Cellular and Molecular Biology
E-ISSN : 1165-158X / P-ISSN : 0145-5680
www.cellmolbiol.org

Identification of novel immune infiltration-related biomarkers of sepsis based on bioinformatics analysis

Hua Li[1#*], Le Yang[2#]

[1] Department of Critical Care Medicine, Songgang People's Hospital, Baoan District, Shenzhen, China
[2] Department of Pediatric Intensive Care Medicine, Children's Medical Center, Hunan Provincial People's Hospital, Changsha, China

[#] Hua Li and Le Yang contributed equally to this work.

ARTICLE INFO

Original paper

Article history:
Received: June 19, 2023
Accepted: September 17, 2023
Published: November 30, 2023

Keywords:

Sepsis; Bioinformatics analysis; Immune infiltration; Biomarker; Lymphocyte; Diagnosis

ABSTRACT

This study aimed to analyze the gene expression characteristics of sepsis and search for potential biomarkers involved in the pathogenesis. The data on sepsis were obtained from the GEO database according to the keyword "sepsis". The CIBERSORT algorithm was applied to determine the immune cell. WGCNA package were applied to build a weighted gene network. Then, a topological overlap matrix was created and dynamic hybrid cutting was applied to categorizing the genes with identical expression patterns. Component analysis of each module was implemented according to module eigengenes. In order to detect the important modules, the connections among the immune infiltration of Mφ and the modules were computed by Pearson's tests. PPI network was made using the STRING database and cytoHubba was applied to find hub genes. A total of 760 sepsis samples as well as 42 healthy control samples were involved. A total of 14 gene modules were generated. Thereinto, the correlations of the yellow (includes 168 hub genes) and blue (includes 526 hub genes) modules with Mφ0 were 0.39 and -0.42, while with Mφ1 were 0.3 and -0.4. 916 up-regulated and 454 down-regulated DEGs were found in the sepsis group. 451 intersected genes were obtained after the intersecting of DEGs and the hub genes from blue and yellow modules. Subsequent GO enrichment analysis suggested that 451 overlapping genes were enriched in "T cell activation", "lymphocyte differentiation" and "T cell differentiation" for biological process. Besides, KEGG enrichment analysis showed that "Human T-cell leukemia virus 1 infection" and "Th17 cell differentiation" were the most enriched pathways. In PPI network, UTP6, RRS1, RRP1B, DDX18, and DDX24 were identified as hub genes. ROC analysis showed the AUC values of these genes were all greater than 0.95. UTP6, RRS1, RRP1B, DDX18, and DDX24 participate in the pathogenetic process of sepsis through regulating the activation and differentiation of lymphocytes. Besides, these five genes could be used for diagnosing sepsis.

Doi: http://dx.doi.org/10.14715/cmb/2023.69.12.33

Introduction

Sepsis, which can cause devastating damage to multiple organs and systems, is a systemic inflammatory syndrome caused by the invasion of bacteria and other pathogenic microorganisms (1). It is characterized by high morbidity and mortality (1,2). About 20 million cases occur per year and can cause more than 5 million deaths, with a mortality rate of about 26 percent (3). Global epidemiological data from the Global Intensive Care study showed that among patients who developed sepsis on admission or while in the ICU, the mortality rate was 25.8% and the hospital mortality rate was 35.3% (4,5). Death rates from sepsis appear to have declined in recent decades, but incidence rates are still rising (5).

As one of the top ten causes of death, sepsis has caused great public health and a huge economic burden to society (6). Unfortunately, treatment options for sepsis are still limited to supportive care, such as control of the primary disease, timely and adequate use of antibiotics, and resuscitation of dysfunctional organs (6, 7). At present, no specific drug has been found to cure sepsis, which may be due to its complex pathogenesis is still unclear (8). Therefore, exploring the pathogenesis of sepsis, identifying key genes, and searching for effective biological targets have become the top priority in the studies of sepsis.

In recent years, the popularity of biomarker-targeted therapy for sepsis has gradually increased, however, no veritably effective targets have been found (9, 10). Previous researches have shown that a series of biomarkers, such as procalcitonin, c-reactive protein, sTREM-1, and suPAR, can play a positive role in clinical diagnosis and treatment, but their specificity is low and can only be used to distinguish sepsis from other inflammatory diseases (11, 12). Because sepsis does not possess a gold diagnostic standard, it is difficult to realize early clinical diagnosis. Thus, the application of biomarkers to diagnose sepsis can facilitate individualized accurate treatment.

