepsis in which the underlying circulatory, cellular, and metabolic abnormalities are sufficiently profound to substantially increase the risk of mortality.[1] Patients with septic shock can be identified with a clinical construct of sepsis with persisting hypotension requiring vasopressors to maintain mean arterial pressure (MAP) ≥ 65 mmHg and having a serum lactate level >2 mmol/L (18 mg/dL) despite adequate volume resuscitation.[1] Sepsis and septic shock are major healthcare problems affecting millions of people worldwide each year and killing over 26.7% of those affected.[4]

The pathogenesis of sepsis is complicated and currently unclear; therefore, the treatment of sepsis is limited.[5] These mechanisms include systemic inflammation, coagulation, fibrinolytic disorders, and excessive production of reactive oxygen and nitrogen.[5] Thus far, no specific therapeutic drugs have been approved for the treatment of sepsis, and treatment options are limited.[6] One important reason for this discrepancy is that the definitions of sepsis and septic shock cover a heterogeneous population of patients. The etiological factors of sepsis and septic shock are so varied that it is difficult to find a common treatment for these conditions.[7] Classification of patients with sepsis or septic shock is one of the key areas of research on sepsis, although biomarkers have been the subject of intensive research for decades. Biomarkers identified by new technologies, such as high-throughput technologies, have been used to better identify subsets of patients with sepsis, identify patients at high risk of developing sepsis, and provide the possibility for rapid and accurate diagnosis of infection.[8]

Microarray technology and bioinformatics analysis have been widely adopted, playing a key role in supporting life sciences and bringing unique opportunities and challenges to human disease research.[9] Based on bioinformatics analyses, many meaningful

CG and WZ contributed equally to this article.

The datasets generated during and/or analyzed during the current study are available in the Gene Expression Omnibus (GEO) database in the National Center for Biotechnology Information at https://pubmed.ncbi.nlm.nih.gov/, Series Accession: GSE95233, GSE33118, GSE57065, GSE26440, and GSE26378.

[a]Emergency Department, The State Key Laboratory for Complex, Severe and Rare Diseases, Peking Union Medical College Hospital, Chinese Academy of Medical Science and Peking Union Medical College, Beijing, China, [b]Department of Emergency, First Hospital of Handan, Handan, Hebei, China.

** Corresponding author. Address: No. 1 Shuaifuyuan, Dongcheng District, Beijing 100730, China. E-mail address: billliyi@126.com (Y. Li).*

Emergency and Critical Care Medicine (2025) Vol:No

Received: 24 September 2024; Accepted: 25 February 2025

Published online: June 5, 2025

http://dx.doi.org/10.1097/EC9.0000000000000147

results and conclusions have been obtained and subsequently verified by other trials.[10,11] Bioinformatics analysis can help identify differentially expressed genes (DEGs) and functional pathways involved in sepsis/septic shock.

However, false-positive rates in independent microarray analyses make it difficult to obtain reliable results. Therefore, in this study, we downloaded and analyzed multiple microarray datasets from public gene expression databases using a series of biological information technologies to identify DEGs between patients with septic shock and healthy people, performed gene enrichment analysis on the DEGs, and constructed a protein–protein interaction (PPI) network to identify hub genes related to sepsis/septic shock. Next, we analyzed the correlation between differentially expressed mRNAs and lncRNAs and clinical prognosis. Our study aimed to detect the neglected biomarkers of septic shock to better distinguish patients with septic shock from healthy controls. We aimed to identify differentially expressed mRNAs and lncRNAs related to clinical prognosis to classify patients with septic shock for treatment and prognosis evaluation.

Materials and methods

Search strategy and selection criteria for included data

We searched "sepsis" in Gene Expression Omnibus (GEO) database,[12] an accessible functional genomics database of high-throughput resources, which was one of the most commonly used sequencing (chip) data bases in the National Center for Biotechnology Information (https://pubmed.ncbi.nlm.nih.gov/), and we found 10 series from platform GPL570. Our inclusion criteria were as follows: whole blood RNA, including necessary clinical information (survival state, which represents the prognosis of disease, and Simplified Acute Physiology Score II (SAPS II), which represents the severity of disease); and samples collected at 12 hours or 24 hours after diagnosis of septic shock. Our exclusion criterion was reanalysis of other samples.

After screening, 279 patients with septic shock and 100 healthy controls from five series were included (51 samples collected on day 1 and 22 controls of GSE95233, 20 samples collected at 12 hours of GSE33118, 28 samples collected at 24 hours and 25 controls of GSE57065, 98 samples collected at 24 hours and 32 controls of GSE26440, and 82 samples collected at 24 hours and 21 controls of GSE26378). We merged the clinical information including sample IDs, age, gender (in this study, biological sex is consistent with sociological gender and there is no difference between 'sex' and 'gender'), survival state and SAPS II of 279 patients (the age range of septic shock patients in GSE33118 were replaced by middle age of each age range, in detail, 30, 50, 70, and 90 years old indicated 20–39, 40–59, 60–79, >80, respectively). A total of 279 patients with septic shock had pneumonia (n = 20) or other infections (n = 259), while 100 healthy controls were uninfected.

Merging samples from different series

First, we used R package "inSilicoMerging" to Merge the gene expression profile of samples from different series, and then removed the batch effects using empirical Bayes methods. After removing the batch effects, the data distribution among different datasets tended to be consistent with similar medians, means, and variances, and the samples among different datasets were clustered together, suggesting good removal of batch effects.

Differentially expressed genes identification

The genes whose false discovery rate (FDR) < 0.05 and fold change (FC) > 2 or FC < –2 were regarded as DEGs. Then, we used the "limma" and "pheatmap" package in R to construct a heatmap. The heatmap shows data in a two-dimensional form, in which colors represent gene expression values, providing an instant visual overview of the information and helping observers understand complex datasets.

A volcano plot was constructed using the website (http://www.ehbio.com/ImageGP/).[13]

Gene Ontology and Kyoto Encyclopedia of Genes and Genomes analyses

Gene Ontology (GO) (http://www.geneontology.org)[14] is the result of efforts to make the functional descriptions of gene products in various databases more consistent. Kyoto Encyclopedia of Genes and Genomes (KEGG) (http://www.genome.jp/kegg/)[15] is a knowledge base for analyzing gene function systems that connect genomic and high-level functional information. The R packages of "ClusterProfiler," "org.Hs.eg.db," "enrichplot," and "ggplot2" were used for enrichment analysis of GO in biological function (BP), cellular component (CC), and molecular function (MF) and KEGG pathways for the previously identified DEGs. These data were used to generate the corresponding histograms and bubble charts. GO data were acquired from "org.Hs.eg.Db" using a cutoff value of q value (a kind of adjusted *P* value, also called Benjamini-Hochberg q value or FDR) < 0.05.

Constructing PPI network of DEGs and screening of hub genes

The Search Tool for the Retrieval of Interacting Genes (STRING) (https://cn.string-db.org/)[16] is a database that searches for interactions between proteins, including both direct physical interactions and indirect functional correlations. The interaction with a combined score >0.4 was considered statistically significant. Therefore, the STRING database was used to build the PPI network.

Cytoscape, an open-source software project for integrating biomolecular interaction networks with high-throughput expression data and other molecular states into a unified conceptual framework, has been extensively used.[17] We used the Cytoscape software (version 3.10.0) to identify hub genes related to septic shock. Plug-in Molecular Complex Detection (MCODE) (version 2.0.0) is an APP of Cytoscape for clustering a given network based on topology to find densely connected regions.[18] CentiScaPe (version 2.2) is an APP to find the most important nodes in a network and calculate the centrality parameters for each node.[19] NetworkAnalyzer (version 4.4.6) is another APP to calculate the topological properties of a network.[20] The PPI network was analyzed using Cytoscape, and the most significant module was identified using the above APPs. The criteria for selection were MCODE scores >6, degree cutoff = 2, node score cutoff = 0.2, maximum depth = 100, and k-score = 2.

Additionally, we listed the 30 genes with the highest adjacent node counts from the PPI network data built using STRING. Finally, the hub genes were selected with degree ≥5 and adjacent node count ≥30.

