



A host immune-related LncRNA and mRNA signature as a discriminant classifier for bacterial from non-bacterial sepsis in children

Chunxia Wang^{a,b,c,1,2,*}, Ting Sun^{a,2}, Yiping Zhou^{a,b,c}, Tiantian Liu^{a,b}, Shuyun Feng^{a,b}, Xi Xiong^{a,b,c}, Jiao Fan^d, Qiming Liang^{e,f}, Yun Cui^{a,b,c,**}, Yucai Zhang^{a,b,c,***}

^a Department of Critical Care Medicine, Shanghai Children's Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai, 200062, China

^b Institute of Pediatric Infection, Immunity, and Critical Care Medicine, Shanghai Jiao Tong University School of Medicine, Shanghai, 200062, China

^c Institute of Pediatric Critical Care, Shanghai Jiao Tong University, Shanghai, 200062, China

^d Institute of Geriatrics, National Clinical Research Center of Geriatrics Disease, Second Medical Center of Chinese PLA General Hospital, Beijing, 100853, China

^e Research Center of Translational Medicine, Shanghai Children's Hospital, Shanghai Jiao Tong University, Shanghai, 200062, China

^f Shanghai Institute of Immunology, Department of Immunology and Microbiology, Shanghai Jiao Tong University School of Medicine, Shanghai, 200025, China

ARTICLE INFO

Keywords:

Immune-related lncRNA and mRNA signature
Discrimination
Machine learning
Bacterial sepsis
Children

ABSTRACT

Background: The variations in non-coding RNA alterations and the host immune response between patients with bacterial and non-bacterial sepsis, along with their clinical characteristics, are largely unknown.

Methods: The landscape of long non-coding RNA (lncRNA) and mRNA in whole blood cells from pediatric patients with bacterial sepsis or non-bacterial sepsis were characterized using an Arraystar human lncRNA microarray. Weighted correlation network analysis (WGCNA) were conducted to identify immune-related lncRNA-mRNA signatures. Least absolute shrinkage and selection operator (*Lasso*) regression and *Ridge* regression analysis were employed to develop a specific lncRNA-mRNA signature, serving as a discriminant classifier for bacterial and non-bacterial sepsis in children.

* Corresponding author. Department of Critical Care Medicine, Shanghai Children's Hospital, Laboratory of Critical Care Translational Medicine, Institute of Pediatric Infection, Immunity, and Critical Care Medicine, Shanghai Jiao Tong University School of Medicine; No. 355 Luding Road, Putuo District, Shanghai, 200062, China.

** Corresponding author. Department of Critical Care Medicine, Shanghai Children's Hospital; Laboratory of Critical Care Translational Medicine, Institute of Pediatric Infection, Immunity, and Critical Care Medicine, Shanghai Jiao Tong University School of Medicine; No. 355 Luding Road, Putuo District, Shanghai, 200062, China.

*** Corresponding author. Department of Critical Care Medicine, Shanghai Children's Hospital; Laboratory of Critical Care Translational Medicine, Institute of Pediatric Infection, Immunity, and Critical Care Medicine, Shanghai Jiao Tong University School of Medicine; No. 355 Luding Road, Putuo District, Shanghai, 200062, China.

E-mail addresses: karencx0465@163.com (C. Wang), sunting@shchildren.com.cn (T. Sun), zhouyp@shchildren.com.cn (Y. Zhou), m17853137921@163.com (T. Liu), fengshuyun2020@163.com (S. Feng), xiongxi3@163.com (X. Xiong), fan.j2007@163.com (J. Fan), liangqiming@shsmu.edu.cn (Q. Liang), cuiyun0815@163.com (Y. Cui), zyucui2018@163.com (Y. Zhang).

¹ lead contact.

² Authorship notes: CW and TS contributed equally to this work.

<https://doi.org/10.1016/j.heliyon.2024.e38728>

Received 5 October 2023; Received in revised form 24 September 2024; Accepted 28 September 2024

Available online 30 September 2024

2405-8440/© 2024 Published by Elsevier Ltd.

This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Results: A total of 33 differentially expressed lncRNAs and 52 mRNAs were identified in pediatric patients with either bacterial sepsis or non-bacterial sepsis. Among these, 69 lncRNAs and mRNAs were pinpointed using WGCNA and found to be significantly correlated with clinical parameters. Further intersection analysis identified 12 lncRNAs and 16 mRNAs as immune-related signature for discerning bacterial infections in children with sepsis. Additionally, the lncRNA-mRNA co-expression network highlighted the key lncRNAs (AC090159.1 and AC080129.2) and mRNAs (S100A8 and TCF7L2) as an infection score model. *Lasso* regression analysis revealed that this infection score model achieved an area under the received operating curve (AUROC) of 0.96 in the training set and 0.86 in the validation set. Ultimately, the expression levels of these 4 key lncRNAs and mRNAs showed significant correlation with CRP or PCT levels.

Conclusion: The machine learning model, developed utilizing key lncRNAs (AC090159.1 and AC080129.2) and mRNAs (S100A8 and TCF7L2), demonstrates robust discrimination and calibration capabilities for distinguishing between bacterial and non-bacterial sepsis in children.

1. Introduction

Sepsis is associated with high morbidity and mortality rates in critically ill children [1,2]. While adult sepsis was defined in 2016 as life-threatening organ dysfunction caused by a dysregulated host response to infection [3], these guidelines do not encompass pediatric cases. The 2005 International Pediatric Sepsis Consensus Conference (IPSCC) established criteria for pediatric sepsis based on the systemic inflammatory response syndrome (SIRS) [4]. However, studies indicate that SIRS criteria have low specificity for identifying infected children at a high risk of mortality. Consequently, the Pediatric Sepsis Definition Taskforce updated these criteria in 2024 using Phoenix score [5]. Despite improvements in the accurate recognition of pediatric sepsis, the primary step in sepsis treatment remains the etiological diagnosis, which is critical for targeted treatment.

Responsible pathogens are identified in only 60–70 % of cases. While bacterial infection is the primary cause of sepsis, viral sepsis, particularly heightened during the COVID-19 pandemic, often remains overlooked [6,7]. Early identification of the pathogen is crucial for guiding the appropriate use of antimicrobials or antiviral drugs. However, routine blood cultures are time-consuming, and a high proportion of negative results can lead to the misuse of broad-spectrum antimicrobials [8,9]. Reducing unnecessary use may help diminish the risk of developing antimicrobial resistance and support the development of the microbiome to enhance long-term health outcomes [10]. The duration of antibiotic therapy varies depending on the diversity among bacteria, virus, or other pathogens and their hosts. This necessitates an individualized approach to antibiotic durations using biomarker-assisted guidance. Generally, procalcitonin (PCT) -guided antibiotic use can enhance patient quality of life and reduce hospital costs for patients with undifferentiated sepsis [11,12]. Additionally, C-reactive protein (CRP) serves as a highly sensitive marker of the inflammatory response and is a well-known biomarker for guiding antibiotic duration [13,14]. Nevertheless, a deeper understanding of the host-response to bacterial and non-bacterial sepsis is essential for developing more effective biomarkers or targets for early diagnosis and treatment strategies.

