Cuproptosis-Related Biomarkers and Characterization of Immune Infiltration in Sepsis

Yuanfeng Wang1,*, Xu Qiu1,*, Jiao Liu1,*, Xuanyi Liu1, Jialu Pan2, Jiayi Cai2, Xiaodong Liu1,3, Shugen Qu1,3

1College of Public Health and Management, Zhejiang Provincial Key Laboratory of Watershed Science and Health, Wenzhou Medical University, Wenzhou, People’s Republic of China; 2The First Clinical Medical College, Wenzhou Medical University, Wenzhou, People’s Republic of China; 3South Zhejiang Institute of Radiation Medicine and Nuclear Technology, Wenzhou, People’s Republic of China

*These authors contributed equally to this work.

Correspondence: Xiaodong Liu; Shugen Qu, Email forget45@wmu.edu.cn; shugenju@wmu.edu.cn

Introduction: Sepsis is a worldwide epidemic, with high morbidity and mortality. Cuproptosis is a form of cell death that is associated with a wide range of diseases. This study aimed to explore genes associated with cuproptosis in sepsis, construct predictive models and screen for potential targets.

Methods: The LASSO algorithm and SVM-RFE model has been analysed the expression of cuproptosis-related genes in sepsis and immune infiltration characteristics and identified the marker genes under a diagnostic model. Gene-drug networks, mRNA-miRNA networks and PPI networks were constructed to screen for potential biological targets. The expression of marker genes was validated based on the GSE57065 dataset. Consensus clustering method was used to classify sepsis samples.

Results: We found 381 genes associated with the development of sepsis and discovered significantly differentially expressed cuproptosis-related genes of 16 cell types in sepsis and immune infiltration with CD8/CD4 T cells being lower. NFE2L2, NLRP3, SLC31A1, DLD, DLAT, PDHB, MTF1, CDKN2A and DLST were identified as marker genes by the LASSO algorithm and the SVM-RFE model. AUC > 0.9 was constructed for PDHB and MTF1 alone respectively. The validation group data for PDHB (P<0.00099) and MTF1 (P=7.2e-14) were statistically significant. Consistent clustering analysis confirmed two subtypes. The C1 subtype may be more relevant to cellular metabolism and the C2 subtype has some relevance to immune molecules. The results of animal experiments showed that the gene expression was consistent with the bioinformatics analysis.

Discussion: Our study systematically explored the relationship between sepsis and cuproptosis and constructed a diagnostic model. And, several cuproptosis-related genes may interfere with the progression of sepsis through immune cell infiltration.

Keywords: machine learning, cuproptosis, immune infiltration, sepsis, diagnostic model

Introduction

Sepsis has a high morbidity and mortality rate, is widespread worldwide and the main cause of death is related to an overreaction and dysfunction of the body’s immune system. Epidemiological findings show that 48.9 million people were diagnosed with sepsis and 11 million died as a result of sepsis in 2017. Other than supportive care there is no other way to treat sepsis directly and effectively. A growing number of studies have demonstrated that the hyperinflammatory response caused by sepsis leads to systemic tissue and organ destruction and ultimately to multiple organ failure. Early diagnosis and intervention in sepsis are crucial to its treatment. Hence, there is a need to explore in greater depth the pathological processes, biomarkers, and potential drug targets related to sepsis. This will help to further enhance our understanding of the disease and reduce mortality from sepsis.

As a micronutrient, copper plays a very significant role in a variety of life activities. In particular, it acts as a cofactor for enzymes in mitochondrial energy metabolism. Copper plays an important role in the tricarboxylic acid cycle, ultimately affecting oxidative phosphorylation, cell growth, iron utilization and various other vital life processes.
Moreover, the results of an epidemiological study based on a prospective cohort study showed the concentration of copper ions in the blood of people with sepsis is higher than normal. Related studies have shown that depletion of Fe-S cluster proteins and excessive accumulation of lipidated mitochondrial enzymes trigger mitochondrial stress and ultimately promote the development of cuproptosis. A rising number of studies indicate that deficiencies in energy metabolism and oxidative stress due to mitochondrial dysfunction may be a key pathogenetic mechanism in sepsis-induced multi-organ failure. The application of mitochondrial protectors can reduce the effects of sepsis-mediated organic damage. In summary, we deduced that cuproptosis is closely related to sepsis-induced multi-organ damage. However, no studies have been reported to date on the mechanisms regulating copper death in the development of sepsis, and further insightful research is needed. Thus, potential biomarkers for early diagnosis and treatment of sepsis may be obtained by further molecular characterisation of cuproptosis-related genes (CRGs).