Differentiating mRNAs and lncRNAs of DEGs and conducting statistical analyses

After identifying the hub genes of septic shock, to study the correlations between mRNA/lncRNA expression and disease severity or prognosis, we used Perl software (version 5.30.0) to differentiate the mRNAs and lncRNAs of DEGs. We then selected 38 mRNAs and two lncRNAs from BP and MF for GO and KEGG analyses. Our inclusion criteria for mRNAs were as follows: $P < 0.05$ in GO or KEGG enrichment analysis, and genes from enriched items with more than eight genes in GO or KEGG analysis. Two differentially expressed lncRNAs were included due to their limited number. Then, we linked the clinical information including sample IDs,

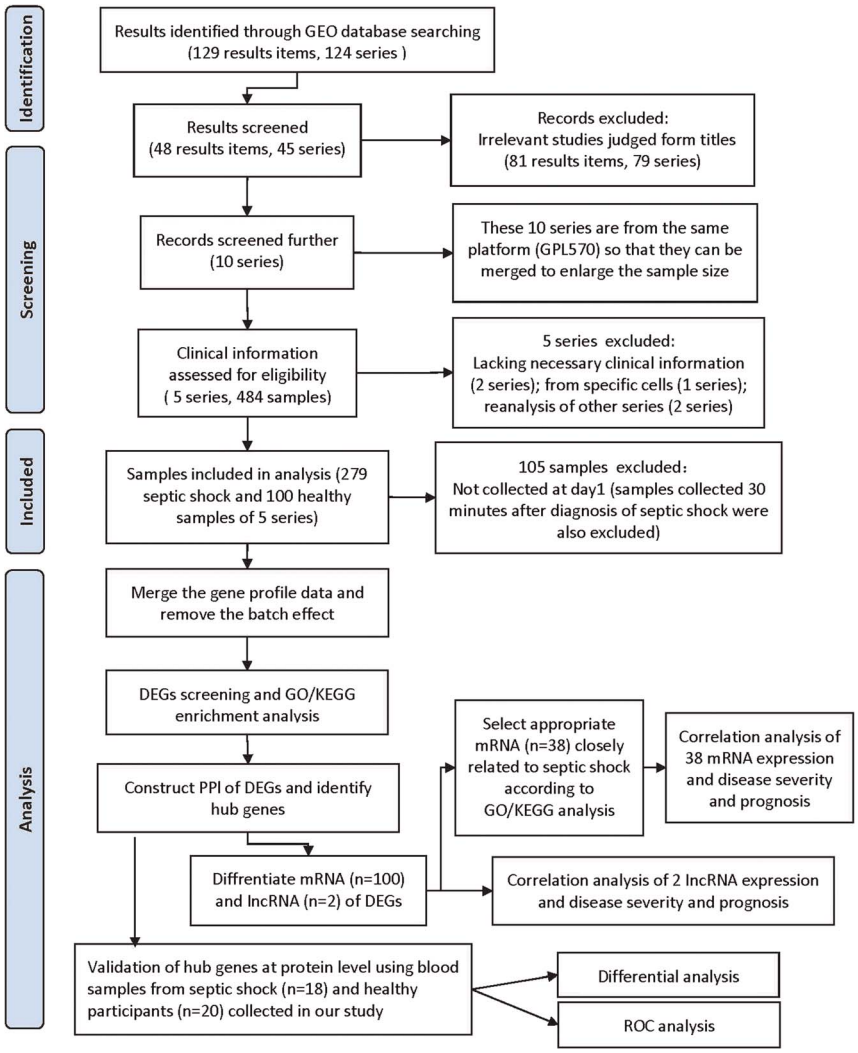

Figure 1. Flow diagram of screening samples and study design. GEO, Gene Expression Omnibus; DEG, differentially expressed gene; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; PPI, protein-protein interaction; ROC, receiver operating characteristic.

age, gender, survival state and SAPS II to mRNA or lncRNA expression profile using "limma" package in R. We used the SAPS II to represent the severity of septic shock, and the disease was more severe with higher SAPS II scores. We used survival state to represent the prognosis of septic shock. SPSS software (version 25) was used for the statistical analysis of the correlation between gene expression and clinical information. We used χ^2 test, Fisher's exact test, or Mann-Whitney *U* test to perform a simple correlation analysis between gene expression and disease severity or prognosis, as appropriate. A multiple logistic regression analysis was performed to further explore this correlation. $P < 0.05$ meant that there was a significant correlation.

Validation of identified hub genes at the protein level

We recruited patients with septic shock from the emergency department of Peking Union Medical College Hospital (PUMCH) and healthy volunteers from physical examination centers and the community. Whole blood samples were collected from the participants and serum was obtained by centrifugation for proteomic analysis. We focused on the proteins encoded by the identified hub genes and performed intergroup differential analysis using an independent two-sample *t* test with Benjamini-Hochberg correction. Receiver operating characteristic (ROC) curves were plotted to evaluate the ability to distinguish septic shock. All participants provided written informed consent, and the hospital ethics committee approved the cohort recruitment and sample collection procedures.

Results

Identification of DEGs in patients with septic shock

We collected whole blood RNA from patients with septic shock caused by pneumonia (n = 20) or other infectious diseases (n = 259) and from uninfected controls (100 healthy people) from the GEO database. The validation cohort included 18 patients with septic shock and 20 age- and sex-matched healthy volunteers. A flow diagram of the screening samples and study design is shown in Fig. 1. The expression of genes in patients with septic shock and healthy controls is shown in a volcano plot (Fig. 2A). Cluster analysis

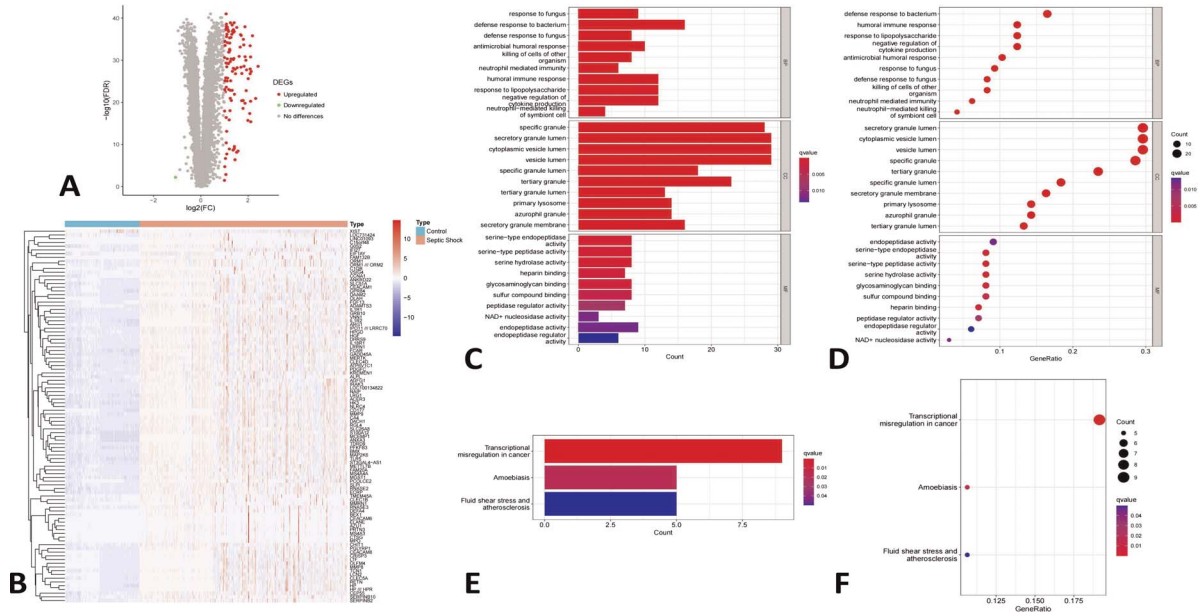

Figure 2. Differential gene expression and enrichment analyses. (A) Volcano plot: DEGs were selected with FC >2 or < −2 and FDR <0.05 among the gene expression profiles of 279 septic shock patients and 100 healthy controls. Upregulated and downregulated genes are indicated in red and green, respectively. (B) Heatmap analysis of identified DEGs between patients with septic shock and uninfected controls. Red indicates upregulated DEGs, and blue indicates downregulated DEGs. (C) GO enrichment analysis: Histograms of BP, CC, and MF analyses of upregulated DEGs in patients with septic shock. Color represents the q-value. (D) GO enrichment analysis: bubble plot of BP, CC, and MF analyses of the upregulated DEGs in patients with septic shock. Color represents the q-value. Size of the bubbles represents the gene count. (E) Histogram of KEGG analysis of the upregulated DEGs in patients with septic shock. Color represents the q-value. (F) Bubble plot of the KEGG analysis of the upregulated DEGs in patients with septic shock. Color represents the q-value. Size of the bubbles represents the gene count. DEG, differentially expressed gene; BP, biological function; CC, cellular component; MF, molecular function; FC, fold change; FDR, false discovery rate; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes.

of the identified DEGs is shown as a heatmap (Fig. 2B). We screened 112 DEGs, including 111 upregulated genes and 1 downregulated gene based on the standard of FDR < 0.05 and FC >2 or FC < –2.