Based on multi-omics data, the identification of “subclasses” with varying risks of outcomes and treatment responses can lead to positive treatment effects [15–17]. Noncoding (nc) RNAs extensively regulate inflammation-related pathways, mitochondrial dysfunction, cell apoptosis, and oxidative stress during sepsis, suggesting their potential as biomarkers and therapeutic targets for sepsis and sepsis-associated organ dysfunction (SAOD) [18]. Long non-coding RNAs (lncRNAs), a distinct category of ncRNA exceeding 200 nucleotides, do not encode protein. lncRNAs play critical roles in immune responses through mechanisms such as alternative splicing, subcellular localization, chromatin modification, transcription, and post-transcriptional processing. In the past 5 years, research has predominantly focused on lncRNAs to explore their molecular mechanisms in sepsis or SAOD [19,20]. However, there are no reports on lncRNA signatures related to host immune response as discriminative classifiers for bacterial versus non-bacterial sepsis. Currently, the distinctions in host-immune lncRNA signatures between bacterial and non-bacterial sepsis remain largely unexplored.

In this prospective cohort study, we examined the differences in lncRNA and mRNA profiles in whole blood cells of pediatric patients with bacterial sepsis or non-bacterial sepsis using comprehensive microarray analysis. We developed a host immune-related lncRNA and mRNA signature that serves an early discriminant classifier for bacterial versus non-bacterial sepsis in children. This research enhances the understanding of lncRNAs and mRNAs in pediatric sepsis, whether bacterial or non-bacterial.

2. Methods

2.1. Patients enrollment

Pediatric patients diagnosed with sepsis were prospectively enrolled in a pediatric intensive care unit (PICU) of Shanghai Children's Hospital between 2018 and 2019. Clinical adjudication, conducted by a team of intensivists with comprehensive access to medical record, occurred post-enrollment and prior to measuring lncRNAs expression. The criteria for pediatric sepsis followed the International Pediatric Sepsis Consensus Conference guidelines established in 2005 [4]. Inclusion criteria were: (1) age between 1 month and 18 years, and (2) diagnosis of sepsis within 24 h of PICU admission. Exclusion criteria included: (1) advanced tumor, (2) severe primary diseases or heredity metabolic diseases, and (3) administration of broad-spectrum antibiotics within 12 h before determining study eligibility.

Bacterial sepsis was defined based on the following criteria: a positive for bacterial culture in blood and PCT $\geq 0.5 \mu\text{g/L}$, accompanied by an elevated white blood cell (WBC) count, increased CRP, elevated erythrocyte sedimentation rate (ESR), and tumor necrosis factor (TNF) levels within 24 h prior to sample collection. Non-Bacterial sepsis was identified by a negative bacterial culture and a PCT $< 0.5 \mu\text{g/L}$ in blood, collected within 24 h prior to analysis, based on previous reports [21,22]. Management of sepsis patients included the implementation of sepsis resuscitation bundles, which involved early goal-directed fluid resuscitation, empiric antibiotic therapy tailored to the presumed site of infection, and early referral for source control procedures when indicated. Baseline characteristics and laboratory indices such as WBC, CRP, PCT, interleukin-1 β (IL-1 β), IL-2, IL-6, IL-8, IL-10, and tumor necrosis factor α (TNF α) and the percentages of CD4⁺CD8⁺ cells, NK cells, and CD19 cells were collected.

The study protocol was approved by the Ethics Committee of Children’s Hospital affiliated to Shanghai Jiao Tong University (approval number: 2018R039-F01). Informed consent was obtained from the patients’ parents or guardians. Privacy and data anonymization are ensured for all patient information in this study.

2.2. The flowchart of this study

The flowchart depicting the steps involved in this study is illustrated in Fig. 1. Differentially expressed genes (DEGs), including lncRNAs and mRNAs, were identified by analyzing lncRNA microarrays data from whole blood samples of patients with bacterial sepsis and non-bacterial sepsis. Subsequently, an immune infiltration analysis was conducted using the “IOBR” R package. Weighted Correlation Network Analysis (WGCNA) was employed to identify potential co-expression modules of lncRNAs and mRNA related to

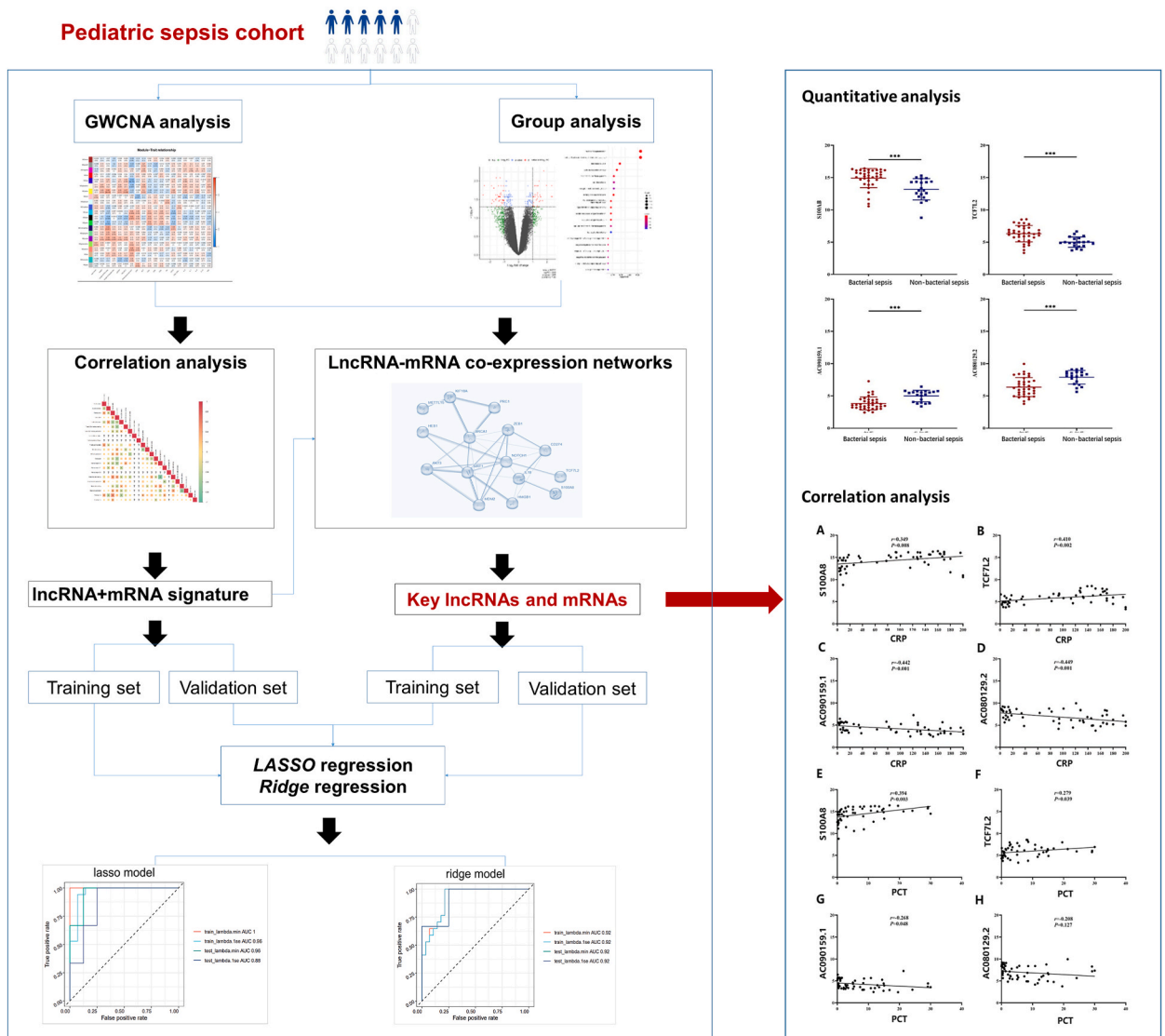


Fig. 1. Flowchart for this study.

clinical characteristics, particularly PCT, CRP, and other immune response indicators. The findings from WGCNA and group analyses were integrated to determine the immune-related lncRNA and mRNA signature. The immune-related lncRNA-mRNA network was analyzed, key lncRNAs and mRNAs were identified, and *lasso* or *ridge* regression model were developed using 80 % of the cases as the training set and 20 % as the validation set. Furthermore, the association between these 4 key lncRNAs and mRNAs with PCT and CRP was validated in this cohort. Additionally, the efficacy of each lncRNA and mRNA as a discriminant classifier for distinguishing bacterial from non-bacterial sepsis was evaluated in this cohort, using CRP as a positive control.