In addition, systemic immune reaction plays an important role in the mechanism of severe sepsis (13). The immune response to sepsis can be separated into two stages: early proinflammatory stage and terminal anti-inflammatory stage (13).

In the early stages, proinflammatory processes domi-

* Corresponding author. Email: 18127014661@163.com

nate the immune response. Infiltration of inflammatory cells such as macrophages, dendritic cells, neutrophils, and T cells is an imperative reason for severe immune disorders in the early inflammatory stage of sepsis (14).

Immunosuppression in patients with sepsis refers to different cell types and characteristics, including increased apoptosis of immune cells, T-cell failure, epigenetic reprogramming of cells, and decreased expression of cell surface molecules (15). There is evidence that sepsis-induced immunosuppression promotes bacterial growth (15). Besides, sepsis immunosuppression boosts the possibility of opportunistic infections and organ dysfunction, as a result, the prognosis is poor (16). Since previous researches have indicated that immune cell infiltration is vital for the development of sepsis, the detection of prognostic markers and the analysis of immune cell infiltration pattern are of great importance for improving the prognosis of sepsis patients. Therefore, the investigation of immune-related prognostic markers particularly is a substantial focus of sepsis research.

Materials and Methods

Attainment and processing of gene data

The gene data of sepsis were downloaded from the GEO database according to the keyword "sepsis". The inclusion criteria were as follows: (1) availability of the patients with sepsis or Healthy from the joint in the datasets; (2) more than six specimens in the dataset. One qualified dataset with accession number GSE65682 was selected (array data, GPL13667). A total of 802 samples, including 760 sepsis samples and 42 healthy control samples, were involved in the data. The information and data of these data were all downloaded from public databases, as a result, patient consent and ethics committee approval were unnecessary.

Assessment of immune cell infiltration

In this study, the CIBERSORT algorithm was used to clarify the immune cell infiltration of GSE65682. The CIBERSORT algorithm is a brilliant approach compared to conventional deconvolution means for assessing infiltrating immunity because it can evaluate unspecified data as well as noise. Therefore, the CIBERSORT algorithm is an outstanding tool for computing the abundance of specialized cells within the mixed matrix.

Establishment and identification of macrophage-related hub modules using weighted gene co-expression network analysis

Different expressed gene variants for the two groups were identified. WGCNA package in R software was used to establish the weighted gene network of the top 5000 genes. Then, Pearson's correlation matrices were ascertained with the similarity matrix altered from the expression of each transcript. The similarity matrix was then altered into the adjacency matrix according to amn=|cmn|β (amn represents the adjacency between paired genes, cmn represents Pearson's correlation coefficient between the paired genes, and β represents the soft-power threshold). Next, a topological overlap matrix was built from the adjacency matrix when β=9. The value of β was also used for assessing the connectivity features in the co-expression network. Besides, we used the average linkage hierarchical clustering as well as a dynamic hybrid cutting method to construct a dendrogram of the topological overlap matrix and classify the genes into different modules, with module minimum size cutoff value being 30 and merging identical modules threshold being 0.25. Module eigengenes were used to make a component analysis of each module. In order to search and identify important modules, the associations among the infiltration levels of Mφ and the modules were assessed by Pearson's tests. The Mφ subset and modules with the maximum association coefficients were deemed as hub modules.

Data process

The limma package in R software was applied to search differentially expressed genes (DEGs) with thresholds of $|\log2FC| \geq 1$ and $P < 0.05$. In the WGCNA result, genes that significance values > 0.3 and module membership values > 0.6 were described as hub genes. To identify latent biomarkers, a Venn diagram was used to visualize the overlap of DEGs and module genes.

Enrichment analysis of GO and KEGG

The overlapping genes were used for functional analysis based on GO and KEGG database using the R entitled org.Hs.eg.db, clusterProfiler, org.Hs.eg.db, and enrichplot package in R software. Therefore, GO annotation contains three parts: biological process (BP), cellular component (CC), and molecular function (MF). The FDR threshold was set as <0.05.