GO and KEGG enrichment analyses of DEGs in septic shock

GO enrichment analysis showed that the BPs of upregulated DEGs were majorly enriched in defense response to bacteria, humoral immune response, response to lipopolysaccharide, negative regulation of cytokine production, antimicrobial humoral response, response to fungus, defense response to fungus, and killing of cells of other organisms, whereas the MFs of upregulated DEGs were mainly enriched in endopeptidase activity, serine-type endopeptidase activity, serine-type peptidase activity, serine hydrolase activity, glycosaminoglycan binding, and sulfur compound binding (Fig. 2C, Fig. 2D). The KEGG pathways of the upregulated DEGs were mainly enriched in transcriptional misregulation in cancer (Fig. 2E, Fig. 2F).

PPI network construction of DEGs and hub gene identification

Subsequently, the PPI network of the DEGs was built using the STRING database, and the most significant module was obtained using Cytoscape. The PPI network comprised 102 nodes (proteins) and 222 edges (interactions) (Fig. 3A). The 30 genes with the highest number of adjacent nodes are shown in the bar diagram (Fig. 3B). Using CentiScaPe in Cytoscape, the centrality parameters for each node, including degree, were obtained, which helped identify the most important nodes in a network. Genes with a higher degree of expression were more closely associated with septic shock. In addition, we screened the most significant module including 12 genes with degree ≥4 (Fig. 3C). The degree was not completely consistent with adjacent node count, so we selected hub genes with degree ≥5 and adjacent node count ≥30, and the final results included seven genes which was showed in Venn diagram (Fig. 3D). The results showed that the degree of LTF was 10, MMP8 was 9, MMP9 was 8, CEACAM8 was 7, CTSG was 7, LCN2 was 6, and PRTN3 was 6. Thus, we concluded that these genes were significantly upregulated hub genes in septic shock.

Correlation between mRNA expression and prognosis

Correlation between relevant mRNAs and SAPS II.

The availability of SAPS II data in 28 patients enabled exploration of correlations between mRNA expression and clinical disease severity. Regarding the correlation between sex and SAPS II, we conducted Fisher's exact test, and the results showed no statistically significant correlation ($\chi^2 = 0.164$, $P > 0.999$). Regarding the correlation between age and 38 mRNA expression levels and SAPS II, we conducted the Mann-Whitney *U* test, which showed that there were statistically significant correlations between HGF ($Z = -2.022$, $P = 0.044$), LCN2 ($Z = -2.619$, $P = 0.008$), MERTK ($Z = -2.389$, $P = 0.016$), and SAPS II, and there were no statistically significant correlations between the other 35 mRNA expression levels and SAPS II (Table 1).

To further analyze the correlation between pertinent mRNA and SAPS II, we included these three significantly statistical mRNA (HGF, LCN2, and MERTK) and age (approaching statistical significance [$Z = -1.966$, $P = 0.050$]) into multiple logistic regression model, the results illustrated that older age was associated with increased severity of septic shock (odds ratio [OR]: 1.171; 95% confidence interval [CI]: 1.023–1.339; $P = 0.022$); the upregulation of LCN2 mRNA expression demonstrated a promising association

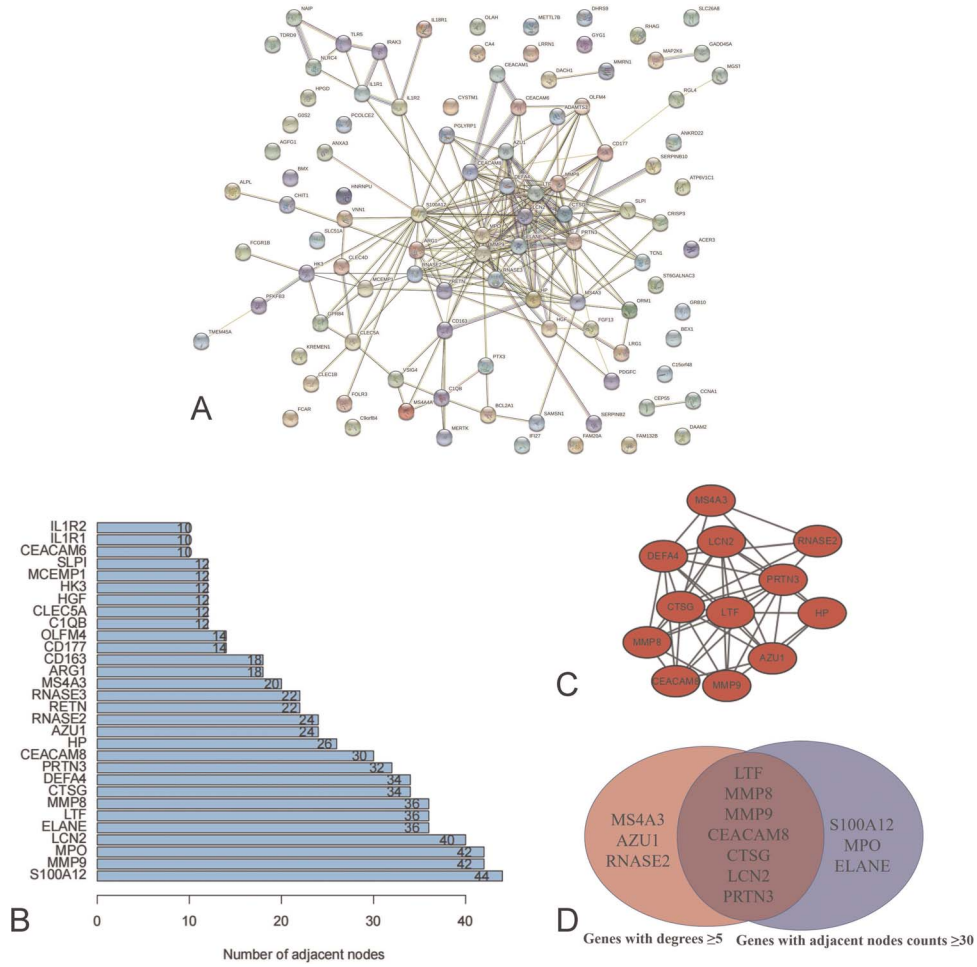

Figure 3. PPI network construction and hub gene selection. (A) PPI network was established using significantly upregulated DEGs, which contains 102 nodes and 222 edges. Nodes represent proteins, and edges indicate protein interactions. (B) First 30 genes with the most adjacent node count are listed from the PPI network data built using STRING. (C) Most significant module was obtained from a PPI network with 12 nodes and 43 edges. Upregulated genes are marked in red, and downregulated genes are marked in blue. (D) Hub genes were selected with degree ≥5 and adjacent node count ≥30. They showed an overlap of seven genes. DEG, differentially expressed gene; PPI, protein-protein interaction; STRING, Search Tool for the Retrieval of Interacting Genes.

with increased severity of septic shock (OR: 1.518; 95% CI: 0.999–2.305; $P = 0.050$), indicating its potential as a clinical biomarker despite the marginal statistical significance. There was no statistically significant correlation between the expression levels of HGF (OR: 0.971; 95% CI: 0.452–2.089; $P = 0.941$) or MERTK (OR: 3.755; 95% CI: 0.813–17.338; $P = 0.090$) and the severity of septic shock (Fig. 4A).