2.3. Blood samples

Blood samples were collected before the initiation of inpatient treatment. EDTA-anticoagulated whole blood was collected and stored at -80°C for further microarray analysis. Samples were removed from -80°C before further microarray analysis is performed to avoid repeated freeze-thaw. The time from blood collection to analysis was controlled within 2 years, and quality control was performed before the samples are sent to the microarray for analysis.

2.4. RNA extraction, labeling and array hybridization

The Arraystar human lncRNA/mRNA microarray was conducted on whole blood samples from 36 bacterial sepsis cases and 20 non-bacterial sepsis cases to identify differentially expressed lncRNAs. All experimental procedures were adhered to the Arraystar standard protocol. Briefly, RNA samples were extracted from whole blood cells using TRIzol reagent, followed by RNA quantification and quality control (QC). Samples that passed QC were labeled using the Quick Amp Labeling Kit, One-Color (Agilent p/n 5190-0442). Subsequently, the Arraystar Flash RNA Labeling Kit was employed for RNA labeling, and hybridization was performed using Agilent Gene Expression Hybridization Kit (Agilent p/n 5188–5242), followed by washing.

2.5. Microarray analysis

The ArrayStar Human lncRNA/mRNA Expression Microarray v3.0 (includes 30,586 lncRNA probes and 26,109 coding transcripts) and v5.0 (includes 39,317 lncRNA probes and 21,174 coding transcripts) provided by KangChen Bio-Tech, Inc (Shanghai, China), were used in our study. The arrays were scanned by Agilent Microarray Scanner (Agilent p/n G2565BA), and data were extracted using Agilent Feature Extraction software. The microarray data are available in the Gene Expression Omnibus (GEO) under Accession No. GSE233322. Comparative analysis between groups was conducted to delineate expression patterns associated with infectious pathogen type: confirmed bacterial *versus* non-bacterial pathogens.

2.6. Removing batch effects

Based on initial results from the Arraystar human lncRNA/mRNA microarray (v3.0) involving 16 cases (12 with bacterial sepsis vs. 4 with non-bacterial sepsis), we first identified a potential key relationship between the lncRNAs landscape and pathogen infection (**preliminary data, not shown**). Subsequently, an additional analysis was performed using the Arraystar human lncRNA/mRNA microarray (v5.0) using 40 cases (24 with bacterial sepsis vs. 16 with non-bacterial sepsis). In total, these two batches included 36 cases with bacterial sepsis and 20 cases with non-bacterial sepsis. To mitigate batch effects, we merged the results from 2 batches. All analyses were conducted using *R* version 4.0.5 and the *sva* package.

2.7. Identification of differentially expressed genes (DEGs)

Utilize the “limma” package to identify differentially expressed genes (DEGs) between bacterial sepsis and non-bacterial sepsis, employing a default screening condition of p -value <0.05 and $|\log_2\text{FoldChange}| (|\log\text{FC}|) \geq 1$. According to the screening criteria $|\log\text{FC}| \geq 1$, the difference between the two groups was at least greater than 2-fold. Differentially expressed lncRNAs and mRNAs were identified using a volcano plot. To explore the potential function of these lncRNAs and mRNAs, related protein-coding genes were examined, and both Gene Ontology (GO) analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis were conducted. Further, gene set enrichment analysis (GSEA) was performed using the “cluster Profiler” *R* package, considering an adjusted p -value <0.05 as statistically significant.

2.8. Weighted correlation network analysis (WGCNA)

The transcriptome of lncRNAs and mRNA were consolidated, extracting 90 % of the genes based on variation for subsequent Weighted Correlation Network Analysis (WGCNA). WGCNA primarily predicts the functions of lncRNAs by analyzing their co-expressed genes and clinical characteristics. Using the WGCNA package, network analysis assessed the relative importance of lncRNAs and their module membership. Hierarchical clustering was employed to produce a sample clustering diagram, and a heat map was generated based on the degree of topological overlap. The link strength between nodes i and j in the co-expression network was calculated using Pearson's correlation coefficient $\text{cor}(i, j)$. The weighted adjacency matrix a_{ij} was calculated as follows: $a_{ij} = (0.5 \times (1 + \text{cor}(i, j)))^\beta$ ($\beta = 7$), focusing on the independence and average connectivity degree of co-expression modules. The reconstructed co-expression network was visualized using Cytoscape software (version 3.7.0), and clustering graphs were created using the dynamic tree

cutting technique of the pheatmap package in R software.

2.9. Immune-related lncRNA and mRNA signature

The module hub genes (KME >0.8) were correlated with clinical parameters using WGCNA. Differentially expressed lncRNAs and mRNAs identified from group analysis were integrated to identify potential lncRNA and mRNA signature associated with bacterial and non-bacterial infections.

2.10. The immune-related lncRNA-mRNA co-expression network and screening the key lncRNAs and mRNAs as a model of infection score

The immune-related lncRNA-mRNA co-expression networks were constructed to explore the association between lncRNA and mRNA. To investigate the biological functions of the immune-related lncRNA-mRNA signature involved in the pathogen-host immune response, we linked them with their co-expressed genes, either reported in the literature or predicted target genes identified using LncRNADisease v2.0 (<http://www.rnanut.net/lncrnadisease/index.php/home/search>). This analysis aims to uncover the pathways that regulate lncRNA in response to bacterial or viral infections, thereby predicting potential mechanisms of lncRNA in disease contexts. Significant correlation pairs were used to establish the network based on Pearson correlation coefficients. The differential co-expression network charts were visualized and analyzed using the STRING online analysis tool (<https://cn.string-db.org/>).