Protein-protein interaction (PPI) network analysis

The overlapping genes were sent to the STRING website to generate the interaction diagram of overlapping genes. Then, the results were imported into Cytoscape (v.3.7.1) to obtain hub genes. Next, in order to identify novel diagnostic markers, ROC curves were made using the hub genes in GraphPad Prism 6.01 (La Jolla, CA, USA) to explore their diagnostic value.

Statistical Analysis

All bioinformatics analyses in this study were conducted using R software (v.4.0.2). GraphPad Prism 6.01 was used for the ROC analyses. Student T-test was performed for continuous variables. Then, P values were adjusted on the basis of Benjamini and Hochberg's test. Statistically, significance was determined so long as P values < 0.05.

Results

Assessment of the immune infiltration

The infiltration pattern of 22 diverse kinds of immune cells was detected using the CIBERSORT algorithm (Figure 1). The complicated pathogenesis of sepsis can bring remarkable alterations in the function of macrophages including phenotype reprogramming, macrophage overactivation, and changes in signaling pathways. As a result, three different types of macrophages were used for further WGCNA analysis.

Identification of hub modules related to macrophages

A total of the top 5000 genes were sent to build a co-expression gene network with the aid of the WGCNA package. A scale-independent topological network (soft-thresholding power nine scale-free R^2= 0.9) as well as a mean

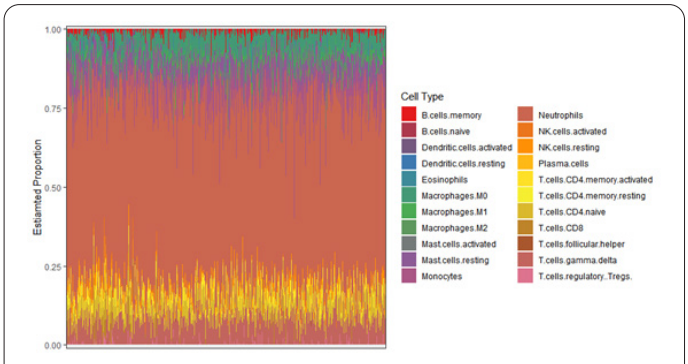

Figure 1. The abundance of infiltrated immune cells in GSE65682.

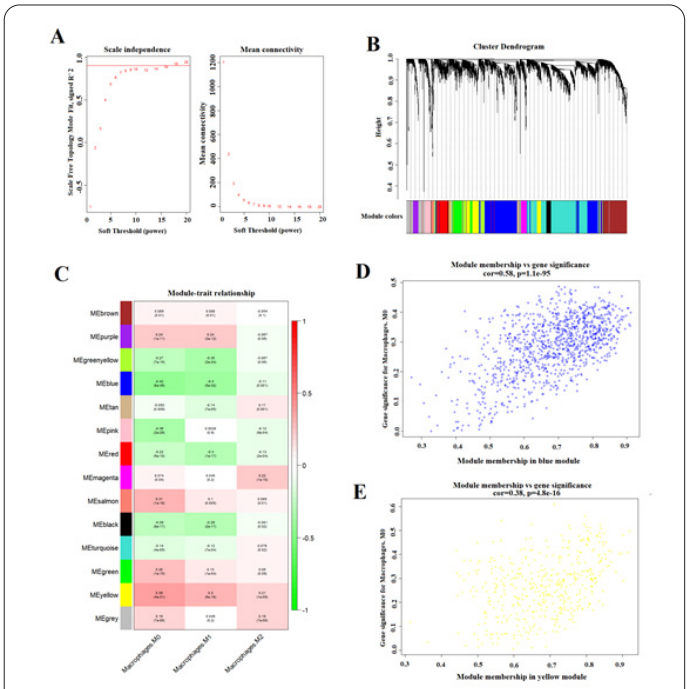

Figure 2. Identification of macrophage-related hub modules. (A) The gene mean connectivity with the scale independence index ranges from 1 to 20 (β=9). (B) Establishment of gene modules. (C) The associations of each module eigengenes with Mφ. The numbers in each box represented the correlation coefficients, along with the p values shown in brackets. (D-E) Genes in yellow and blue modules. Each dot shows a gene. Genes with module membership values >0.6 and gene significances >0.3 were deemed as module genes.