Correlation between relevant mRNAs and survival. Of the 69 patients included in the analysis, clinical information on survival status was available, then we analyzed the correlations between relevant mRNA expression and survival in these patients. Regarding the correlation between sex and survival state, we conducted χ^2 tests, and the results showed no statistically significant correlation ($\chi^2 = 3.423$, $P = 0.064$). As for the correlation between age and 38 mRNA expression level and survival state, we conducted Mann-Whitney U test, and the results showed that there were statistically significant correlation between ADAMTS3 ($Z = -1.984$, $P = 0.047$), ARG1 ($Z = -2.199$, $P = 0.028$), AZU1 ($Z = -3.111$, $P = 0.002$), BCL2A1 ($Z = -2.848$, $P = 0.004$), C1QB ($Z = -2.401$, $P = 0.016$), CCNA1 ($Z = -2.346$, $P = 0.019$), CTSG ($Z = -2.861$, $P = 0.004$), DEFA4 ($Z = -3.987$, $P < 0.001$), ELANE ($Z = -4.652$, $P < 0.001$), IL1R2 ($Z = -2.291$, $P = 0.022$), LCN2 ($Z = -2.677$, $P = 0.007$), MMP8 ($Z = -2.587$, $P = 0.010$), MPO ($Z = -3.847$, $P < 0.001$), PGLYRP1 ($Z = -2.067$, $P = 0.039$), PRTN3 ($Z = -3.347$, $P = 0.001$), PTX3 ($Z = -3.366$, $P = 0.001$), RNASE3 ($Z = -3.708$, $P < 0.001$), and VSIG4 ($Z = -2.775$, $P = 0.006$) and survival state, and there were no statistically significant correlation between other 20 mRNA expression level and survival state (Table 2).

To further analyze the correlation between pertinent mRNA and survival state, we included these 12 significantly statistical independent variables whose P value was less than 0.01 (AZU1, BCL2A1, CTSG, DEFA4, ELANE, LCN2, MMP8, MPO, PRTN3, PTX3, RNASE3, and VSIG4) into multiple logistic regression model, the results illustrated that the upregulation of BCL2A1 mRNA expression correlated with increased mortality risk of septic shock (OR: 1.178; 95% CI: 1.035–1.341; $P = 0.013$). There were no statistically significant correlations between the expression levels of the other 11 mRNAs and the prognosis of septic shock (Fig. 4B).

Correlation between lncRNA expression and prognosis

Correlation between relevant lncRNAs and SAPS II. We included 28 samples, including clinical information on SAPS II, to

Table 1

Simple Correlation Analysis between Pertinent mRNAs and Simplified Acute Physiology Score II

Variations	SAPS II-Low	SAPS II-High	χ^2 Test/Mann-Whitney *U* Test	
			χ^2/Z	*P*
Gender				
Male	10 (52.6%)	9 (47.4%)	0.164	>0.999
Female	4 (44.4%)	5 (55.6%)		
Age	58.5 (44.0,69.5)	74.5 (57.5,80.0)	–1.955	0.050
ADAMTS3	2.68 (2.36,3.98)	5.06 (2.61,8.12)	–1.838	0.069
ALPL	6.66 (6.17,7.76)	5.96 (5.59,7.23)	–1.195	0.246
ANXA3	16.37 (15.33,17.17)	17.61 (15.02,18.31)	–1.332	0.194
ARG1	22.07 (18.18,23.45)	21.93 (20.58,22.96)	0.000	>0.999
AZU1	4.69 (2.02,10.78)	7.11 (2.18,8.41)	–0.138	0.910
BCL2A1	8.39 (7.13,9.35)	8.96 (7.35,9.92)	–0.643	0.541
C1QB	6.46 (5.15,8.62)	6.65 (5.50,8.39)	–0.092	0.946
CCNA1	5.89 (3.61,7.23)	6.34 (3.85,8.92)	–1.011	0.329
CEACAM1	8.07 (7.30,8.90)	10.35 (6.95,11.57)	–1.424	0.164
CLEC4D	9.80 (8.92,11.13)	11.49 (8.74,12.34)	–1.287	0.210
CTSG	4.36 (3.33,9.02)	5.56 (3.44,7.16)	–0.092	0.946
DEFA4	11.75 (6.97,11.66)	12.48 (7.59,15.25)	–0.046	0.982
ELANE	8.00 (4.65,14.13)	10.27 (6.44,12.20)	–0.459	0.667
GADD45	9.35 (8.64,9.91)	10.67 (8.02,11.24)	–1.47	0.150
HGF	6.67 (5.25,8.15)	8.55 (6.04,10.91)	–2.022	0.044*
HP	21.07 (19.62,22.21)	23.52 (19.06,25.48)	–1.057	0.306
HPGD	7.60 (4.60,11.34)	8.67 (6.09,11.15)	–0.505	0.635
IL1R2	14.80 (13.08,15.08)	14.91 (13.79,15.67)	–0.827	0.427
IRAK3	8.33 (7.46,9.54)	8.68 (7.74,9.62)	–0.276	0.804
LCN2	16.80 (15.17,20.21)	21.69 (19.42,25.29)	–2.619	0.008**
LTF	17.54 (14.04,21.61)	20.45 (18.88,21.83)	–1.746	0.085
MERTK	4.01 (3.43,4.66)	6.35 (4.03,7.95)	–2.389	0.016*
MGST1	5.95 (5.04,6.91)	6.20 (5.27,7.19)	–0.597	0.571
MMP8	63.76 (60.23,66.39)	68.08 (61.35,70.60)	–1.700	0.094
MMP9	13.33 (12.12,14.12)	13.20 (11.94,14.77)	–0.459	0.667
MPO	4.80 (2.79,8.85)	4.96 (4.12,7.17)	–0.551	0.603
NAIP	9.56 (8.33,11.01)	9.92 (9.04,10.99)	–0.965	0.352
NLRC4	8.01 (7.13,8.78)	8.07 (7.24,8.97)	0.000	>0.999
ORM1	7.22 (5.18,8.40)	6.12 (4.64,8.01)	–0.827	0.427
PCOLCE2	8.87 (6.68,11.08)	12.19 (4.70,13.87)	–1.103	0.285
PGLYRP1	7.83 (6.62,8.66)	8.18 (6.94,8.54)	–0.276	0.804
PRTN3	4.08 (1.85,8.83)	4.01 (2.69,7.70)	–0.368	0.734
PTX3	4.42 (3.12,7.08)	4.65 (3.44,6.08)	0.000	>0.999
RNASE3	5.78 (4.82,7.97)	6.86 (6.25,7.61)	–1.195	0.246
S100A12	9.82 (8.74,10.13)	10.29 (9.37,10.54)	–1.792	0.077
SLPI	7.88 (6.88,8.34)	6.18 (4.60,7.96)	–1.700	0.094
TLR5	8.17 (7.79,8.61)	8.35 (7.58,8.88)	–0.046	0.982
VSIG4	7.22 (6.63,8.96)	9.54 (6.28,10.28)	–1.011	0.329

Data are reported as n (%) for categorical variables and median (IQR) for continuous variables, as appropriate.

*$P < 0.05$; **$P < 0.01$. $P < 0.05$ was considered statistically significant.

IQR, interquartile range.

analyze the correlations between pertinent lncRNAs and SAPS II. As for the correlations between the expression levels of the two lncRNAs and SAPS II, we conducted the Mann-Whitney *U* test and found no statistically significant correlation between LINC01093 (Z = –0.092, *P* = 0.946), XIST (Z = –1.378, *P* = 0.178), and SAPS II (Table 3).

To further analyze the correlation between pertinent lncRNAs and SAPS II, we included these three independent variables (age, LINC01093, and XIST) in the multiple logistic regression model and found that older age was slightly associated with increased severity of septic shock (OR: 1.075; 95% CI: 1.005–1.149; *P* = 0.035). There was no statistically significant correlation between the expression levels of LINC01093 (OR: 0.601; 95% CI: 0.273–1.327; *P* = 0.208) and XIST (OR: 1.086; 95% CI: 0.899–1.313; *P* = 0.391) and the severity of septic shock (Fig. 4C).

Correlation between pertinent lncRNAs and survival. We included 69 samples, including clinical information on survival state, to analyze the correlation between pertinent lncRNAs and survival state. As for the correlations between the expression levels of the two lncRNAs and survival state, we conducted the Mann-Whitney *U* test and found no statistically significant correlation between LINC01093 (Z = –1.044, *P* = 0.296), XIST (Z = –0.551, *P* = 0.581), and survival state (Table 4).

To further analyze the correlation between pertinent lncRNAs and survival, we included these four independent variables (age, sex, LINC01093, and XIST) in the multiple logistic regression model, which revealed that there was no statistically significant correlation between them and the prognosis of septic shock (Fig. 4D).