In recent years, bioinformatics analysis has been used to explore the process of disease development and to find potential biological targets, which will help to enable early diagnosis of diseases, prognostic evaluation and treatment of diseases through drug screening. In this study, differentially expressed CRGs and immune profiles between normal and septicemic individuals were examined using the dataset (GSE65682). We used machine learning models to filter for key genes, construct a diagnostic model for sepsis and validate the performance of the predictive model in an external dataset (GSE57065). Potential drugs and microRNAs were predicted based on key genes in the model. We divided sepsis patients into two subtypes, analysed their CRGs expression and immune infiltration and performed GSEA analysis according to the CRGs expression landscapes. This will assist us in further elucidating the heterogeneity of sepsis. We believe our findings will provide greater insight into the characterization of cuproptosis progression in sepsis and provide potential prognostic biomarkers for the design of rational treatment regimens.

Materials and Methods
Data Source
The original expression profile data for the study was downloaded from the NCBI Gene Expression Omnibus (GEO). In this study, the GSE65682 dataset was used as the training set. The GSE57065 dataset was validated for expression of marker genes. And the detailed datasets are listed in Table 1. The GSE65682 dataset included 46 control blood samples and 760 blood samples that met the diagnostic criteria for sepsis. Specific clinical information is provided in Table S1. In addition, the GSE57065 dataset included 25 control blood samples and 82 blood samples that met the diagnostic criteria for sepsis. Specific clinical information is also shown in Table S1.

The Drug-Gene Interaction database (DGIdb) is a database that holds a large number of drug-gene interactions and we use it to screen for drugs that interact with marker genes. Three databases, miRanda, miRDB and TargetScan, were used to analyze mRNA-miRNA interaction pairs.

Differential Expression Analysis
The limma package was applied to screen for differentially expressed cuproptosis-associated genes. If $P < 0.05$ and $\log_{2}FC > 1$ will be identified as differentially expressed genes (DEGs).

Weighted Gene Co-Expression Network Analysis (WGCNA)
WGCNA was performed for the identification of co-expression modules using the R package “WGCNA”. To guarantee accurate results, the top 25% of genes with the highest variance were utilized for subsequent WGCNA analyses. To construct a weighted adjacency matrix, an optimal soft power was selected. This was then transformed into a topological

<table>
<thead>
<tr>
<th>Table 1 Data Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data</td>
</tr>
<tr>
<td>GSE65682</td>
</tr>
<tr>
<td>GSE57065</td>
</tr>
</tbody>
</table>
overlap matrix (TOM). Relying on the hierarchical clustering tree algorithm, TOM dissimilarity measure (1-TOM) was used to acquire modules, where minimum module size was set as 100. A random color was designated to every module. The global gene expression profiles were each represented by module eigengenes. Module Significance (MS) described the relationship between the modules and the disease status. Gene Significance (GS) was stated as the correlation between a gene and the clinical phenotype.

**Functional Enrichment**
Gene ontology (GO) analysis can reveal molecular functions (MF), biological processes (BP) and cellular components (CC) based on gene sets. In addition, KEGG can suggest possible cellular signalling pathways, biological processes, and more. Based on sepsis-related DEGs and differentially expressed genes associated with cuproptosis, analyses were performed using the cluster Profiler package in R.\(^{23}\) \(P < 0.05\) was used as a significant enrichment threshold.

**Immune Infiltration Analysis**
Using the CIBERSORT algorithm and the LM22 feature matrix, the corresponding ratios of the 22 immune cell types in each sample can be estimated. Only specimens with \(P < 0.05\) were regarded as an accurate fraction of immune cells.\(^{24}\)

**Correlation Analysis Between CRGs and Infiltrated Immune Cells**
The correlation coefficient between CRGs expression and the relative percentage of immune cells was analyzed using the Spearman correlation coefficient, with \(p < 0.05\) indicating a relevant correlation. The conclusion was presented using the R package “corrplot”.

**Identifying the Best Diagnostic Genetic Biomarkers for Sepsis**
The least absolute shrinkage and selection operator (LASSO) algorithms is exerted in the glmnet package to decrease the dimensionality of the data.\(^{25,26}\) For feature selection, LASSO algorithms were used to identify gene biomarkers for sepsis by retaining the differently expressed CRGs between patients with sepsis and normal samples. At the same time, the SVM package was used to construct models to screen for key genes.\(^{27}\) The crossover genes obtained by the two algorithms were considered as the best marker genes and logistic regression models were constructed for the diagnosis of the disease. The diagnostic power of the logistic regression model was then assessed using ROC curves.

**Single-Gene Gene Set Enrichment Analysis (GSEA) Enrichment Analysis**
This analysis is implemented in the GSEA package in R. To further explore the related pathways of the marker genes, the correlation between the marker genes and all the other genes in the GSE65682 data set was calculated. Then, all the genes were sorted from the top to the bottom according to their correlations, and these sorted genes were contemplated to be tested. In addition, the GSEA analysis in our study was implemented based on the gene set of KEGG.\(^{28}\)

**Single-Gene Gene Set Variation Analysis (GSVA) Enrichment Analysis**
GSVA package based gene set variation analysis was implemented.\(^{29}\) We performed GSVA analysis for each marker gene. Differences in scores between high and low expression groups of marker genes were analysed using the limma package. If |\(t\)| \(> 2\), \(P < 0.05\), we consider the two to be different. And \(t > 0\) means that the pathway is promoted in the high expression group. Conversely, \(t < 0\) means that the pathway is upregulated in the low expression group.