2.11. Statistical analysis

For clinical data analysis, continuous data were presented as mean ± standard deviation (SD) for normal distributed variables or as

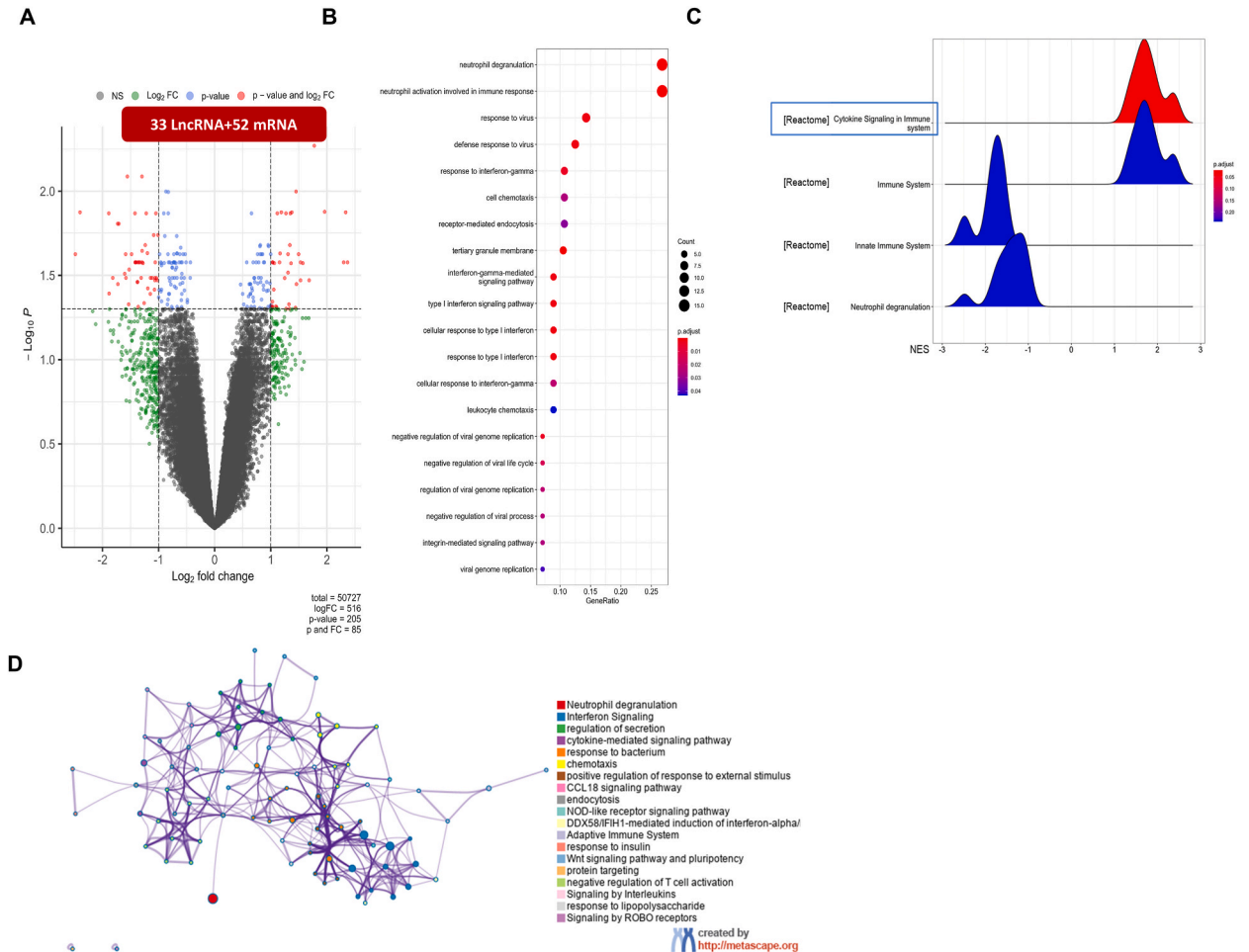


Fig. 2. Differentially expressed lncRNAs and mRNAs between bacterial or non-bacterial-response (A) Volcano plot, (B) GO analysis, (C) GSEA using the REACTOME database, (D) network of enriched terms, colored by cluster ID with nodes sharing the same cluster ID generally positioned in close proximity, created by Metascape.

3.2. Assessing technical variation

The study was comprised of two distinct batches: the first batch included 16 cases (12 cases with bacterial sepsis and 4 with non-bacterial sepsis), and the second batch included 40 cases (24 with bacterial sepsis and 16 with non-bacterial sepsis), collectively forming the study sample. There was no significant difference in the proportion of bacterial sepsis cases between the first and second batches (75 % vs. 60 %, $P = 0.290$). Principal Coordinates Analysis (PCoA) indicated that the normalized gene expression was consistent before and after correcting for batch effects using both the “limma” and “sva” R packages (Supplementary Fig. 1).

3.3. Differential expressed lncRNA and mRNAs involved in infection response to bacteria or non-bacteria

After removing batch effects, the cohorts with bacterial and non-bacterial infections exhibited 33 differentially expressed lncRNAs and 52 differentially expressed mRNAs (Fig. 2A and Supplementary Table 3). Gene Ontology (GO) enrichment analysis identified key signaling pathways implicated in the differential immune responses to bacterial or non-bacterial infections, including neutrophil degranulation, activation of neutrophils, response to viruses, defense response to viruses, and response to interferon- γ (IFN γ). The pathways also encompass IFN γ -mediated signaling, type I interferon (IFN-1) signaling pathway, cellular responses to IFN-1, leukocyte chemotaxis, integrin-mediated signaling pathway, cytokine-mediated signaling pathway, the adaptive immune system, and negative regulation of T cell activation (Fig. 2B and Supplementary Table 4). Moreover, we utilized GSEA with the REACTOME database to identify major functional pathways, finding significant enrichment ($p < 0.05$) in cytokine signaling in the immune system (Fig. 2C). Furthermore, KEGG enrichment analysis showed that these lncRNAs and mRNAs are likely to be involved in pathways associated with O-glycan biosynthesis, Type II diabetes mellitus, osteoclast differentiation, hepatitis C, folate biosynthesis, hematopoietic cell lineage, biosynthesis of unsaturated fatty acids, lysosome, glycosphingolipid biosynthesis - lacto and neolacto series, pentose phosphate pathway, galactose metabolism, and thyroid cancer (Fig. 2D and Supplementary Table 5). These results suggest that infections in sepsis, whether bacterial or non-bacterial, engage multiple pathways primarily related to immune response and metabolic regulation.

3.4. WGCNA network and module-trait relationship analysis

The lncRNAs and mRNAs identified using the Arraystar Human lncRNA Microarray were analyzed through WGCNA to identify modules and hub genes. Initially, sample clustering, cluster dendrogram, and eigengene networks were constructed (Fig. 3A–C). Subsequently, the correlation between these co-expression modules and clinical traits was evaluated, utilizing phenotypic data such as age, weight, length of PICU and hospital stays, pSOFA, PELD2, WBC, CRP, PCT, and various immune cells and inflammatory markers. A significant co-expression module, closely associated with PCT, CRP, or other inflammatory factors, was determined. Module-trait relationships were quantified using Pearson’s correlation tests, with a P -value < 0.05 indicating significant correlation. Hub genes from the green, cyan, black, pink, yellow, and grey modules were identified (Fig. 3D).

3.5. Immune-related lncRNA-mRNA signature

To further develop the immune-related lncRNA-mRNA signature, the black module encompassed 16 mRNAs (PLP2, NLRC4, XRN2,