connectivity network were generated (Figure 2A). Then, we establish dynamic hybrid cutting in order to build a hierarchical clustering tree through dividing the dendrogram at relevant transition points. That is, each gene was set as the leaves, while a bunch of genes with similar expression patterns were set as branches (Figure 2B). Therefore, the above-mentioned branches that included similarly expressed genes were deemed as gene modules. Finally, a total of 14 gene modules were created after blending similar modules (Figure 2C). Among these 14 modules, the R^2 of the yellow module for Mφ0 was 0.39, while the blue module was -0.42. In addition, the R^2 of the yellow module for Mφ1 was 0.3, while the blue module was -0.4.

Thus, the yellow and blue modules were identified as the hub modules (Figure 2C). According to module membership values > 0.6 and gene significance values > 0.3, the blue hub module possessed 526 module genes (Figure 2D) and the yellow hub module possessed 168 module genes (Figure 2E).

Macrophage-related DEGs in sepsis samples

A total of 760 sepsis and 42 control samples were involved in this study. Through differential expression analysis of the dataset, a total of 1370 DEGs including 916 up-regulated and 454 down-regulated genes were differentially expressed in the disease group. The expressions of the DEGs were pictured on a volcano map (Figure 3A). The top 10 up-regulated as well as down-regulated genes were shown in the heatmap (Figure 3B).

For analysis of hub genes, the DEGs and the genes from the blue and yellow modules were intersected, as shown in Fig 3C. The intersected 451 genes were analyzed in the next step.

Gene function analysis of GO and KEGG

451 overlapping genes were utilized to implement GO and KEGG enrichment analysis. For the GO analysis, these genes were enriched in “T cell activation”, “lymphocyte differentiation” as well as “T cell differentiation” when considering BP (Figure 4A). The top two enriched items of CC were “cytoplasmic vesicle lumen” and “vesicle lumen” (Figure 4A). For MF, the top two enriched items were “immune receptor activity” as well as “MHC protein complex binding” (Figure 4A). KEGG enrichment analysis demonstrated that the top-2 vital pathways were “Human T-cell leukemia virus 1 infection” as well as “Th17 cell differentiation” (Figure 4B).

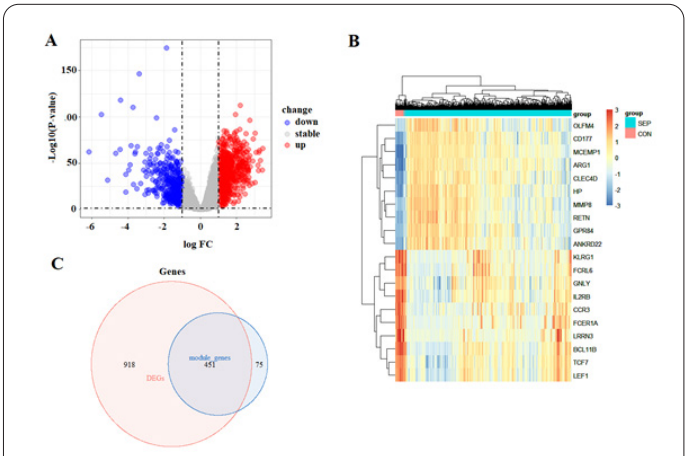

Figure 3. Screening of macrophage-related DEGs. (A)Volcano plots exhibited DEGs in sepsis and control samples. (B) Heatmap exhibited the expressions of the Top 10 up-regulated as well as down-regulated DEGs in sepsis and control. (C)The Venn diagrams of overlapped genes from DEGs and module genes. A total of 451 overlapped mRNAs are identified.

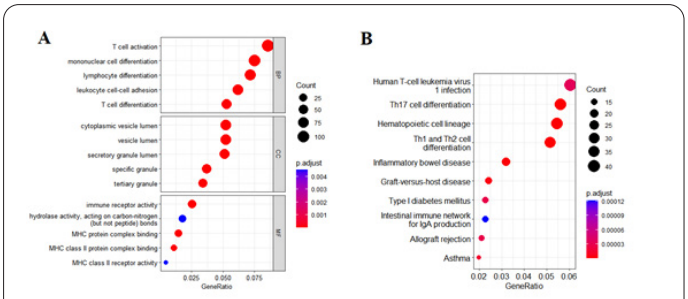

Figure 4. Enrichment Analysis of GO and KEGG. (A)The result of GO enrichment analyses for the overlap genes. (B) The result of KEGG pathway enrichment analyses for the overlap genes.