**GeneMANIA Analysis**
GeneMANIA (http://genemania.org) is a flexible, user-friendly website to hypothesize gene function, analyze gene lists and prioritize genes to be tested functionally. Using the identified genetic biomarkers as a gene list, GeneMANIA finds functionally similar genes based on genomic and proteomic data. In this mode, it weights each functional genomic dataset in accordance with its predictive value for the query.\(^{30}\)
Unsupervised Clustering of Patients with Sepsis

According to the characteristics of the 14 differentially expressed CRGs, we performed an unsupervised cluster analysis using the “ConsensusClusterPlus” R package, grouping the 760 sepsis samples into different clusters using the k-means algorithm with 1000 iterations.

Animal Grouping and Model Establishment

C57BL/6JGpt mice (Strain NO. N000013) were purchased from GemPharmatech (Nanjing, China). Ten 12-week-old C57BL/6JGpt male mice (wydw2023-0561) were randomly divided into two groups, labeled and weighed. Mice in the experimental group were injected LPS (lipopolysaccharide) into the intraperitoneal cavity at the dose of 0.01mg LPS/g, and the control group was injected with the corresponding dose of normal saline for 24 hours. Blood was collected from the eyeballs of mice 24 hours later for follow-up experiments. The experiment has been approved by the Animal Ethics Committee of Wenzhou Medical University.

Materials and Reagents

RNAiso Blood (Takara) (9112), PrimeScript™ RT reagent Kit with gDNA Eraser (Perfect Real Time) (Takara) (RR047A), Mouse CRP (C-Reactive Protein) ELISA Kit (Sangon Biotech) (JL47089), 2X SanTaq PCR Master Mix (with Blue Dye) (Sangon Biotech) (B532061-0005).

Experimental Method

Each mouse was scored according to the scoring scale (Table S2) at 2, 4, 8, 12 and 24 hours after injection of LPS. Then we analyzed the score. Centrifuge mouse blood, perform the operation according to the instructions of C-reactive protein ELISA kit, centrifuge the blood sample according to the instructions, take the supernatant, and dilute it five times. Each sample is made three multiple holes, and finally converted into concentration according to absorbance, multiplied by five, and analyzed. Another 40ul of blood was taken from each mouse and divided into two tubes of 20ul each. Blood count results were measured using SYSMEX-XN-1000 with 20ul of whole blood in 140ul of diluent. The sending agency measures the number of various types of cells in the blood for analysis. After extracting RNA from mouse plasma, reverse transcription and amplification, gel electrophoresis was performed, and the trend of the results was calculated by gray value analysis. Primer sequences are listed in Table S3.

Statistical Analysis

All of the analyses were carried out using R version 4.2.1, 64-bit6 and its support package. To test the correlation between two groups of continuous parameters, the non-parametric Wilcoxon rank-sum test was used. Correlation coefficients were examined using Spearman correlation analysis. And, P < 0.05 was considered to be statistically significant. The glm package is used to build logistic regression models. The receiver operating characteristic curve (ROC) and the area under the curve (AUC) are used to evaluate the effectiveness of the model. And PCA plots using the ggplot2 package.

Results

Identification of DEGs in GSE65682

A total of 1058 DEGs were identified from the screening of whole blood-derived RNA samples from both normal controls and septic patients. The genes ABLIM1, CD96, CD160, GATA3, NMT2, PTCH1 were down-regulated (Figure 1A, green dots) and MAPK14, SRPK1, S100A8, TXN, UPP1 were up-regulated (Figure 1A, red dots). Figure 1B shows 50 up-regulated genes and 50 down-regulated genes by heatmap.

Screening of Gene Modules and Construction of Co-Expression Networks

We constructed co-expression networks and modules for normal and sepsis samples and selected the genes with the highest 25% variation for further analysis. Further, we set the scale-free $R^2$ to 0.9 and screened for soft power (soft power = 7), at which
point the co-expressed gene modules were able to be successfully identified. (Figure 2A). Subsequently, we obtained eight co-expression modules in different colours and the topological overlap matrix (TOM) heat map is shown (Figure 2B–D). We found that the blue module was most strongly associated with sepsis and that \( P < 2 \times 10^{-73} \), the module contained a total of 585 genes (Figure 2E and F).

