

·生物信息学·

朱思哲¹,叶成林²

(1.华中科技大学同济医学院附属同济医院临床免疫研究室,湖北 武汉 430030;
2.武汉大学人民医院病理科,湖北 武汉 430060)

摘要:目的 基于生物信息学分析构建免疫相关基因脓毒症预后模型。**方法** 从 GEO 数据库下载脓毒症相关的基因表达矩阵,基于 CIBERSORTx 数据库分析脓毒症组和正常组的免疫细胞浸润差异,通过 R 包 limma 进行差异分析;对差异基因和免疫相关基因取交集,并进行生物功能富集分析;通过 LASSO 回归、多因素 COX 回归筛选免疫相关独立预后基因,并构建脓毒症预后风险模型,通过训练集和验证集利用预后相关基因构建脓毒症诊断模型。**结果** 脓毒症组与正常组间存在 12 种表达差异的免疫细胞,共发现 245 个差异表达的免疫相关基因;生物功能富集分析显示,245 个基因主要富集在免疫和感染相关信号通路;LASSO 回归和多因素 COX 回归显示,DEFA4、CAMP、CX3CR1 和 PRKCA 可作为脓毒症的独立预后因素,其中 DEFA4 和 CAMP 在脓毒症组高表达,CX3CR1 和 PRKCA 在正常组中高表达($P<0.05$);基于 DEFA4、CAMP、CX3CR1 和 PRKCA 成功构建脓毒症风险预后模型,DEFA4、CAMP、CX3CR1 在训练集和验证集中的 ROC 曲线下面积(AUC)均 >0.750 ,PRKCA 在训练集中的 AUC 为 0.893,在验证集中为 0.673。**结论** 基于 DEFA4、CAMP、CX3CR1 和 PRKCA 构建的风险预后模型具有良好的预测能力,同时 DEFA4、CAMP、CX3CR1 和 PRKCA 也具备脓毒症诊断能力,可能成为脓毒症新的生物标志物和靶点。

关键词:脓毒症;免疫细胞浸润;预后模型

中图分类号:R631

文献标识码:A

DOI:10.3969/j.issn.1006-1959.2022.15.001

文章编号:1006-1959(2022)15-0001-07

Identification of Immune-related Gene Prognostic Model for Sepsis Based on Bioinformatics Analysis

ZHU Si-zhe¹,YE Cheng-lin²

(1.Department of Clinical Immunology,Tongji Hospital,Tongji Medical College,Huazhong University of Sciences and Technology,Wuhan 430030,Hubei,China;

2.Department of Pathology,Renmin Hospital of Wuhan University,Wuhan 430060, Hubei,China)

Abstract: Objective To construct a prognostic model for immune-related genes in sepsis based on bioinformatics analysis.**Methods** The sepsis-related gene expression matrix was downloaded from GEO database. The difference of immune cell infiltration between sepsis group and normal group was analyzed based on CIBERSORTx database, and the difference was analyzed by R package limma. The intersection of differential genes and immune-related genes was collected, and the biological function enrichment analysis was performed. The immune-related independent prognostic genes were screened by LASSO regression and multivariate COX regression, and the prognostic risk model of sepsis was constructed. The prognostic-related genes were used to construct the diagnostic model of sepsis through training set and validation set.**Results** There were 12 immune cells with different expression between sepsis group and normal group, and 245 immune-related genes with different expression were found. Biofunctional enrichment analysis showed that 245 genes were mainly enriched in immune and infection-related signaling pathways. LASSO regression and multivariate COX regression showed that DEFA4, CAMP, CX3CR1 and PRKCA were independent prognostic factors for sepsis, while DEFA4 and CAMP were highly expressed in sepsis group, CX3CR1 and PRKCA were highly expressed in normal group ($P<0.05$). The risk prognosis model of sepsis was successfully constructed based on DEFA4, CAMP, CX3CR1 and PRKCA. The area under the ROC curve (AUC) of DEFA4, CAMP and CX3CR1 in the training set and the validation set was >0.750 . The AUC of PRKCA in the training set was 0.893 and 0.673 in the validation set.**Conclusion** The risk prognosis model based on DEFA4, CAMP, CX3CR1 and PRKCA has good predictive ability. At the same time, DEFA4, CAMP, CX3CR1 and PRKCA also have the diagnostic ability of sepsis, which may become new biomarkers and targets for sepsis.

Key words:Sepsis;Immune cell infiltration;Prognostic mode

脓毒症(sepsis)是宿主对感染的免疫反应失调而导致危及生命的器官功能性障碍^[1]。据保守估计,脓毒症是全世界死亡和危重病的主要原因^[2]。尽管在过去的几十年中,脓毒症的预后有所改善,但脓毒症休克的死亡率仍高于 25%~30%,甚至 40%~50%^[3]。此外,幸存的脓毒症患者通常要承受长期的身体、心理和认知障碍^[4]。脓毒症与其他主要流行病不同,其治疗是非特异性的,没有批准的药物,只能给予维持器官功能、抗生素和保障血液动力学稳定治疗。因此,迫切需要探索更多潜在的脓毒症诊断生物标志