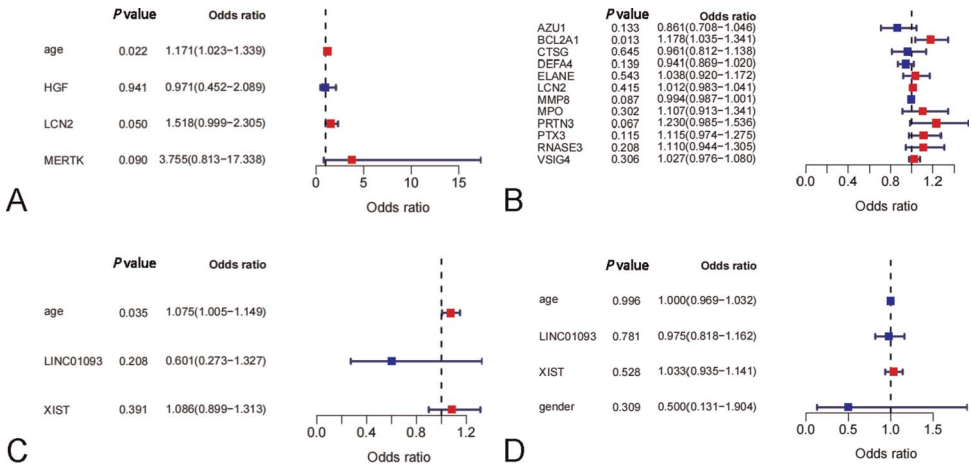

Figure 4. Multiple logistic regression between mRNAs/lncRNAs and the prognosis of septic shock. (A) Multiple logistic regression analysis of pertinent mRNAs and SAPS II. (B) Multiple logistic regression analysis of pertinent mRNA and survival. (C) Multiple logistic regression analysis of pertinent lncRNAs and SAPS II. (D) Multiple logistic regression analysis of pertinent lncRNAs and survival. SAPS II, Simplified Acute Physiology Score II.

Validation of the identified hub genes at the protein level

Among the proteins encoded by the identified hub genes, CEACAM8 was not detected in our cohort, whereas the levels of LTF, MMP8, MMP9, CTSG, LCN2, and PRTN3 were successfully measured. Comparative analysis between patients with septic shock and healthy controls revealed no significant differences in PRTN3 levels. However, the remaining five proteins were significantly upregulated in the patients with septic shock. Notably, four of these proteins demonstrated an area under the curve (AUC) greater than 0.8 for discriminating septic shock patients, as illustrated in Fig. 5.

Discussion

Sepsis/septic shock was recognized as a global health priority by the World Health Assembly and World Health Organization in 2017, owing to its high mortality rate.[21] In high-income countries, approximately 28 million deaths per year are attributed to sepsis/septic shock.[22] Overall, however, there is probably a substantial underreporting of the incidence of sepsis/septic shock, and with an aging population, the incidence will continue to increase.[23] As the pathogenesis and treatment methods remain unclear, the clinical treatment of sepsis/septic shock has placed a great burden on the society and patients.[24] Therefore, it is important to identify potential biomarkers for sepsis treatment and prognosis. Although Li et al. reported that CEACAM8, MPO, and RETN were hub genes of sepsis,[25] they lacked correlation analysis of genes and clinical information, and their results require further verification. Other studies have identified several candidate biomarkers, including CX3CR1 and LILRB2, as prognostic biomarkers of sepsis,[26,27] they did not reach a conclusion consistent with that of Li et al. Therefore, we utilized data from multiple datasets to identify DEGs in septic shock and further identified the genes most associated with septic shock and biomarkers predicting patient severity and prognosis.

Our results showed that the defense response to bacteria, humoral immune response, response to lipopolysaccharide, negative regulation of cytokine production, antimicrobial humoral response, response to fungus, defense response to fungus, and killing of cells of other organisms were the most significantly enriched BPs of upregulated DEGs. These biological processes are all related to antimicrobial effects, indicating that a large number of genes related to antimicrobial effects are expressed at higher levels in patients with septic shock. In addition, endopeptidase activity, serine-type endopeptidase activity, serine-type peptidase activity, serine hydrolase activity, glycosaminoglycan binding, and sulfur compound binding were the most significantly enriched MFs among the upregulated DEGs. Transcriptional misregulation in cancer was the most significantly enriched KEGG pathway for the upregulated DEGs. This pathway has been extensively studied in hematopoietic cancers and solid tumors.[28–31] However, its significance in sepsis and septic shock remains unclear.

Using the STRING database, Cytoscape, and R software, we found that the genes most strongly associated with septic shock were LTF, MMP8, MMP9, CEACAM8, CTSG, LCN2, and PRTN3. This included the same three genes as Pedro Martínez-Paz's results, which identified DEGs of IGHG1, IL1R2, LCN2, LTF, MMP8, and OLFM4 to distinguish septic shock from nonseptic shock in postsurgical patients.[32] Xu et al. also reported the top hub genes as MMP9, CEACAM8, ARG1, MCEMP1, LCN2, RETN, S100A12, GPR97, and TRAT1 in pediatric septic shock, three of which were the same as our results.[33]

LTF, a lactotransferrin, is a member of the transferrin family of genes and its protein product is found in the secondary granules of neutrophils. The protein and its peptides have been found to have antimicrobial, antifungal, antiparasitic, and antiviral activities, including activity against both DNA and RNA viruses and against SARS-CoV-2 and HIV.[34–38] We conclude that LTF upregulation indicates infection by various pathogens, which is consistent with evidence that LTF is involved in the infectious disease pathway (R-HSA-5663205).[39] Infection is necessary evidence for the diagnosis of sepsis/septic shock, which indicates that LTF can help diagnose the disease.

MMP8, matrix metallopeptidase 8, is a member of the MMP family of proteins.[34] It is involved in the collagen catabolic process,[35] positive regulation of neuroinflammatory response, positive regulation of interleukin-6 production, positive regulation of nitric oxide biosynthesis, and positive regulation of reactive oxygen species biosynthetic process, and so on.[40] Wong et al. regarded MMP8 as biomarker therapeutic target in septic shock,[41,42] which is consistent with our results. MMP9, matrix metallopeptidase 9, also encodes proteins of MMP family, which were involved in extracellular matrix disassembly.[34]