**Functional Annotation of Sepsis-Associated Genes**

We obtained 381 genes by taking the intersection of the differentially expressed genes and the modular genes obtained by WGCNA (Figure 3A). Next, these genes were analysed by KEGG and GO for functional annotation. In the KEGG analysis (Figure 3B), the first 30 enrichment pathways were mainly Th17 cell differentiation, T cell receptor signaling pathway, PD-L1 expression and PD-1 checkpoint pathway in cancer, etc. In the GO-BP analysis (Figure 3C), the main pathway is also lymphocyte differentiation (GO:0030098). In GO-CC analysis (Figure 3C), the external side of the plasma membrane (GO:0009897) and immunological synapse (GO:0001772) were significantly enriched. The results of GO-MF (Figure 3C) are also mainly related to immunological processes. Thus, the enrichment analysis results indicated that the activation and regulation of cell death and resistance to infection may be of vital importance in sepsis.

**Expression of CRGs, Biological Function and Immune Infiltration in Patients with Sepsis**

To further explore the biological role of CRGs in the development of sepsis and in causing multi-organ damage, we investigated the expression of 19 CRGs in GSE65682. The 14 CRGs were eventually identified as differentially expressed cuproptosis genes. NFE2L2, FDX1, LIAS, LIPT1, DLAT, PDHA1, PDHB, GLS, CDKN2A and DLST gene expression levels were greatly lower in the sepsis samples than in the non-sepsis controls. And, NLRP3, SLC31A1, DLD and MTF1 expression levels were upregulated in the sepsis samples (Figure 4A–C). Furthermore, we analysed the correlation of differentially expressed CRGs, which will be useful in revealing the mutual regulation of CRGs in sepsis. The results showed that DLAT and DLD exhibited a strong synergistic effect and that NFE2L2 and SLC31A1 had a definite antagonistic effect (Figure 4D and E). The immune system plays a crucial role in the development of sepsis, so an immune infiltration analysis was performed based on the CIBERSORT algorithm for differences to detect differences in the percentages of 22 infiltrating immune cell types between sepsis samples and non-sepsis control samples.
Figure 2 Co-expression network of differentially expressed genes in sepsis. (A) The choice of the soft threshold of power. (B) Dendrogram of co-expression module clusters. And, co-expression modules were represented by different colours. (C) Representative clusters of modular autologous genes. (D) Representative heatmap. (E) Correlation analysis of module-related genes with clinical traits (normal and sepsis). Rows represent modules, columns represent clinical traits and numbers in brackets are p-values. (F) Scatter plot.
The results showed significant differences in the proportions of immune cells such as naïve B cells, B cells memory, plasma cells, T cells CD8, Monocytes and Dendritic cells resting, etc. Meanwhile, the correlation analysis showed that CRGs were associated with immune cells such as the activation of NK cells and the activation of CD4 memory in T cells, etc. (Figure 4G). These results indicate that CRGs may be crucial factors regulating the molecular and immune aggressiveness state in sepsis patients.

The effect of differentially expressed CRGs in the development of sepsis was explored by GO and KEGG analysis. It was shown in GO enrichment analysis that these genes play a major role in the acetyl-CoA metabolic process, mitochondrial matrix and oxidoreductase activity (Figure 4H). In addition, the Citrate cycle (TCA cycle), Carbon metabolism and Pyruvate metabolism were the main enrichment results in the KEGG analysis (Figure 4I).

Identification of Diagnostic Genes and Construction of Models

We constructed diagnostic models for differentiating septicemic samples from non-septic control samples based on differentially expressed CRGs. The LASSO and SVM-RFE algorithms were used in the GSE65682 dataset to screen for important differentially expressed CRGs to differentiate between sepsis patients and normal individuals. The LASSO logistic regression algorithm was applied to identify 10 sepsis-related features. The penalty parameters were tuned by 10-fold cross-validation (Figure 5A and B). Meanwhile, we filtered 11 genes(maximal accuracy =0.994, minimal RMSE =0.006) out of 14 differentially expressed CRGs as optimal signature genes in the SVM-RFE algorithm (Figure 5C). The genes screened by the two algorithms were intersected to obtain a total of 9 genes (NFE2L2, NLRP3, SLC31A1, DLD, DLAT, PDHB, MTF1, CDKN2A, DLST) as the final identified genes for subsequent analysis (Figure 5D). ROC curves were generated for each marker gene in an attempt to explore the ability of individual genes to distinguish sepsis samples from normal samples (Figure 5E). Each gene had an AUC greater than 0.6 and some had an AUC greater than 0.9 (PDHB, AUC=0.970; MTF1, AUC=0.974). In addition, using the 9 marker genes mentioned above, we constructed logistic regression models using the R package glm. Using ROC curves, logistic
Figure 4 Differential expression and function of CRGs in sepsis, immune infiltration in patients with sepsis. (A) Boxplots showed the expression of CRGs between sepsis and non-sepsis controls. (B) The heatmap shows the expression patterns of 14 differentially expressed CRGs. (C) Chromosomes locations of the 19 CRGs. (D) Correlation analysis, red represents positive correlations, green represents negative correlations, and the correlation coefficient is represented by the area of the pie chart. (E) Correlation analysis of the 14 differentially expressed CRGs is represented by chord plots. (F) Differences in immune infiltration between sepsis and non-sepsis controls were shown in boxplots. (G) Correlation analysis between 14 differentially expressed CRGs and infiltrating immune cells. Enrichment analysis of differentially expressed CRGs in sepsis. (H) GO pathway. (I) KEGG pathway. *P < 0.05, **P < 0.01, ***P < 0.001.

https://doi.org/10.2147/JIR.S452980

Journal of Inflammation Research 2024:17
regression models based on 9 markers discriminated between normal and septic samples with AUC=1.000 (Figure 5F). The final results showed that the accuracy and specificity of the logistic regression model was superior to that of individual marker genes.