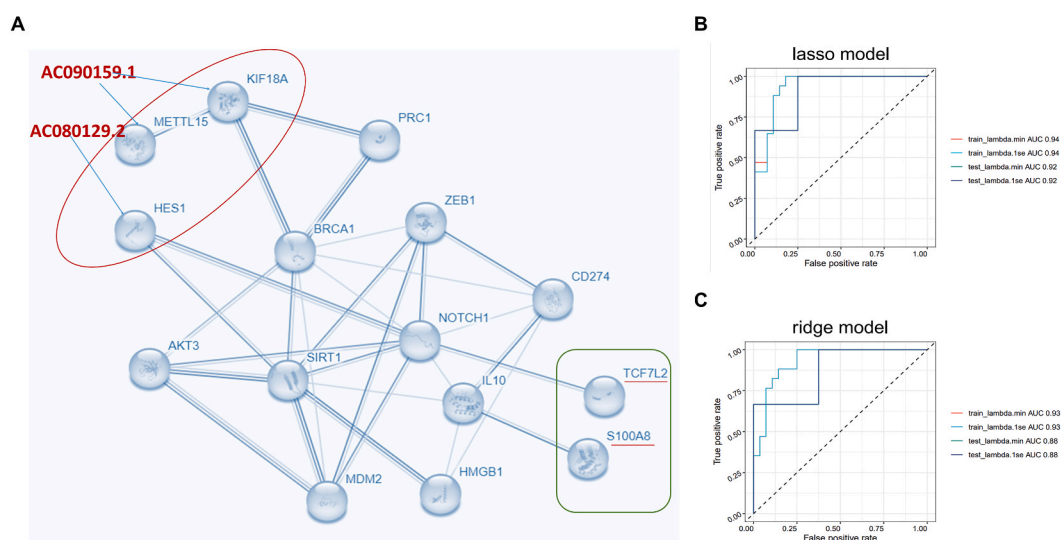


Fig. 4. The immune-related lncRNA-mRNA co-expression network identifies key lncRNAs and mRNAs employing λ_{\min} (min) or $\lambda_{\min} + 1\text{se}$ (1se) to distinguish between bacterial and non-bacterial infections in both the training set (train) or validation set (test). (A) Immune-related lncRNA-mRNA co-expression network, (B) lasso regression model, (C) ridge regression model.

COLGALT1, ITGAM, LAPTM5, LILRB3, ACTA2, HK3, TCF7L2, CAP1, NBEAL2, S100A8, EMB, SYNE1, ZNF467), while the pink module, significantly correlated to CRP, included 12 lncRNAs (Supplementary Table 6). Additionally, the potential immune-related lncRNA-mRNA signature was analyzed using *lasso* and *ridge* regression to construct an infection score. The AUC of infection score, using *lasso* or *ridge* regression based on λ_{\min} for differentiating bacterial from non-bacterial sepsis, was 1.0 or 0.92 in the training set, and 0.96 or 0.92 in the validation set, respectively (Supplementary Figs. 2A and 2B).

3.6. The immune-related lncRNA-mRNA co-expression networks to further focus on the key lncRNAs and mRNAs

According to the reported signaling pathway involving immune-related lncRNAs, mRNAs, or predicted target genes of lncRNAs, the co-expressed immune-related lncRNA-mRNA signature notably includes S100A8, TCF7L2, AC090159.1, and AC080129.2 (Fig. 4A). The differential expression of S100A8, TCF7L2, AC090159.1, and AC080129.2 between bacterial and non-bacterial sepsis was significant according to Arraystar Human lncRNA microarray results. The mRNA levels of S100A8 and TCF7L2 were elevated in whole blood cells of patients with bacterial sepsis compared to those with non-bacterial sepsis, whereas the expressions of lncRNAs AC090159.1 and AC080129.2 were significantly lower (Supplementary Table 7).

Genes annotated as related to the activation of immune cells were implicated in the host immune response (Supplementary Table 8). The *lasso* and *ridge* regression model, which incorporated key biomarkers comprising 2 lncRNAs and 2 mRNAs based on λ_{\min} , achieved an AUC of 0.94 or 0.93 in the training set, and 0.92 and 0.88 in the validation set, respectively (Fig. 4B and C). Furthermore, using S100A8, TCF7L2, AC090159.1, or AC080129.2 as single biomarkers yielded AUCs of 0.825 (95 % CI: 0.718–0.932), 0.810 (95 % CI: 0.696–0.924), 0.822 (95 % CI: 0.712–0.932), and 0.799 (95 % CI: 0.683–0.914), respectively (Supplementary Table 9).

3.7. The correlation of key lncRNAs and mRNAs signature to the levels of CRP and PCT

There was a significantly positive correlation between S100A8 ($r = 0.349$, $P = 0.008$) and CRP, as well as TCF7L2 ($r = 0.410$, $P = 0.002$) and CRP. Conversely, a significantly negative correlation was observed between AC090159.1 ($r = -0.442$, $P = 0.001$) and CRP, and between AC080129.2 ($r = -0.449$, $P = 0.001$) and CRP. Interestingly, similar correlations were also noted between these four key lncRNAs and mRNAs and PCT levels (Supplementary Fig.3 and Supplementary Table 10).

4. Discussion

The COVID-19 pandemic has heightened concerns regarding viral sepsis. In septic patients without evidence of bacterial, parasitic or fungal infection, viral sepsis should be considered and appropriate viral laboratory tests conducted [6]. Rapid developments have occurred in targeting the host response to infection and targeting virus through antiviral replication and neutralizing antibody therapies during COVID-19 epidemic [23]. Additionally, immune checkpoint inhibitors (ICIs), such as programmed cell death protein 1/programmed death-ligand 1 blockade, are garnering attention in the treatment of viral infections [24]. Consequently, the early and effective identification of bacterial versus non-bacterial sepsis gains importance for precision treatment. In this study, we report the key role of immune-related lncRNA-mRNA signatures in effectively distinguishing bacterial sepsis from non-bacterial sepsis, mainly including S100A8, TCF7L2, AC090159.1, and AC080129.2. Due to the small sample size, the conclusions need to be verified by more clinical studies.

The reliance on bacterial culture as the gold standard for diagnosing bacterial infection presents challenges due to its time-consuming nature and potential for false negative or positive results, which deviate from precision therapy. Furthermore, despite for considering well-established biomarkers (CRP, PCT, etc.) for guiding antibiotic use, most sepsis patients can be effectively treated with relatively short antibiotic courses of approximately 7 days even without biomarker guidance [11]. Thus, a reliable biomarker remains a critical, unmet need for promptly identifying patients who require antibiotic therapy and for assessing the response or duration of antibiotics use among septic patients. In this study, the criteria for determining bacterial and non-bacterial sepsis were determined by PCT and CRP levels, which may limit the generality of the main findings of this study.

There is significant heterogeneity in the pathogens infecting PICU patients, as well as substantial differences in the quality of the hosts that they infect [25]. Consequently, the diverse pathogen-host immune response could represent the primary, common clinical characteristics regardless of the specific infecting pathogen. To the best of our knowledge, this study is the first to report significant differences in the non-coding RNA and mRNA profiles in whole blood cells from patients with bacterial or non-bacterial sepsis. Additionally, the differentially expressed mRNA and lncRNAs are involved in immune response and metabolic regulation, including neutrophil degranulation, neutrophil activation, IFN γ or IFN-1-mediated signaling pathway, cytokine signaling, and O-glycan, folate, unsaturated fatty acids, glycosphingolipid biosynthesis, as well as the pentose phosphate pathway, and galactose metabolism, among others. Among these critical pathways, neutrophil granulocytes play a pivotal role in the initial host defense against invading pathogens. Oxidants such as MPO or NADPH functions as essential signaling molecules are essential for the regulation of neutrophil trafficking, activation, phagocytosis, and formation of extracellular traps [26–28]. In predicting bacterial infections, the AUC of modified neutrophil CD64 (nCD64) was significantly higher than that of CRP [29]. Many identified glycosphingolipids, stored in granules, posit glycosphingolipid beta-glucosylceramide as a potent ligand of macrophage-inducible C-type lectin, which regulates NET formation and has diagnostic and therapeutic potential in inflammatory disease [30,31]. Persistently activated type I IFN signaling drives pathogenic neutrophil responses and highlights IL-18 as a novel component of disease during genital Herpes S Viruses 2 infection [32]. In severe COVID-19, neutrophils exhibits a metabolic shift with reduced activity of the glycolytic enzyme GAPDH, leading to inhibited glycolysis and enhanced activity of the pentose phosphate pathway activity, contributing to their dysfunction

[33]. Furthermore, lipid droplets (LDs) provide essential substrates for intracellular viruses, bacteria, and other invaders, facilitating host colonization; LDs also possess protein-mediated antibiotic activity, which is upregulated in response to danger signals and sepsis [34]. Thus, the synthesis of unsaturated fatty acids signaling pathway may be another critical regulator in pathogen-host immune responses. Although further research is necessary to delineate the relationship between neutrophils activation and metabolism, our findings provide initial insight into the interplay between neutrophils, metabolism, and IFN signaling pathway, which could be an important immune marker for distinguishing between bacterial and non-bacterial sepsis.