Table 2

Simple Correlation Analysis Between Pertinent mRNAs and Survival State

Variations	Survival	Nonsurvival	χ^2 Test/Mann-Whitney *U* Test	
			χ^2/Z	*P*
Gender				
Male	31 (72.1%)	12 (27.9%)	3.423	0.064
Female	13 (50.0%)	13 (50.0%)		
Age	4.2 (1.5,9.7)	9.6 (0.8,66.8)	–1.863	0.062
ADAMTS3	2.69 (2.09,4.68)	3.29 (2.42,4.67)	–1.984	0.047*
ALPL	4.31 (3.29,5.86)	4.35 (3.39,5.37)	–0.714	0.475
ANXA3	14.61 (9.94,19.31)	14.88 (11.96,16.98)	–0.213	0.831
ARG1	17.35 (7.90,24.42)	18.92 (15.73,26.39)	–2.199	0.028*
AZU1	2.47 (1.94,4.35)	3.75 (2.30,6.78)	–3.111	0.002**
BCL2A1	7.10 (5.21,9.09)	8.36 (6.55,9.26)	–2.848	0.004**
C1QB	4.15 (2.43,6.92)	5.55 (3.37,7.61)	–2.401	0.016*
CCNA1	3.34 (2.51,6.41)	5.21 (3.30,6.96)	–2.346	0.019*
CEACAM1	6.83 (4.44,10.26)	7.89 (5.81,10.10)	–1.688	0.091
CLEC4D	8.00 (5.04,10.89)	8.91 (6.36,10.48)	–1.203	0.229
CTSG	2.84 (2.54,3.98)	3.39 (2.72,7.74)	–2.861	0.004**
DEFA4	3.68 (2.23,12.03)	9.76 (4.38,12.28)	–3.987	<0.001***
ELANE	2.92 (1.92,7.13)	6.86 (3.37,15.07)	–4.652	<0.001***
GADD45	7.96 (4.98,9.79)	8.14 (6.87,9.76)	–1.618	0.106
HGF	4.42 (2.19,8.31)	5.18 (4.00,7.09)	–1.343	0.179
HP	19.29 (9.22,26.19)	20.39 (16.62,24.99)	–1.159	0.246
HPGD	4.48 (2.45,8.72)	5.25 (3.09,8.65)	–0.944	0.345
IL1R2	12.51 (5.97,16.60)	13.44 (11.36,16.37)	–2.291	0.022*
IRAK3	6.48 (4.38,8.38)	6.36 (5.20,8.43)	–0.123	0.902
LCN2	14.87 (4.13,24.32)	20.11 (14.40,24.12)	–2.677	0.007**
LTF	16.21 (4.31,24.99)	18.42 (13.96,21.97)	–1.852	0.064
MERTK	3.51 (2.64,5.45)	4.22 (3.39,5.36)	–1.629	0.103
MGST1	4.96 (3.53,6.20)	4.71 (3.43,6.27)	–0.247	0.805
MMP8	57.60 (3.57,95.11)	64.70 (57.90,92.35)	–2.587	0.010**
MMP9	11.82 (6.31,15.72)	11.85 (7.26,15.32)	–0.444	0.657
MPO	3.08 (2.64,4.55)	4.55 (2.97,9.70)	–3.847	<0.001***
NAIP	8.00 (5.93,10.76)	7.83 (5.52,9.79)	–1.025	0.305
NLRC4	6.72 (4.69,8.46)	6.84 (5.46,7.98)	–0.255	0.799
ORM1	4.25 (3.22,6.41)	5.65 (3.46,7.65)	–1.559	0.119
PCOLCE2	4.66 (2.48,10.25)	5.95 (2.71,10.94)	–0.833	0.405
PGLYRP1	5.87 (4.51,7.83)	6.48 (5.40,8.29)	–2.067	0.039*
PRTN3	2.19 (1.80,3.68)	3.70 (2.11,9.22)	–3.347	0.001**
PTX3	2.75 (2.19,4.57)	4.26 (2.57,6.69)	–3.366	0.001**
RNASE3	4.27 (3.15,6.17)	5.71 (4.13,7.90)	–3.708	<0.001***
S100A12	9.04 (7.66,9.88)	9.08 (8.15,9.76)	–0.583	0.56
SLPI	5.12 (3.96,7.69)	6.16 (4.35,7.97)	–1.176	0.24
TLR5	7.16 (5.72,8.66)	6.64 (5.17,8.01)	–1.88	0.06
VSIG4	4.29 (2.68,7.18)	5.41 (3.98,9.27)	–2.775	0.006**

Data are reported as n (%) for categorical variables and median (IQR) for continuous variables, as appropriate.

*$P < 0.05$; **$P < 0.01$; ***$P < 0.001$. $P < 0.05$ was considered statistically significant.

IQR, interquartile range.

CEACAM8, CEA cell adhesion molecule 8, belongs to the carcinoembryonic antigen (CEA) family of the immunoglobulin.[43] It is involved in heterophilic cell-cell adhesion via plasma membrane cell adhesion molecules and the immune response.[35,44] Several studies have reported that CEACAM8 exerts immunomodulatory effects on sepsis occurrence.[45]

Literature retrieval revealed that the interaction between septic shock and the hub genes CTSG and PRTN3 has not been widely reported. Myeloperoxidase (MPO) is the main enzyme produced by neutrophils and has antibacterial function in sepsis.[46] Demaret et al. believed that upregulated MPO expression was the best predictor for identifying a subgroup of high-risk death patients.[47] However, MPO was not screened as a hub gene in our study because of differences in the methods and criteria.

LCN2, lipocalin 2, encodes a protein belonging to the lipocalin family. This protein is a neutrophil gelatinase-associated lipocalin that plays a role in innate immunity by limiting bacterial growth.[48] We found that the upregulation of LCN2 mRNA expression might be indicative of an increased severity of septic shock and that the upregulation of BCL2A1 mRNA expression could indicate an increased mortality risk in septic shock. Therefore, these two mRNA can potentially be used as biomarkers to predict the severity and prognosis of septic shock.

Compared to traditional markers, such as cytokines, gene biomarkers may exhibit greater advantages in reflecting an individual's susceptibility to septic shock.[49] Although traditional cytokine biomarkers (e.g., IL-6 and IL-10) demonstrate diagnostic value in septic shock, their utility is limited by their significant dynamic variability

Table 3

Simple Correlation Analysis between Pertinent lncRNAs and Simplified Acute Physiology Score II

Variations	SAPS II-Low	SAPS II-High	χ^2 Test/Mann-Whitney *U* Test	
			χ^2/Z	*P*
Gender				
Male	10 (52.6%)	9 (47.4%)	0.164	>0.999
Female	4 (44.4%)	5 (55.6%)		
Age	58.5 (44.0,69.5)	74.5 (57.5,80.0)	–1.955	0.050
LINC01093	4.32 (3.99,6.45)	4.75 (4.17,5.22)	–0.092	0.946
XIST	6.86 (6.38,14.56)	7.09 (6.77,17.18)	–1.378	0.178

Data are reported as n (%) for categorical variables and median (IQR) for continuous variables, as appropriate. $P < 0.05$ was considered statistically significant.

IQR, interquartile range.

and susceptibility to multiple confounding factors. In contrast, genomic biomarkers (e.g., mRNA expression profiles) exhibit superior stability and specificity. For instance, while IL-6 and IL-10 levels show significant disparities between patients with gram-positive and gram-negative sepsis,[50] transcriptomic analyses revealed no substantial differences between these groups, indicating a shared host response at the transcriptional level.[51] In clinical practice, these biomarkers can be assessed at the genetic level through mRNA quantification, or at the protein level by measuring the expression of corresponding proteins, potentially providing valuable insights into the diagnosis of septic shock and clinical prognosis.

Limitations

This study had some limitations. First, the specific mechanisms of hub genes in septic shock need to be further explored. Second, we lacked data on survival or in-hospital time; therefore, we could not further analyze the correlation between key genes and survival or in-hospital time. Detailed clinical information, such as the infection site and pathogens, could not be obtained, so its impact on mRNA and lncRNA expression was not further discussed. In future studies, we will collect more clinical information to further explore the correlation between gene expression and the prognosis of patients with sepsis, and investigate the influence of other clinical factors on the expression of these biomarkers.

Conclusion

In this study, we identified seven hub genes of septic shock among the DEGs, particularly MMP9, LCN2, CTSG, and LTF, which were significantly upregulated in septic shock and were validated at the protein level with strong diagnostic performance. We found that the upregulation of LCN2 mRNA expression showed a trend toward an association with increased severity, and the upregulation of BCL2A1 mRNA expression correlated with increased mortality in patients with septic shock. This study will help us further understand the molecular mechanisms underlying septic shock and provide candidate biomarkers and targets for the rapid and accurate diagnosis of septic shock and prediction of prognosis.

Conflict of interest statement

The authors declare no conflict of interest.

Author contributions

Gong C participated in concept, design, data acquisition, data analysis, manuscript preparation, and manuscript editing. Zhang W participated in design, literature search, data acquisition, data analysis, statistical analysis, and manuscript preparation. Lu X, Yu S, Ge Z, and Qin M participated in data analysis, statistical analysis, and manuscript editing. Zhu H participated in manuscript review. Li Y participated in manuscript editing and review. Gong C and Zhang W contributed equally to the study.

Funding

This work was supported by the National High Level Hospital Clinical Research Funding [grant number 2022-PUMCH-B-109] and the Chinese Academy of Medical Science Innovation Fund for Medical Sciences (CIFMS) [grant number 2021-I2M-1-020].

Ethical approval of studies and informed consent

The use of public databases does not require ethical approval or informed consent. For our own cohort, it was approved by the ethics committee of Peking Union Medical College Hospital (K2424) on April 10, 2023, and written informed consent was obtained from all participants.

Table 4

Simple Correlation Analysis Between Pertinent lncRNAs and Survival State

Variations	Survival	Nonsurvival	χ^2 Test/Mann-Whitney *U* Test	
			χ^2/Z	*P*
Gender				
Male	31 (72.1%)	12 (27.9%)	3.423	0.064
Female	13 (50.0%)	13 (50.0%)		
Age	4.2 (1.5,9.7)	9.6 (0.8,66.8)	–1.863	0.062
LINC01093	2.43 (1.62,4.69)	2.89 (1.73,5.77)	–1.044	0.296
XIST	5.24 (–0.61,18.43)	7.71 (–1.13,17.29)	–0.551	0.581

Data are reported as n (%) for categorical variables and median (IQR) for continuous variables, as appropriate. $P < 0.05$ was considered statistically significant.