Marker Genes Have Been Associated with a Variety of Pathways in the Development of Sepsis

We used a GSEA-KEGG single-gene analysis pathway to analyse the biological functions of marker genes. The enrichment pathway for the top 6 of each marker gene is shown in the figure (Figure 6A–I). Our analysis has revealed that these marker genes are associated with the cell cycle, oxidative phosphorylation, immune response (Natural killer cell-mediated cytotoxicity and Presentation and handling of antigens), multiple organelles (Ribosomes, Spliceosomes and lysosomes), anabolism of substances (DNA replication, Pyrimidine metabolism and Histidine metabolism), various diseases (Parkinson’s disease, Asthma, Primary immunodeficiency disease, Type 1 diabetes, Systemic lupus erythematosus and Autoimmune thyroid disease) and numerous cellular signaling pathways (Nod-like receptor signaling pathway, Toll-like receptor signaling pathway, Neurotrophin signaling pathway and T-cell receptor signaling pathway).

Using GSV A analysis, we examined the signaling pathways that were differentially activated between the high and low expression groups. We found that overexpression of CDKN2A-activated processes such as GLYCOSPHINGOLIPID BIOSYNTHESIS LACTO AND NEOLACTO SERIES, STARCH AND SUCROSE METABOLISM, LINOLEIC ACID METABOLISM. However, low expression of CDKN2A in sepsis may contribute to the development of this disease through ONE CARBON POOL BY FOLATE, INTESTINAL IMMUNE NETWORK FOR IGA PRODUCTION (Figure 7A). The low expression group of DLAT was enriched in BASAL TRANSCRIPTION FACTORS, VALINE
LEUCINE, AND ISOLEUCINE BIOSYNTHESIS (Figure 7B). Low expression of DLD is associated with TERPENOID BACKBONE BIOSYNTHESIS, O GLYCAN BIOSYNTHESIS and high expression is associated with PRIMARY IMMUNODEFICIENCY, TYROSINE METABOLISM (Figure 7C). Low expression of DLST in disease activates GLYCOSAMINOGLYCAN BIOSYNTHESIS KERATAN SULFATE, GLYCOSAMINOGLYCAN BIOSYNTHESIS CHONDROITIN SULFATE, but the high expression is only strongly correlated with FOLATE BIOSYNTHESIS (Figure 7D). Notably, high expression of MTF1 is associated with diseases such as AUTOIMMUNE THYROID DISEASE, TYPE I DIABETES MELLITUS and GRAFT VERSUS HOST DISEASE that has a relevant association with sepsis. However, at low MTF1 expression, it is mainly associated with STARCH AND SUCROSE METABOLISM (Figure 7E). The low-expression group of NFE2L2 was predominantly enriched in GLYCOSAMINOGLYCAN BIOSYNTHESIS CHONDROITIN SULFATE and enriched in DRUG METABOLISM CYTOCHROME P450 in the high-expression group (Figure 7F). The highly expressed NLRP3 plays a central role in the anabolism of substances (DNA REPLICATION, MISMATCH REPAIR, TYROSINE METABOLISM, BIOSYNTHESIS OF UNSATURATED FATTY ACIDS) (Figure 7G). Interestingly, low expression of PDHB is mainly enriched with amino acid biosynthesis and metabolisms such as VALINE LEUCINE AND ISOLEUCINE BIOSYNTHESIS, BETA-ALANINE METABOLISM and OTHER GLYCAN DEGRADATION (Figure 7H). In addition, SLC31A1 expression was upregulated in sepsis samples, closely related to PRIMARY IMMUNODEFICIENCY, GLYCOSAMINOGLYCAN BIOSYNTHESIS KERATAN SULFATE (Figure 7I).
Multiple Interaction Networks Based on Marker Genes

In order to screen for potential drugs for sepsis treatment, we filtered through the DGIdb database based on marker genes. The results are visualised using Cytoscape software as follows (Figure 8A). In the DGIdb database, five of the marker genes (NFE2L2, NLRP3, SLC31A1, MTF1, CDKN2A) had interacting drugs. Unfortunately, the remaining marker genes were not queried for information on interacting drugs. The miRanda, miRDB and TargetScan databases are used to screen for miRNAs that interact with mRNAs of marker genes. And we use Cytoscape software for visualisation (Figure 8B). The miRNAs interacting with the nine marker genes are presented. The details of the network are shown in “gene-miRNA.txt”. In addition, we constructed Protein-Protein Interaction Networks based on the GeneMANIA database by using marker genes (Figure 8C). In the generated PPI network, proteins are presented as nodes, with differently coloured lines between the nodes representing specific meanings.