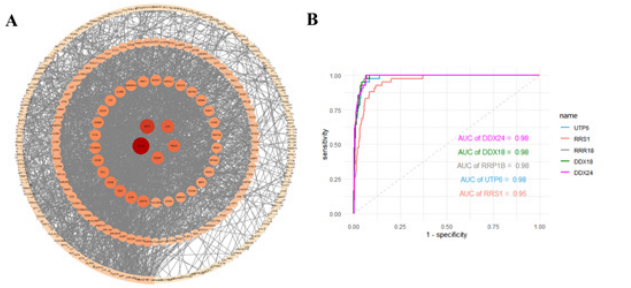

Figure 5. Identification of hub genes with diagnostic values for sepsis. (A)Visualization of the interaction network of genes by Cytoscape. (B)ROC curves of DDX24, DDX18, RRP1B, UTP6 and RRS1 for diagnosing sepsis.

Identification of biomarkers for sepsis

The PPI network according to 451 overlapping genes was established on the basis of the STRING database. The original network (Figure 5A) consisted of 385 middle nodes and 2248 edge nodes. Based on cytoHubba, five genes (UTP6, RRS1, RRP1B, DDX18, and DDX24) were then identified in this cluster and were considered as hub genes.

To further confirm the diagnostic value of these hub genes, ROC curves were drawn (Figure 5B). The AUC values of these hub genes were all greater than 0.95, of which DDX24 (AUC =0.98) had the highest AUC value and RRS1 (AUC =0.95) had the minimum AUC value.

Discussion

Currently, various studies increasingly suggest that finding biomarkers may facilitate the personalized treatment for sepsis, for it gives patients access to tailored treatment on the basis of their particular characteristics. The speedy prosperity of bioinformatics has offered a series of novel analytical tools, which can supply technical support for identifying biomarkers. These bioinformatic tools, such as CIBERSORT, WGCNA, limma, clusterProfiler, and enrichplot, have been extensively applied in the study of sepsis, particularly playing significant roles in studying molecular level-related pathway analysis. Therefore, in this study, array data from whole blood samples were obtained from GSE65682, and 22 distinct cell types of immune infiltration in 802 samples were analyzed.

Monocytes/macrophages are one of the most vital participants in the pathogenesis of sepsis. It can differentiate into various functional phenotypes, after being induced by pathogens as well as cytokines, and carry out different functions such as phagocytosing pathogenic microorganisms and producing cytokine and chemokine. The complicated pathogenesis of sepsis can bring remarkable alterations in the function of macrophages including phenotype reprogramming, macrophage overactivation, and changes in signaling pathways. These alterations further confound the pathophysiological status of sepsis and cause the intensively decreased curative effects of traditional therapy.

Therefore, three distinct kinds of macrophages in sepsis samples were used for further WGCNA analysis. Next, a total of the top 5000 genes were sent to the established gene co-expression network. Finally, a total of 14 modules were obtained after merging parallel modules. Among these 14 modules, the R2 of the yellow module for Mφ0 was 0.39, while the blue module was -0.42. In addition, the R2 of the yellow module for Mφ1 was 0.3, while the blue module was -0.4. Therefore, the yellow and blue modules that have the highest connectivity with Mφ were identified as hub modules. According to module membership values > 0.6 and gene significance values > 0.3, the blue hub module possessed 526 hub genes and the yellow hub module possessed 168 hub genes.

Through differential expression analysis of the dataset, 916 up-regulated as well as 454 down-regulated DEGs were found in the sepsis group. Then, the DEGs and the hub genes from the blue and yellow modules were intersected, and 451 intersected genes were obtained. Subsequent GO enrichment analysis suggested that 451 overlapping genes were enriched in “T cell activation”, “lymphocyte differentiation” as well as “T cell differentiation” for biological processes. Besides, KEGG enrichment analysis showed that “Human T-cell leukemia virus 1 infection” and “Th17 cell differentiation” were the most enriched pathways.