Non-coding RNAs (ncRNAs) are increasingly recognized as significant contributors to various human diseases and present potential therapeutic targets [35]. Recent evidence has elucidated the crucial roles of lncRNA in the onset of sepsis, sepsis-associated organ dysfunction, and sepsis-related immune responses through lncRNA-microRNA signaling networks and as circulating biomarkers [36–38]. This study, focusing on distinguishing bacterial sepsis from non-bacterial sepsis, is the first to report a host immune-related lncRNA and mRNA signature that includes S100A8, TCF7L2, and 2 lncRNAs (AC090159.1 and AC080129.2), achieving an AUC of 0.94 and 0.92 in the testing and validation sets respectively, based on the lasso regression model. Using CRP and PCT as positive controls, the signature proved comparable as a discriminant classifier for bacterial sepsis, highlighting the significance of the altered expression of S100A8, TCF7L2, and 2 lncRNAs (AC090159.1 and AC080129.2) in whole blood.

Until now, S100A8 has been recognized as a predictive marker for neonatal sepsis and a therapeutic target for long-term pulmonary complications following sepsis [31–33]. Additionally, S100A8 supplementation has been shown to mitigate injury in the lungs, kidneys, and livers associated with sepsis, exhibiting protective and anti-inflammatory effects [39,40]. However, genetic ablation of the S100A8/A9-TLR4 signaling axis enhances survival in mice with CLP-induced sepsis by inhibiting platelet pyroptosis [20]. In this study, S100A8 levels were significantly elevated in the whole blood cells of patients with bacterial sepsis compared to those with non-bacterial sepsis, correlating positively with CRP and PCT levels. Furthermore, the AUC for S100A8 alone was 0.825 (95 % CI: 0.718–0.932), suggesting its potential as a novel biomarker for early identification of bacterial sepsis. TCF7L2 plays roles in macrophage differentiation [41], virus replication [42], and the Wnt/ β -catenin signaling pathway [43,44]. Notably, the AUC for TCF7L2 alone reached 0.810 (95 % CI: 0.696–0.924), indicating the need for further investigation into the mechanisms underlying TCF7L2-mediated pathogen-host immune response. Additionally, AC090159.1 and AC080129.2 were identified as two lncRNA biomarkers with significant capacity to discriminate bacterial sepsis. Although direct reports on the function of AC090159.1 and AC080129.2 are lacking, predictive target genes such as KIF18A, METTL15, and HES1 suggest their involvement in the host-pathogen immune response. Specifically, METTL15 is essential for mitochondrial function and mitoribosome biogenesis [45,46]. These target genes also show network interaction with S100A8 and TCF7L2, associated with IL10 [41], NOTCH 1 [47,48], and sirtuin 1 (SIRT1) [49]. The Notch1/HES1 signaling pathway promotes IL-10 production in macrophages, which helps suppress inflammation in LPS-induced sepsis and liver injury [45]. To the best of our knowledge, this is the first report about the association of lncRNA with the NOTCH1-HES1 in whole blood cells across different pathogen infections, underscoring the need for further research in this area.

5. Limitation

This study has several limitations. Although two batch cohorts were utilized to investigate the roles of 2 key lncRNAs and 2 mRNAs in early identification of bacterial sepsis, this is still a single center study with a small sample size. Consequently, our current findings are limited in determining the causative agent of severe infections, particularly in distinguishing bacterial from non-bacteria origins. Furthermore, the definition of bacterial sepsis relied on a positive for bacterial culture and PCT ≥ 0.5 $\mu\text{g/L}$ within 24 h prior to sample collection. However, this study lacked precise methods for accurate pathogen detection. Additionally, due to limitations in obtaining blood samples from children, data confirming the expression of these 4 key genes through RT-PCR were insufficient. The conclusions drawn from this study require further validation in an external cohort and additional samples. Moreover, the underlying mechanisms by which these 4 key genes contribute to the host-immune response to bacterial or non-bacterial agents warrant more thorough investigation.

6. Conclusion

In conclusion, we identified and validated the first non-coding signature, comprising 2 lncRNAs and 2 mRNA, that effectively distinguishes bacterial sepsis from non-bacterial sepsis. Despite limitations such as a limited number of lncRNA features and a small sample size, we demonstrated that lncRNAs can serve as markers for the discriminant classifier of bacteria versus non-bacterial infections. The proposed model offers a potential alternative or complementary diagnostic tool for sepsis.

Data availability statement

All data generated or analyzed used in this study can be found at GEO (<https://www.ncbi.nlm.nih.gov/geo/>) under the accession number GSE233322. Software and resources used for the analyses are described in each method section.

Ethics approval and consent to participate

The study protocol received approval from the Ethics Committee of Children's Hospital affiliated to Shanghai Jiao Tong University (approval number: 2018R039-F01). Informed consent was obtained from the patients' parents or relatives.

Additional information

All data generated or analyzed during this study are accessible at the GEO database (<https://www.ncbi.nlm.nih.gov/geo/>) under accession number GSE233322. The software and resources employed for the analyses are detailed in the respective methods sections.

Funding

This study was supported by the National Natural Science Foundation of China (82171729, 82100085), the Shanghai Municipal Commission of Health and Family Planning (202340076), the National Key Research and Development Program of China (2021YFC2701800, 2021YFC2701805), the Science and Technology Commission of Shanghai Municipality (18411951000, and 21Y11902600). Dr. Chunxia Wang is funded by Shanghai Municipal Education Commission–Gaofeng Clinical Medicine Grant (20171928), the talent program from Shanghai Jiao Tong University School of Medicine (17XJ11018), and discipline building project from Institute of Pediatric Infection, Immune, and Critical Care Medicine, Shanghai Children's Hospital, Shanghai Jiao Tong University School of Medicine.

CRedit authorship contribution statement

Chunxia Wang: Writing – review & editing, Writing – original draft, Supervision, Investigation, Funding acquisition. **Ting Sun:** Software, Methodology, Formal analysis, Data curation. **Yiping Zhou:** Writing – review & editing, Funding acquisition, Data curation. **Tiantian Liu:** Writing – review & editing, Data curation. **Shuyun Feng:** Writing – review & editing, Methodology, Data curation. **Xi Xiong:** Writing – review & editing, Methodology, Data curation. **Jiao Fan:** Writing – review & editing, Software, Formal analysis. **Qiming Liang:** Writing – review & editing, Methodology. **Yun Cui:** Writing – review & editing, Investigation, Funding acquisition. **Yucai Zhang:** Writing – review & editing, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.heliyon.2024.e38728>.