IQR, interquartile range.

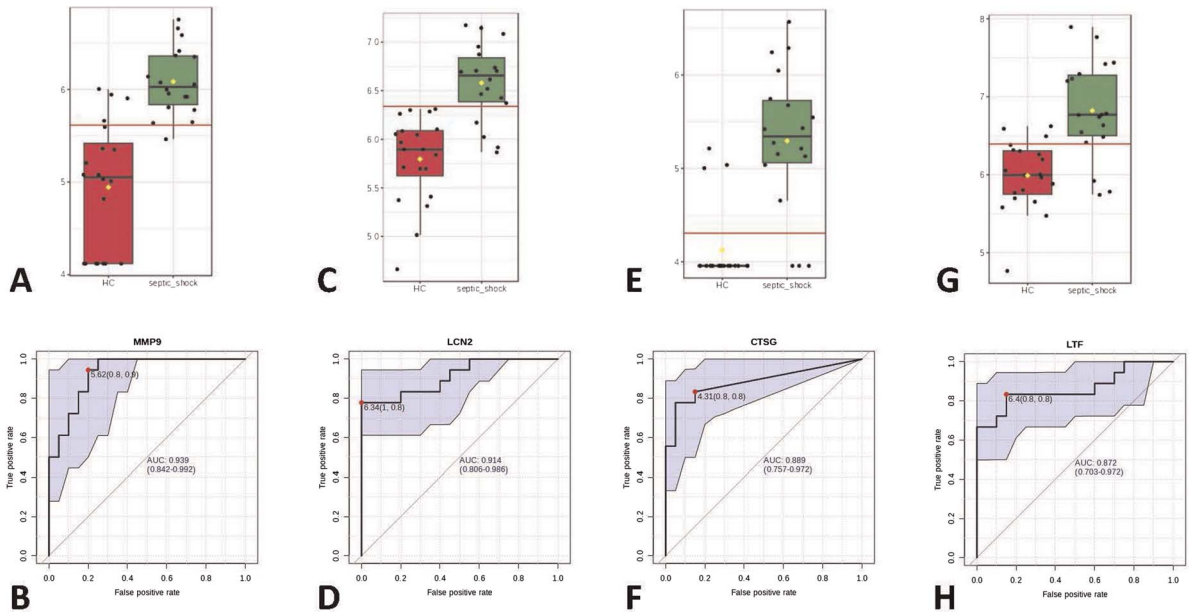

Figure 5. Validation of identified hub genes at the protein level using blood samples from patients with septic shock and healthy control (HC) subjects. (A) Boxplot showing the abundance of MMP9 protein in the two groups, with a fold change (FC) of 7.22 and false discovery rate (FDR) of 2.21e-06. (B) ROC curve demonstrating the ability of MMP9 protein to discriminate between patients with septic shock and healthy controls. (C) Box plot showing the abundance of LCN2 protein in the two groups (FC = 5.99 and FDR = 3.97e-06). (D) ROC curve demonstrating the ability of LCN2 protein to discriminate between patients with septic shock and healthy controls. (E) Box plot showing the abundance of CTSG protein in the two groups (FC = 23.73 and FDR = 3.50e-06). (F) ROC curve demonstrating the ability of CTSG protein to discriminate between patients with septic shock and healthy controls. (G) Box plot showing the abundance of LTF protein in the two groups (FC = 10.82 and FDR = 4.81e-05). (H) ROC curve demonstrating the ability of LTF protein to discriminate between patients with septic shock and healthy controls. Comparison of protein abundance between the two groups was performed using an independent two-sample *t* test with Benjamini-Hochberg correction. AUC, area under the curve; ROC, receiver operating characteristic.

Acknowledgments

We would like to extend our sincere gratitude to Professor Xiaomin Hu and Ms. Siqi Sun (both from Peking Union Medical College Hospital) for their invaluable support and assistance with proteomic data analysis. We would like to thank the funds and all original authors who provided publicly available data.

References

[1] Singer M, Deutschman CS, Seymour CW, et al. The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). *JAMA*. 2016;315(8):801–810. doi:10.1001/jama.2016.0287

[2] Evans L, Rhodes A, Alhazzani W, et al. Surviving sepsis campaign: international guidelines for management of sepsis and septic shock 2021. *Intensive Care Med*. 2021;47(11):1181–1247. doi:10.1007/s00134-021-06506-y

[3] Salomão R, Ferreira BL, Salomão MC, Santos SS, Azevedo LCP, Brunialti MKC. Sepsis: evolving concepts and challenges. *Braz J Med Biol Res*. 2019;52(4):e8595. doi:10.1590/1414-431x20198595

[4] Fleischmann-Struzek C, Mellhammar L, Rose N, et al. Incidence and mortality of hospital- and ICU-treated sepsis: results from an updated and expanded systematic review and meta-analysis. *Intensive Care Med*. 2020;46(8):1552–1562. doi:10.1007/s00134-020-06151-x

[5] Huang M, Cai S, Su J. The pathogenesis of sepsis and potential therapeutic targets. *Int J Mol Sci*. 2019;20(21):5376. doi:10.3390/ijms20215376

[6] Haak BW, Prescott HC, Wiersinga WJ. Therapeutic potential of the gut microbiota in the prevention and treatment of sepsis. *Front Immunol*. 2018;9:2042. doi:10.3389/fimmu.2018.02042

[7] Zhou J, Dong S, Wang P, Su X, Cheng L. Identification of nine mRNA signatures for sepsis using random forest. *Comput Math Methods Med*. 2022;2022:5650024. doi:10.1155/2022/5650024

[8] Bourcier S, Hindlet P, Guidet B, Dechartres A. Reporting of organ support outcomes in septic shock randomized controlled trials: a methodologic review-the sepsis organ support study. *Crit Care Med*. 2019;47(7):984–992. doi:10.1097/ccm.0000000000003746

[9] Ditty JL, Kvaal CA, Goodner B, et al. Incorporating genomics and bioinformatics across the life sciences curriculum. *PLoS Biol*. 2010;8(8):e1000448. doi:10.1371/journal.pbio.1000448

[10] Du W, Sun J, Gu J, Zhang S, Zhang T. Bioinformatics analysis of LINC00426 expression in lung cancer and its correlation with patients' prognosis. *Thorac Cancer*. 2020;11(1):150–155. doi:10.1111/1759-7714.13228

[11] Zhang G, Tang X, Liang L, et al. DNA and RNA sequencing identified a novel oncogene VPS35 in liver hepatocellular carcinoma. *Oncogene*. 2020;39(16):3229–3244. doi:10.1038/s41388-020-1215-6

[12] Barrett T, Wilhite SE, Ledoux P, et al. NCBI GEO: archive for functional genomics data sets—update. *Nucleic Acids Res*. 2013;41(Database issue):D991–D995. doi:10.1093/nar/gks1193

[13] Chen T, Liu YX, Huang L. ImageGP: an easy-to-use data visualization web server for scientific researchers. *Imeta*. 2022;1(1):e5. doi:10.1002/imt2.5

[14] Gene Ontology Consortium. The Gene Ontology resource: enriching a GOld mine. *Nucleic Acids Res*. 2021;49(D1):D325–D334. doi:10.1093/nar/gkaa1113

[15] Kanehisa M, Sato Y, Furumichi M, Morishima K, Tanabe M. New approach for understanding genome variations in KEGG. *Nucleic Acids Res*. 2019;47(D1):D590–D595. doi:10.1093/nar/gky962

[16] Szklarczyk D, Gable AL, Nastou KC, et al. The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res*. 2021;49(D1):D605–D612. doi:10.1093/nar/gkaa1074

[17] Shannon P, Markiel A, Ozier O, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*. 2003;13(11):2498–2504. doi:10.1101/gr.1239303

[18] Bader GD, Hogue CW. An automated method for finding molecular complexes in large protein interaction networks. *BMC Bioinformatics*. 2003;4:2. doi:10.1186/1471-2105-4-2

[19] Scardoni G, Petterlini M, Laudanna C. Analyzing biological network parameters with CentiScaPe. *Bioinformatics*. 2009;25(21):2857–2859. doi:10.1093/bioinformatics/btp517

[20] Assenov Y, Ramírez F, Schelhorn SE, Lengauer T, Albrecht M. Computing topological parameters of biological networks. *Bioinformatics*. 2008;24(2):282–284. doi:10.1093/bioinformatics/btm554