Sepsis causes a series of changes in both the presentation and function of immune cells, which promote the pathophysiology of sepsis and lead to immunoparalysis (16). Sepsis is initially featured by leukocytosis for 2–4 days, with prominent increases in monocyte and neutrophil populations, which is followed instantly by the condition of lymphopenia (17, 18). Lymphocyte populations are exceptionally sensitive to apoptosis, and the amount of B cells and T cells are distinctly reduced after sepsis onset (19, 20). It would cause increased mortality if the cell numbers of leukocytosis or lymphopenia can not be normalized during the stages (20-28). This study suggested that the stress induced by sepsis mainly influences the macrophage function related to T cell activation and differentiation, and lymphocyte differentiation. That is, the 451 overlapping genes may be closely related to the prognosis of sepsis.

To further reveal the relationship among these 451 overlapping genes, the PPI network was performed on the STRING website. The original network consisted of 385 middle nodes and 2248 edge nodes. Using the cytoHubba, five genes (UTP6, RRS1, RRP1B, DDX18, and DDX24) were defined as hub genes. In addition, ROC curves suggested that these genes all can be used for diagnosing sepsis with AUC values greater than 0.95. To our knowledge, these five genes were first reported as the biomarkers of sepsis in this study. Chao Xu et al. (11) found that MMP9 as well as C3AR1 were significantly connected with the prognosis of sepsis through regulating immune infiltration. Ye Chen et al. (9) found that CD81 could play important roles in both immunological and pathological progressions of sepsis. Therefore, this study enormously added to previous studies by finding UTP6, RRS1, RRP1B, DDX18, and DDX24 participate in the pathogenetic process of sepsis, regulating the activation and differentiation of lymphocytes, and could be used for diagnosing sepsis.

Conclusion

This study enormously added to previous studies by finding that UTP6, RRS1, RRP1B, DDX18, and DDX24 participate in the pathogenetic process of sepsis through regulating the activation and differentiation of lymphocytes. Besides, these five genes could be used for diagnosing sepsis.

Conflict of Interests

The authors declared no conflict of interest.