Expression of Marker Genes in the Validation Set

To make our experimental results more realistic and credible, we confirmed marker genes expression in the GSE57065 dataset. The results demonstrated that the expression trends for PDHB, DLAT, DLD, DLST, MTF1, SLC31A1 and NLRP3 were consistent with the data from GSE65682 and were statistically significant (Figure 9A–G). The expression trend of CDKN2A was not statistically significant (Figure 9H). And the expression trend of NFE2L2 was opposite to that of the GSE65682 dataset (Figure 9I).
Figure 8 Interaction networks. (A) Prediction of marker gene-targeted drugs. Red nodes represent genes up-regulated in sepsis samples, green nodes represent genes down-regulated in sepsis samples and blue nodes represent drugs that have interactions with the target gene. (B) The gene-miRNA action network. The red nodes represent marker genes and the blue nodes represent miRNAs that interact with the genes. (C) PPI network. Nodes represent proteins. The connecting lines represent network types and the colour of the node represents the enriched function.
The Differences in Expression of CRGs, Immune Infiltration Characteristics and Functions Between the Two Subtypes

The results in the above validation experiments differed somewhat from the analytical findings. We speculate that this may be related to the category of gene chips or the heterogeneity of sepsis. Therefore, based on the expression profiles of 14 differentially expressed CRGs, we attempted to explore the heterogeneity of sepsis by clustering sepsis samples using a consensus clustering method. Based on the combined scoring of the clusters, the optimum number of subtypes was determined to be 2 and was shown in the consensus matrix (Figure 10A), and representative cumulative distribution function (CDF) curves, triangular CDF area plots, consensus clustering scores, and range plots (Figure 10B-E). We named the two subtypes C1 and C2, respectively, and used the PCA plot to represent the distribution of the two subtypes (Figure 10F).

We first analyzed the expression of CRGs between the two subtypes, showing that the expression of NLRP3, SLC31A1 and MTF1 was higher in the C1 subtype and that NFE2L2, FDX1, LIISA, LIPT1, DLD, DLAT, PDHB, GLS and CDKN2A were upregulated in the C2 subtype (Figure 10G-H). In addition, immunological infiltration analysis between the two subtypes showed that the C1 subtype exhibited higher B-cell memory, Plasma cells, Macrophages M0 and Macrophages M1 (Figure 10I). Moreover, GSVA analysis between the two subtypes showed that the C1 subtype was

Figure 9 Marker genes’ expression in the GSE57065 dataset. PDHB(A), DLAT(B), DLD(C), DLST(D), MTF1(E), SLC31A1(F), NLRP3(G), CDKN2A(H) and NFE2L2(I).
Figure 10. The differences in expression of CRGs, immune infiltration characteristics and functions between the two subtypes. (A) Consensus clustering matrix when $k=2$. (B) The cumulative distribution function (CDF) curve. (C) CDF delta area curves. (D) Consensus clustering and trajectory plot scores for $k = 2$–9. (E) Pattern classification trajectories for $k=2$–9. (F) The PCA diagram represents the effective division of sepsis samples into two subtypes (C1 and C2). (G) Heatmap is used to represent the expression of CRGs between the two subtypes. (H) The expression of CRGs between the two subtypes is represented by boxplots. (I) The differences were shown in immune infiltrating between the two subtypes. (J) Ranked by GSVA t value, differences in biological functions between C1 and C2 subtypes. Blue represents the C1 subtype and red represents the C2 subtype. *P < 0.05, **P < 0.01, ***P < 0.001.
predominantly enriched in SPLICEOSOME, CYSTEINE AND METHIONINE METABOLISM, VALINE LEUCINE AND ISOLEUCINE DEGRADATION and the C2 subtype was enriched in CYTOKINE CYTOKINE RECEPTOR INTERACTION, VASCULAR SMOOTH MUSCLE CONTRACTION and CHEMOKINE SIGNALING PATHWAY (Figure 10J). In conclusion, we found that the C1 subtype may be more relevant to cellular metabolism and the C2 subtype has some relevance to immune molecules.

**Animal Experiments Verified the Expression of CRGs**

We constructed sepsis models by intraperitoneal injection of LPS in the mice grouped equally according to body weight (Figure 11A). The results showed that c-reactive protein levels were elevated in the serum of sepsis mice, and the

![Animal experiments verified the expression of CRGs.](image-url)

Figure 11 Animal experiments verified the expression of CRGs. (A) Comparison of body weights of control and sepsis group mice. (B) Comparison of c-reactive protein in serum of control and sepsis mice. (C) Curves of change in observational index scores over time in control and sepsis group mice. (D) Blood cell ratio chart. (E) Comparative analysis of differences in blood cells. (F) Differential analysis of CRGs expression in serum of control and sepsis mice. *P < 0.05, **P < 0.01, ***P < 0.001.
observational index scores increased with time (Figure 11B–C). In addition, peripheral blood cell analysis showed that sepsis mice had significantly higher levels of neutrophils and lower levels of monocytes (Figure 11D–E). In an examination of the expression of CRGs in serum samples, we found that the expression of Dld, Dlst, Nlrp3, Nfe2l2, and Slc31a1 was up-regulated in sepsis mice (Figure 11F).