References

- [1] E.Y. Killien, R.W.D. Farris, R.S. Watson, L.A. Dervan, J.J. Zimmerman, Health-related quality of life among survivors of pediatric sepsis, *Pediatr. Crit. Care Med.* 20 (2019) 501–509.
- [2] J.J. Zimmerman, R. Banks, R.A. Berg, A. Zuppa, C.J. Newth, D. Wessel, M.M. Pollack, K.L. Meert, M.W. Hall, M. Quasney, et al., Critical illness factors associated with long-term mortality and health-related quality of life morbidity following community-acquired pediatric septic shock, *Crit. Care Med.* 48 (2020) 319–328.
- [3] M. Shankar-Hari, G.S. Phillips, M.L. Levy, C.W. Seymour, V.X. Liu, C.S. Deutschman, D.C. Angus, G.D. Rubenfeld, M. Singer, Sepsis definitions task F: developing a new definition and assessing new clinical criteria for septic shock: for the third international consensus definitions for sepsis and septic shock (Sepsis-3), *JAMA* 315 (2016) 775–787.
- [4] B. Goldstein, B. Giroir, A. Randolph, International pediatric sepsis consensus conference: definitions for sepsis and organ dysfunction in pediatrics, *Pediatr. Crit. Care Med.* 6 (2005) 2–8.
- [5] L.J. Schlapbach, R.S. Watson, L.R. Sorce, A.C. Argent, K. Menon, M.W. Hall, S. Akech, D.J. Albers, E.R. Alpern, F. Balamuth, et al., International consensus criteria for pediatric sepsis and septic shock, *JAMA* 331 (2024) 665–674.
- [6] G.L. Lin, J.P. McGinley, S.B. Drysdale, A.J. Pollard, Epidemiology and immune pathogenesis of viral sepsis, *Front. Immunol.* 9 (2018) 2147.
- [7] D. Liu, Q. Wang, H. Zhang, L. Cui, F. Shen, Y. Chen, J. Sun, L. Gan, J. Sun, J. Wang, et al., Viral sepsis is a complication in patients with novel corona virus disease (COVID-19), *Med Drug Discov* 8 (2020) 100057.
- [8] S.P. Costa, C.M. Carvalho, Burden of bacterial bloodstream infections and recent advances for diagnosis, *Pathog Dis* 80 (2022) ftac027.
- [9] L. Samuel, Direct detection of pathogens in bloodstream during sepsis: are we there yet? *J Appl Lab Med* 3 (2019) 631–642.
- [10] E. Giannoni, V. Dimopoulou, C. Klingenberg, L. Naver, V. Nordberg, A. Berardi, S. El Helou, G. Fusch, J.M. Bliss, D. Lehnick, et al., Analysis of antibiotic exposure and early-onset neonatal sepsis in Europe, North America, and Australia, *JAMA Netw. Open* 5 (2022) e2243691.
- [11] W.C. Albrich, S. Harbarth, Pros and cons of using biomarkers versus clinical decisions in start and stop decisions for antibiotics in the critical care setting, *Intensive Care Med.* 41 (2015) 1739–1751.
- [12] E. de Jong, J.A. van Oers, A. Beishuizen, P. Vos, W.J. Vermeijden, L.E. Haas, B.G. Loef, T. Dormans, G.C. van Melsen, Y.C. Kluiters, et al., Efficacy and safety of procalcitonin guidance in reducing the duration of antibiotic treatment in critically ill patients: a randomised, controlled, open-label trial, *Lancet Infect. Dis.* 16 (2016) 819–827.
- [13] C.F. Oliveira, F.A. Botoni, C.R. Oliveira, C.B. Silva, H.A. Pereira, J.C. Serufo, V. Nobre, Procalcitonin versus C-reactive protein for guiding antibiotic therapy in sepsis: a randomized trial, *Crit. Care Med.* 41 (2013) 2336–2343.
- [14] E. von Dach, W.C. Albrich, A.S. Brunel, V. Prendki, C. Cuvelier, D. Flury, A. Gayet-Ageron, B. Huttner, P. Kohler, E. Lemmenmeier, et al., Effect of C-reactive protein-guided antibiotic treatment duration, 7-day treatment, or 14-day treatment on 30-day clinical failure rate in patients with uncomplicated gram-negative bacteremia: a randomized clinical trial, *JAMA* 323 (2020) 2160–2169.
- [15] K.M. DeMerle, D.C. Angus, J.K. Baillie, E. Brant, C.S. Calfee, J. Carrillo, C.H. Chang, R. Dickson, I. Evans, A.C. Gordon, et al., Sepsis subclasses: a framework for development and interpretation, *Crit. Care Med.* 49 (2021) 748–759.