[21] Reinhart K, Daniels R, Kissoon N, Machado FR, Schachter RD, Finfer S. Recognizing sepsis as a global health priority - a WHO resolution. *N Engl J Med*. 2017;377(5):414–417. doi:10.1056/NEJMp1707170
[22] Adhikari NK, Fowler RA, Bhagwanjee S, Rubenfeld GD. Critical care and the global burden of critical illness in adults. *Lancet*. 2010; 376(9749):1339–1346. doi:10.1016/s0140-6736(10)60446-1
[23] Cecconi M, Evans L, Levy M, Rhodes A. Sepsis and septic shock. *Lancet*. 2018;392(10141):75–87. doi:10.1016/s0140-6736(18)30696-2
[24] Kalil AC, Opal SM. Sepsis in the severely immunocompromised patient. *Curr Infect Dis Rep*. 2015;17(6):487. doi:10.1007/s11908-015-0487-4
[25] Li Y, Zhang H, Shao J, et al. Bioinformatics analysis for identifying pertinent pathways and genes in sepsis. *Comput Math Methods Med*. 2021;2021:2085173. doi:10.1155/2021/2085173
[26] Tabone O, Mommert M, Jourdan C, et al. Endogenous retroviruses transcriptional modulation after severe infection, trauma and burn. *Front Immunol*. 2018;9:3091. doi:10.3389/fimmu.2018.03091
[27] Venet F, Schilling J, Cazalis MA, et al. Modulation of LILRB2 protein and mRNA expressions in septic shock patients and after ex vivo lipopolysaccharide stimulation. *Hum Immunol*. 2017;78(5–6):441–450. doi:10.1016/j.humimm.2017.03.010
[28] Alcalay M, Orleth A, Sebastiani C, et al. Common themes in the pathogenesis of acute myeloid leukemia. *Oncogene*. 2001;20(40): 5680–5694. doi:10.1038/sj.onc.1204642
[29] Scandura JM, Boccuni P, Cammenga J, Nimer SD. Transcription factor fusions in acute leukemia: variations on a theme. *Oncogene*. 2002; 21(21):3422–3444. doi:10.1038/sj.onc.1205315
[30] Roumier C, Fenaux P, Lafage M, Imbert M, Eclache V, Preudhomme C. New mechanisms of AML1 gene alteration in hematological malignancies. *Leukemia*. 2003;17(1):9–16. doi:10.1038/sj.leu.2402766
[31] Koschmieder S, Halmos B, Levantini E, Tenen DG. Dysregulation of the C/EBPalpha differentiation pathway in human cancer. *J Clin Oncol*. 2009;27(4):619–628. doi:10.1200/jco.2008.17.9812
[32] Martínez-Paz P, Aragón-Camino M, Gómez-Sánchez E, et al. Distinguishing septic shock from non-septic shock in postsurgical patients using gene expression. *J Infect*. 2021;83(2):147–155. doi:10.1016/j.jinf.2021.05.039
[33] Xu Z, Jiang M, Bai X, Ding L, Dong P, Jiang M. Identification and verification of potential core genes in pediatric septic shock. *Comb Chem High Throughput Screen*. 2022;25(13):2228–2239. doi:10.2174/1386207325666220310110902
[34] DiStefano MT, Goehringer S, Babb L, et al. The Gene Curation Coalition: a global effort to harmonize gene-disease evidence resources. *Genet Med*. 2022;24(8):1732–1742. doi:10.1016/j.gim.2022.04.017
[35] Gaudet P, Livstone MS, Lewis SE, Thomas PD. Phylogenetic-based propagation of functional annotations within the Gene Ontology consortium. *Brief Bioinform*. 2011;12(5):449–462. doi:10.1093/bib/bbr042
[36] Lupetti A, Paulusma-Annema A, Welling MM, Senesi S, van Dissel JT, Nibbering PH. Candidacidal activities of human lactoferrin peptides derived from the N terminus. *Antimicrob Agents Chemother*. 2000; 44(12):3257–3263. doi:10.1128/aac.44.12.3257–3263.2000
[37] Chapple DS, Hussain R, Joannou CL, et al. Structure and association of human lactoferrin peptides with Escherichia coli lipopolysaccharide. *Antimicrob Agents Chemother*. 2004;48(6):2190–2198. doi:10.1128/aac.48.6.2190-2198.2004
[38] Nozaki A, Ikeda M, Naganuma A, et al. Identification of a lactoferrin-derived peptide possessing binding activity to hepatitis C virus E2 envelope protein. *J Biol Chem*. 2003;278(12):10162–10173. doi:10.1074/jbc.M207879200
[39] National Center for Biotechnology Information (2022). PubChem pathway summary for pathway R-HSA-5663205, infectious disease, source: reactome. Assessed May 31, 2022. https://pubchem.ncbi.nlm.nih.gov/pathway/Reactome:R-HSA-5663205
[40] Lee EJ, Han JE, Woo MS, et al. Matrix metalloproteinase-8 plays a pivotal role in neuroinflammation by modulating TNF-α activation. *J Immunol*. 2014;193(5):2384–2393. doi:10.4049/jimmunol.1303240
[41] Wong HR, Salisbury S, Xiao Q, et al. The pediatric sepsis biomarker risk model. *Crit Care*. 2012;16(5):R174. doi:10.1186/cc11652
[42] Wong HR, Cvijanovich NZ, Anas N, et al. Improved risk stratification in pediatric septic shock using both protein and mRNA biomarkers. PERSEVERE-XP. *Am J Respir Crit Care Med*. 2017;196(4):494–501. doi:10.1164/rccm.201701-0066OC
[43] Jog NR, Rane MJ, Lominadze G, Luerman GC, Ward RA, McLeish KR. The actin cytoskeleton regulates exocytosis of all neutrophil granule subsets. *Am J Physiol Cell Physiol*. 2007;292(5):C1690–C1700. doi:10.1152/ajpcell.00384.2006
[44] Berling B, Kolbinger F, Grunert F, et al. Cloning of a carcinoembryonic antigen gene family member expressed in leukocytes of chronic myeloid leukemia patients and bone marrow. *Cancer Res*. 1990;50(20):6534–6539.
[45] Ribon M, Mussard J, Semerano L, Singer BB, Decker P. Extracellular chromatin triggers release of soluble CEACAM8 upon activation of neutrophils. *Front Immunol*. 2019;10:1346. doi:10.3389/fimmu.2019.01346
[46] Yu H, Liu Y, Wang M, et al. Myeloperoxidase instigates proinflammatory responses in a cecal ligation and puncture rat model of sepsis. *Am J Physiol Heart Circ Physiol*. 2020;319(3):H705–H721. doi:10.1152/ajpheart.00440.2020
[47] Demaret J, Venet F, Friggeri A, et al. Marked alterations of neutrophil functions during sepsis-induced immunosuppression. *J Leukoc Biol*. 2015;98(6):1081–1090. doi:10.1189/jlb.4A0415-168RR
[48] Shields-Cutler RR, Crowley JR, Miller CD, Stapleton AE, Cui W, Henderson JP. Human metabolome-derived cofactors are required for the antibacterial activity of siderocalin in urine. *J Biol Chem*. 2016; 291(50):25901–25910. doi:10.1074/jbc.M116.759183
[49] Wu M, Mi B, Liu L, et al. Genetic polymorphisms, biomarkers and signaling pathways associated with septic shock: from diagnosis to therapeutic targets. *Burns Trauma*. 2024;12:tkae006. doi:10.1093/burnst/tkae006
[50] Zhang Y, Li B, Ning B. Evaluating IL-6 and IL-10 as rapid diagnostic tools for gram-negative bacteria and as disease severity predictors in pediatric sepsis patients in the intensive care unit. *Front Immuno*. 2022;13:1043968. doi:10.3389/fimmu.2022.1043968
[51] Tang BM, McLean AS, Dawes IW, Huang SJ, Cowley MJ, Lin RC. Gene-expression profiling of gram-positive and gram-negative sepsis in critically ill patients. *Crit Care Med*. 2008;36(4):1125–1128. doi:10.1097/CCM.0b013e3181692c0b

How to cite this article: Gong C, Zhang W, Lu X, et al. Transcriptome analysis for identifying hub genes and prognosis biomarkers of mrna/lncrna in septic shock. *Emerg Crit Care Med*. 2025;00(00):00–00. doi: 10.1097/EC9.0000000000000147