Discussion

Sepsis has a high morbidity and mortality rate, with multiple organ dysfunction, and can even develop into septic shock. A growing number of studies have concluded that disturbances in substance and energy metabolism are highly relevant to the progression of sepsis. And cuproptosis is closely linked to the TCA cycle. Therefore, it is reasonable to assume that cuproptosis may play a vital role in the disease progression of sepsis, multi-organ damage and other processes. We now note that former studies did not examine the relevance between CRG and sepsis’s advancement. Strikingly, many CRGs were differentially expressed between septic and normal samples, and most of these genes were associated with immune function and may influence sepsis typing. This suggests a potential role for cuproptosis in sepsis.

In recent years, our understanding of sepsis which is a systemic, complex and systemic disease is still incomplete. The key biological processes underlying disease progression need to be further explored. Therefore, we took intersections of differentially expressed genes between sepsis patients and healthy individuals and genes screened by the WGCNMA method as disease-related genes. This not only helps to pick out latent diagnostic biomarkers and therapeutic targets of sepsis, but also helps to further clarify the molecular mechanism of sepsis pathogenesis. We picked out 381 genes as relevant for sepsis and analysed their function. Cytokines are important in organ damage due to sepsis. The use of appropriate inhibitors can provide some protection during the acute phase of sepsis. Platelet MHC class I in sepsis inhibits CD8+ T-cells and exacerbates sepsis injury. The results of these studies are consistent with our functional enrichment results. In addition, we have revealed a number of other cellular signaling pathways that may be involved in the development of sepsis and may be able to provide new ideas for mitigating organ damage due to sepsis.

Moreover, we observed differences in the expression of CRGs, kind and abundance of immune cell population infiltration between the two groups. This shows that immune cells take significant part in sepsis and that CRGs may play a regulatory role in immune cells. It has been shown that depletion of B-cell memory contributes to sepsis-induced immunosuppression and exacerbates disease damage and that reduced circulating B-cell and IgM levels are associated with reduced sepsis survival. During sepsis, the systemic inflammatory reaction induced by circulating intrinsic immune cells also affects tissue-resident immune cells, compromising vital organ function. T-cell CD8 and T-cell CD4 dysfunction and decreased cell counts in sepsis patients. In our study, the memory of B cells, plasma cells, M0 macrophages, and monocytes was significantly higher in sepsis samples than in controls. The relative frequencies of CD8 T cells and CD4 T cells were significantly reduced. Although T cells gamma delta have been shown to be a specialized antigen-presenting cell, the antigen-presenting function of T cells gamma delta is severely impaired in patients with sepsis. In our study, there was more T cell gamma delta in sepsis samples than in normal samples. We speculate that this may be related to a compensatory increase following impaired antigen presentation function. Notably, in a relevance analysis that between CRGs and immune cell infiltration, we found a significant negative correlation between Nfe2l2 (Nrf2) and infiltration of neutrophils. This is consistent with the conclusion that inhibition of the NLRP3-GSDMD pathway using melatonin activation of the Nrf2/HO-1 signaling axis alleviated acute lung injury in sepsis characterised by neutrophil infiltration. In our study, numerous results have shown that the expression of CRGs correlates significantly with the infiltration of multiple immune cells. Thus CRGs may play a key role in the regulation of the immune microenvironment in sepsis.