- [16] E. Cano-Gamez, K.L. Burnham, C. Goh, A. Allcock, Z.H. Malick, L. Overend, A. Kwok, D.A. Smith, H. Peters-Sengers, D. Antcliffe, et al., An immune dysfunction score for stratification of patients with acute infection based on whole-blood gene expression, *Sci. Transl. Med.* 14 (2022) eabq4433.
- [17] M.J. Cummings, B. Bakamutumaho, A. Price, N. Owor, J. Kayiwa, J. Namulondo, T. Byaruhanga, M. Muwanga, C. Nseroko, S. Sameroff, et al., Multidimensional analysis of the host response reveals prognostic and pathogen-driven immune subtypes among adults with sepsis in Uganda, *Crit. Care* 26 (2022) 36.
- [18] T. Zhang, Y.H. Yang, Y.P. Liu, T.N. Zhang, N. Yang, Regulatory role of noncoding rna in sepsis and sepsis-associated organ dysfunction: an updated systematic review, *Shock* 58 (2022) 434–456.
- [19] X. Zhang, X. Tang, L. Pan, Y. Li, J. Li, C. Li, Elevated lncRNA-UCA1 upregulates EZH2 to promote inflammatory response in sepsis-induced pneumonia via inhibiting HOXA1, *Carcinogenesis* 43 (2022) 371–381.
- [20] M. Su, C. Chen, S. Li, M. Li, Z. Zeng, Y. Zhang, L. Xia, X. Li, D. Zheng, Q. Lin, et al., Gasdermin D-dependent platelet pyroptosis exacerbates NET formation and inflammation in severe sepsis, *Nat Cardiovasc Res* 1 (2022) 732–747.
- [21] P.E. Charles, S. Ladoire, S. Aho, J.P. Quenot, J.M. Doise, S. Prin, N.O. Olsson, B. Blettery, Serum procalcitonin elevation in critically ill patients at the onset of bacteremia caused by either Gram negative or Gram positive bacteria, *BMC Infect. Dis.* 8 (2008) 38.
- [22] V. Nobre, S. Harbarth, J.D. Graf, P. Rohner, J. Pugin, Use of procalcitonin to shorten antibiotic treatment duration in septic patients: a randomized trial, *Am. J. Respir. Crit. Care Med.* 177 (2008) 498–505.
- [23] N. Murakami, R. Hayden, T. Hills, H. Al-Samkari, J. Casey, L. Del Sorbo, P.R. Lawler, M.E. Sise, D.E. Leaf, Therapeutic advances in COVID-19, *Nat. Rev. Nephrol.* 19 (2023) 38–52.
- [24] K. Tsiakos, N. Gavrielatou, I.A. Vathiotis, L. Chatzis, S. Chatzis, G. Poulakou, E. Kotteas, N.K. Syrigos, Programmed cell death protein 1 Axis inhibition in viral infections: clinical data and therapeutic opportunities, *Vaccines (Basel)* 10 (2022) 1673.
- [25] K.N. Iskander, M.F. Osuchowski, D.J. Stearns-Kurosawa, S. Kurosawa, D. Stepien, C. Valentine, D.G. Remick, Sepsis: multiple abnormalities, heterogeneous responses, and evolving understanding, *Physiol. Rev.* 93 (2013) 1247–1288.
- [26] S.A. Rizo-Tellez, M. Sekheri, J.G. Filep, Myeloperoxidase: regulation of neutrophil function and target for therapy, *Antioxidants* 11 (2022) 2302.
- [27] M.Y. Zeng, I. Miralda, C.L. Armstrong, S.M. Uriarte, J. Bagaitkar, The roles of NADPH oxidase in modulating neutrophil effector responses, *Mol Oral Microbiol* 34 (2019) 27–38.
- [28] E.C. Britt, J. Lika, M.A. Giese, T.J. Schoen, G.L. Seim, Z. Huang, P.Y. Lee, A. Huttenlocher, J. Fan, Switching to the cyclic pentose phosphate pathway powers the oxidative burst in activated neutrophils, *Nat. Metab.* 4 (2022) 389–403.
- [29] Z. He, J.J. Liu, G.F. Shu, S.L. Ma, G.Q. Wu, Diagnostic utility of CD64 and CD38 biomarkers for the differential diagnosis of infections, *Genet. Test. Mol. Biomarkers* 26 (2022) 133–139.
- [30] A. Sharma, A. Chauhan, P. Chauhan, D.L. Evans, R.E. Szlabick, M.O. Aaland, B.B. Mishra, J. Sharma, Glycolipid metabolite beta-glucosylceramide is a neutrophil extracellular trap-inducing ligand of mincle released during bacterial infection and inflammation, *J. Immunol.* 209 (2022) 391–400.
- [31] A. Karlsson, H. Miller-Podraza, P. Johansson, K.A. Karlsson, C. Dahlgren, S. Teneberg, Different glycosphingolipid composition in human neutrophil subcellular compartments, *Glycoconj. J.* 18 (2001) 231–243.
- [32] T. Lebratti, Y.S. Lim, A. Cofie, P. Andhey, X. Jiang, J. Scott, M.R. Fabbri, A.N. Ozanturk, C. Pham, R. Clemens, et al., A sustained type I IFN-neutrophil-IL-18 axis drives pathology during mucosal viral infection, *Elife* 10 (2021) e65762.
- [33] Y. Li, J.S. Hook, Q. Ding, X. Xiao, S.S. Chung, M. Mettlen, L. Xu, J.G. Moreland, M. Agathocleous, Neutrophil metabolomics in severe COVID-19 reveal GAPDH as a suppressor of neutrophil extracellular trap formation, *Nat. Commun.* 14 (2023) 2610.
- [34] R. Safi, M. Sanchez-Alvarez, M. Bosch, C. Demangel, R.G. Parton, A. Pol, Defensive-lipid droplets: cellular organelles designed for antimicrobial immunity, *Immunol. Rev.* 317 (2023) 113–136.
- [35] K. Nemeth, R. Bayraktar, M. Ferracin, G.A. Calin, Non-coding RNAs in disease: from mechanisms to therapeutics, *Nat. Rev. Genet.* 25 (2024) 211–232.
- [36] Y. Wu, P. Li, L. Liu, A.J. Goodwin, P.V. Halushka, T. Hirose, S. Nakagawa, J. Zhou, M. Liu, H. Fan, lncRNA Neat1 regulates neuronal dysfunction post-sepsis via stabilization of hemoglobin subunit beta, *Mol. Ther.* 30 (2022) 2618–2632.
- [37] X. Pan, L. He, lncRNA MEG3 expression in sepsis and its effect on LPS-induced macrophage function, *Cell Mol Biol (Noisy-le-grand)* 66 (2020) 131–136.
- [38] Q. Huang, C. Huang, Y. Luo, F. He, R. Zhang, Circulating lncRNA NEAT1 correlates with increased risk, elevated severity and unfavorable prognosis in sepsis patients, *Am. J. Emerg. Med.* 36 (2018) 1659–1663.
- [39] S.J. Denstaedt, J.L. Spencer-Segal, M.W. Newstead, K. Laborc, A.P. Zhao, A. Hjelmaas, X. Zeng, H. Akil, T.J. Standiford, B.H. Singer, S100A8/A9 drives neuroinflammatory priming and protects against anxiety-like behavior after sepsis, *J. Immunol.* 200 (2018) 3188–3200.
- [40] D. Miyashita, R. Inoue, T. Tsuno, T. Okuyama, M. Kyohara, C. Nakahashi-Oda, K. Nishiyama, S. Fukushima, Y. Inada, Y. Togashi, et al., Protective effects of S100A8 on sepsis mortality: links to sepsis risk in obesity and diabetes, *iScience* 25 (2022) 105662.
- [41] Y.S. Baek, S. Haas, H. Hackstein, G. Bein, M. Hernandez-Santana, H. Lehrach, S. Sauer, H. Seitz, Identification of novel transcriptional regulators involved in macrophage differentiation and activation in U937 cells, *BMC Immunol.* 10 (2009) 18.
- [42] L. Ma, H. Zhang, Y. Zhang, H. Li, M. An, B. Zhao, H. Ding, J. Xu, H. Shang, X. Han, Integrated analysis of lncRNA, miRNA and mRNA profiles reveals potential lncRNA functions during early HIV infection, *J. Transl. Med.* 19 (2021) 135.
- [43] N. Hou, B. Ye, X. Li, K.B. Margulies, H. Xu, X. Wang, F. Li, Transcription factor 7-like 2 mediates canonical Wnt/beta-catenin signaling and c-Myc upregulation in heart failure, *Circ Heart Fail* 9 (2016).
- [44] W. Shao, V. Szeto, Z. Song, L. Tian, Z.P. Feng, M.C. Nostro, T. Jin, The LIM homeodomain protein ISL1 mediates the function of TCF7L2 in pancreatic beta cells, *J. Mol. Endocrinol.* 61 (2018) 1–12.
- [45] H. Chen, Z. Shi, J. Guo, K.J. Chang, Q. Chen, C.H. Yao, M.C. Haigis, Y. Shi, The human mitochondrial 12S rRNA m(4)C methyltransferase METTL15 is required for mitochondrial function, *J. Biol. Chem.* 295 (2020) 8505–8513.
- [46] L. Van Haute, A.G. Hendrick, A.R. D'Souza, C.A. Powell, P. Rebelo-Guiomar, M.E. Harbour, S. Ding, I.M. Fearnley, B. Andrews, M. Minczuk, METTL15 introduces N4-methylcytidine into human mitochondrial 12S rRNA and is required for mitoribosome biogenesis, *Nucleic Acids Res.* 47 (2019) 10267–10281.
- [47] Y. Xu, Y. Li, X. Liu, Y. Pan, Z. Sun, Y. Xue, T. Wang, H. Dou, Y. Hou, SPIONs enhances IL-10-producing macrophages to relieve sepsis via Cav1-Notch1/HES1-mediated autophagy, *Int J Nanomedicine* 14 (2019) 6779–6797.
- [48] X. Wang, Y. Yu, MiR-146b protect against sepsis induced mice myocardial injury through inhibition of Notch1, *J. Mol. Histol.* 49 (2018) 411–417.
- [49] X. Wang, N.L. Buechler, A.G. Woodruff, D.L. Long, M. Zabalawi, B.K. Yoza, C.E. McCall, V. Vachharajani, Sirtuins and immuno-metabolism of sepsis, *Int. J. Mol. Sci.* 19 (2018).