References

1. Evans L, Rhodes A, Alhazzani W, et al. Executive Summary: Surviving Sepsis Campaign: International Guidelines for the Management of Sepsis and Septic Shock 2021. Crit Care Med 2021; 49(11): 1974-1982.
2. Sha Y, Wang J, Wang Z, Zhang X, Chen J. Analysis of the effect of a new type of nano-drug carrier preparation on the pathological changes in lung and inducible nitric-oxide synthase expression in severe sepsis. Cell Mol Biol 2022; 68(8): 1-7.
3. Feng Q, Zhang Y, Xu C, et al. Metabolite Study of Cecal Ligation and Puncture-Induced Sepsis Rats Based on Macro Metabolomics. J Biomed Nanotechnol 2022; 18(11): 2585-2591.
4. Gavelli F, Castello LM, Avanzi GC. Management of sepsis and septic shock in the emergency department. Intern Emerg Med 2021; 16(6): 1649-1661.
5. Taj M, Brenner M, Sulaiman Z, Pandian V. Sepsis protocols to reduce mortality in resource-restricted settings: A systematic review. Intens Crit Care Nur 2022; 72: 103255.
6. Liu D, Huang SY, Sun JH, et al. Sepsis-induced immunosuppression: mechanisms, diagnosis and current treatment options. Military Med Res 2022; 9(1): 56.
7. Burgunder L, Heyrend C, Olson J, et al. Medication and Fluid Management of Pediatric Sepsis and Septic Shock. Pediatr Drugs 2022; 24(3): 193-205.
8. Dai W, Zheng P, Luo D, et al. LPIN1 Is a Regulatory Factor Associated With Immune Response and Inflammation in Sepsis. Front Immunol 2022; 13: 820164.
9. Chen Y, Qiu C, Cai W. Identification of key immune genes for sepsis-induced ARDS based on bioinformatics analysis. Bioengineered 2022; 13(1): 697-708.
10. Yang K, Fan M, Wang X, et al. Lactate promotes macrophage HMGB1 lactylation, acetylation, and exosomal release in polymicrobial sepsis. Cell Death Differ 2022; 29(1): 133-146.
11. Xu C, Xu J, Lu L, Tian W, Ma J, Wu M. Identification of key genes and novel immune infiltration-associated biomarkers of sepsis. Innate Immun-London 2020; 26(8): 666-682.
12. Downes KJ, Fitzgerald JC, Weiss SL. Utility of Procalcitonin as a Biomarker for Sepsis in Children. J Clin Microbiol 2020; 58(7): e01851-19.
13. Gong T, Liu Y, Tian Z, et al. Identification of immune-related endoplasmic reticulum stress genes in sepsis using bioinformatics and machine learning. Front Immunol 2022; 13: 995974.
14. Kumar V. Pulmonary Innate Immune Response Determines the Outcome of Inflammation During Pneumonia and Sepsis-Associated Acute Lung Injury. Front Immunol 2020; 11: 1722.
15. Cao C, Yu M, Chai Y. Pathological alteration and therapeutic implications of sepsis-induced immune cell apoptosis. Cell Death Dis 2019; 10(10): 782.
16. Zeng YP, Wu D, Zhuo XM, Song J. Effects of Continuous Blood Purification without Heparin on Strem-1, NSE, and IL-10 Levels in Patients with Sepsis. Cell Mol Biol 2022; 68(4): 178-187.
17. Rubio I, Osuchowski MF, Shankar-Hari M, et al. Current gaps in sepsis immunology: new opportunities for translational research. Lancet Infect Dis 2019; 19(12): e422-e436.
18. Wu J, Liu Q, Zhang X, et al. The interaction between STING and NCOA4 exacerbates lethal sepsis by orchestrating ferroptosis and inflammatory responses in macrophages. Cell Death Dis 2022; 13(7): 653.
19. Huang Z, Fu Z, Huang W, Huang K. Prognostic value of neutrophil-to-lymphocyte ratio in sepsis: A meta-analysis. Am J Emerg Med 2020; 38(3): 641-647.
20. Shamakhteh F, Ibrahim SK, Saki M et al. Investigating the antibiotic resistance pattern of Staphylococcus aureus isolated from clinical samples of patients hospitalized in the special care departments of Ahvaz teaching hospitals. Cell Mol Biomed Rep 2021;1(4):191-198. doi: https://doi.org/10.55705/cmbr.2021.419144.1182.
21. Keikha M, Shahramian I and Saravani K. Investigation of non-invasive indicators of liver fibrosis in thalassemia major patients. Cell Mol Biomed Rep 2021;1(4):181-190. doi: https://doi.org/10.55705/cmbr.2021.411709.1168.
22. Dominguez Huarcaya LR, Mohammadi MR and Dominguez Rios MF. Gut microbiota and parkinson's disease. Cell Mol Biomed Rep 2021;1(4):168-179. doi: https://doi.org/10.55705/cmbr.2021.410426.1165.
23. Saravani K, Saravani S and Dadras F. Investigating attitudinal barriers to breast cancer mammography screening among women in Zabol city. Cell Mol Biomed Rep 2021;1(4):158-167. doi: https://doi.org/10.55705/cmbr.2023.411556.1167.
24. Omidi AH, Sabati H, Amini S et al. Staphylococcus aureus in the environment of healthcare centers. Cell Mol Biomed Rep 2021;1(4):147-157. doi: https://doi.org/10.55705/cmbr.2021.403541.1156
25. Soumya P, Priya TS, Qureshi HA et al. A retrospective observational study to assess the cirrhosis risk in alcoholic patients. Cell Mol Biomed Rep 2021;1(3):129-137. doi: Https://doi.org/10.55705/cmbr.2021.361362.1063.
26. Rezaie-Kahkhaie L, Saravani K, Rezaie-Keikhaie K et al. Prevalence of hepatitis B in HIV-positive patients in Zabol. Cell Mol Biomed Rep 2021;1(3):105-112. doi: Https://doi.org/10.55705/cmbr.2021.356667.1058.
27. Martin MD, Badovinac VP, Griffith TS. CD4 T Cell Responses and the Sepsis-Induced Immunoparalysis State. Front Immunol 2020; 11: 1364.
28. Song C, Xu J, Gao C, Zhang W, Fang X, Shang Y. Nanomaterials targeting macrophages in sepsis: A promising approach for sepsis management. Front Immunol 2022; 13: 1026173.