Furthermore, we used two machine learning algorithms to identify 9 marker genes (Nfe2l2, Nlrp3, Slc31a1, Dld, Dlst, Pdhb, Mtf1, Cdkn2a, Dlst) and construct a diagnostic model. Nfe2l2 is a key regulator of cellular redox homeostasis. It plays a critical role in the regulation of mitochondrial function. Clinical studies have found that the Nfe2l2’s mRNA levels have a negative relation to the inflammation of the lungs and the severity of the disease in patients with sepsis. And Nlrp3 is a cytosolic innate immune signaling receptor that promotes inflammatory vesicle formation and induces inflammatory cellular scavenging. It is also associated with auto-inflammation and autoimmune diseases. In addition, the Nlrp3 inflammasome activates interleukin 1β (IL-1β). Then interleukin 1β (IL-1β) damages cardiac ability in patients with sepsis. Slc31a1 is a copper transporter protein and upregulation of ATF3/SPI1/
SLC31A1 signaling. Copper homeostasis and promote cuproptosis would be disrupted by SLC31A1. FDXI, DLD, DLAT and PDHB are the lipoic acid way’s core genes and their low expression in ovarian tissue would promote cuproptosis. According to the report, MTF1 is a Cys2-His2 zinc finger protein that is closely related to intracellular metal metabolism and metal homeostasis and is involved in the regulation of cuproptosis. CDKN2A is a cell cycle control gene and its absence disrupts glucose homeostasis. DLST is the E2 component of the alpha-ketoglutarate (aKg) dehydrogenase complex, which controls the entry of glutamine into the tricarboxylic acid cycle (TCA) for oxidative decarboxylation. In our study, the expression trends of genes for which the function of sepsis development has been clarified (NFE2L2 and NLRP3) are consistent with the above experimental findings. However, the effect of sepsis’ other genes (SLC31A1, FDXI, DLD, DLAT, PDHB, MTF1, CDKN2A and DLST) has not yet been revealed and further studies are needed. Nonetheless, our findings support that they promote cuproptosis in sepsis samples and that cuproptosis may be a cause of multi-organ damage. The diagnostic model we have developed is able to accurately identify sepsis in the dataset and provides new insights into the diagnosis of sepsis. Next, our marker gene-based ssGSEA and GSVA have a significant meaning for the molecular mechanisms of sepsis and cuproptosis.

In order to search for drugs and latent biomarkers for the therapy of sepsis, we analysed the gene-drug network, the mRNA-miRNA network, and the PPI network. Some of these drugs have been shown to be effective in the treatment of sepsis such as ETODOLAC, KETOCONAZOLE and TRAMETINIB. According to the report, miRNA dysregulation plays a role in sepsis. It has been demonstrated that exosomal miR-30d-5p of neutrophils induces M1 macrophage polarization and triggers macrophage scorching in sepsis-associated acute lung injury. And, exosomal miR-1-3p in sepsis leads to endothelial cell dysfunction by targeting SERP1. Surprisingly, these studies are consistent with the results predicted by our mRNA-miRNA network. Therefore, we can prospectively study selected drugs and miRNAs to explore whether they can have some effect on the treatment of sepsis. Moreover, the PPI network provides us with additional targets for action that can be explored. Interestingly, we found a loss of differential significance for CDKN2A and NLRP3 in the validation set data. NFE2L2 showed the opposite trend. The expression trends of other genes were consistent with the original results. We believe this may be due to the different patient sources of the two transcriptomic data, the difference in assay methods and the staging of sepsis. Therefore, We chose to explore the heterogeneity of sepsis, dividing it into two subtypes (C1 and C2) based on differentially expressed CRGs. And, we found that the C1 subtype may be more relevant to cellular metabolism and the C2 subtype has some relevance to immune molecules. At present, studies have revealed the relationship between sepsis and glutamine metabolism, and this will affect the tolerance of sepsis. In addition, the activation of some immune molecules will aggravate the damage of sepsis. However, the relationship between the two subtypes and the prognosis of sepsis still needs to be explored in more sophisticated experiments. This may help inform the implementation of personalised and precise treatment for different subtypes of sepsis, and hopefully improve the cure rate of the disease.

Finally, there are some limitations to our experiment. Due to the limitation of the number of experimental samples and the method of sepsis modelling, the expression of the genes in mice may be somewhat different in humans.

Conclusions
In conclusion, our results suggest that cuproptosis is critical in the progression of sepsis and may able to regulate the immune microenvironment of sepsis. We have screened some potential drugs and targets, which may have a positive effect on the treatment of sepsis.

Availability of Data and Materials
The dataset analyzed in the current study is available on GEO database. In addition, according to China’s “Measures for Ethical Review of Life Science and Medical Research Involving Humans”, our research involving humans meets the relevant requirements for ethical approval exemption. The original contributions presented in the study are included in the article/ Supplementary Material. Further inquiries can be directed to the corresponding author.

Animal and Ethical Matters
The study was approved by the Animal Ethics Committee of Wenzhou Medical University (Ethics approval number: wydw2023-0561). The research followed Laboratory animals—Guideline ethical review of animal welfare (GB/T 35892-2018). In addition,
all operations on laboratory animals are carried out in accordance with the Measures for the Management of Laboratory Animals of Zhejiang Province (Zhejiang Provincial Government Order No. 263).

**Acknowledgments**

Yuanfeng Wang, Xu Qiu and Jiao Liu are co-first authors for this study. We appreciate the National Natural Science Foundation of China and the National Innovative Training Program for college students. We acknowledge the editors and reviewers for their helpful comments on this paper.

**Author Contributions**

YF wrote and submit the manuscript. QX, LJ collected and analyzed data. JL, JY, XY made key changes to the manuscript and drew pictures. XD and SG provide ideas and participated in the coordination of the study and reviewed the manuscript. All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

**Funding**

This work was funded by the National Natural Science Foundation of China (82103774). In addition, this work was also funded by the National Innovative Training Program for college students (202310343013).

**Disclosure**

All authors declare that they have no conflict of interest.

**References**