物并构建预后模型,从而提高治疗决策的有效性。免疫反应失调在脓毒症中扮演着重要的角色。单核细胞和巨噬细胞是先天免疫系统的主要组成部分,在脓毒症期间协调宿主免疫反应方面发挥着关键作用。在脓毒症的早期阶段,两种不同的淋巴细胞分泌的促炎因子和趋化因子水平升高,这会加剧炎症反应,增加死亡率^[5]。中性粒细胞是体循环中最丰富的白细胞。研究表明,脓毒症患者中循环中性粒细胞增加可导致免疫失调,通过释放细胞因子和活性氧导致多器官衰竭,CD4⁺T 细胞可以直接介导宿主对脓毒症的反应,脓毒症导致的细胞凋亡会诱导 CD4⁺T 细胞的耗竭^[6]。而脓毒症导致记忆 CD8⁺T 细胞表型、保护功能和定位的长期改变,可能会改变宿主对再

作者简介:朱思哲(1992.2-),男,湖北武汉人,硕士,技师,主要从事脓毒症发病机制的研究

感染的反应能力^[7]。因此,免疫相关基因在脓毒症的诊断与预后预测方面具有巨大潜力。目前,关于免疫相关基因在脓毒症的诊断与预后方面的研究报道罕见。基于此,本研究通过生物信息学分析获得具有表达差异的免疫相关基因,并探索其生物学功能。进一步通过 LASSO 回归和多因素 COX 回归分析筛选脓毒症独立预后基因,并构建预后风险模型,以及探索所筛选的基因在脓毒症中的诊断价值,为脓毒症的诊断、治疗和预后提供新的思路。

1 资料与方法

1.1 数据来源 通过 GEO 数据库(<https://www.ncbi.nlm.nih.gov/geo/>) 下载脓毒症基因芯片表达数据集 GSE134347、GSE13904 和 GSE65682,每个数据集样本信息见表 1。通过 CIBERSORTx 数据库(<https://cibersortx.stanford.edu/>) 分析 GSE134347 数据集中每个样本中 22 种常见免疫细胞的浸润程度。

表 1 GSE134347、GSE13904 和 GSE65682 数据集信息

样本量	GSE134347	GSE13904	GSE65682
正常标本	83	18	46
脓毒症标本	156	158	467
标本总数	239	178	513

1.2 差异基因分析 利用 R 包 limma 对 GSE134347 数据集进行差异分析, adjust *P*-value < 0.05, |Log₂FC| > 1 的基因被定义为差异表达基因(DEGs)。通过 ImmPort 数据库下载免疫相关基因,并与差异表达基因取交集。

1.3 生物学功能富集分析 利用 R 包 clusterProfile 进行生物学功能 GO(gene ontology)和 KEGG(kyoto encyclopedia of genes and genomes)富集分析,利用

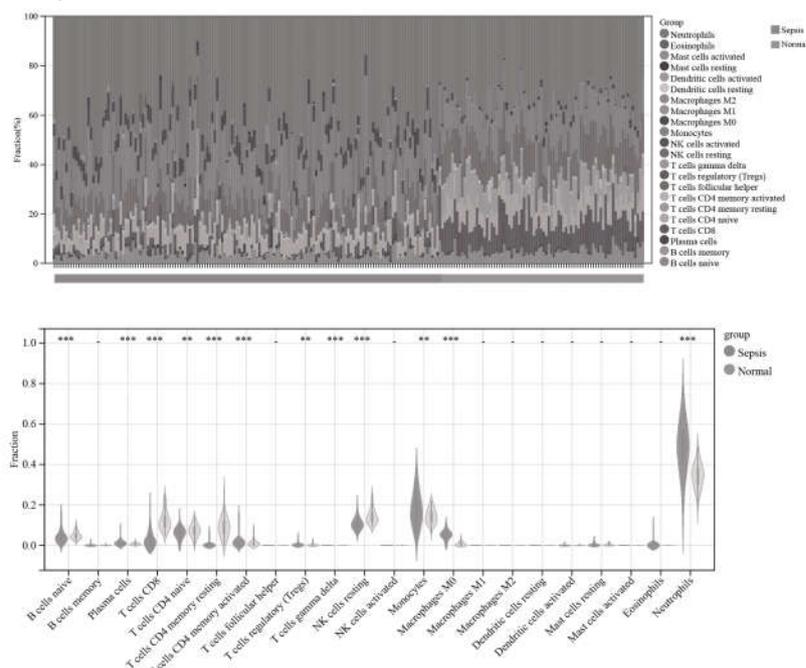
R 包 GOplot 计算 z-score。

1.4 构建风险预后模型 通过 R 包 glmnet 对 GSE65682 数据集中含有预后信息的标本进行 LASSO 回归分析,设置 10 折交叉验证,以获得最优模型。利用 R 包 survival 进行多因素 COX 回归筛选独立预后基因(*P*<0.01),风险比(Hazard ratio, *HR*)>1 为危险因素, *HR*<1 为保护因素。进一步通过多因素回归分析确定每个独立预后基因的回归系数,建立预后风险模型,预测患者生存率。公式为: 风险分数 = $\sum \beta_i \times x_i$, β_i 为基因对应的多元 COX 回归 beta 值, x_i 为基因表达值。

1.5 绘制诊断 ROC 曲线及验证 利用 R 包 pROC 对 GSE134347 训练集和 GSE13904 验证集中的独立预后基因绘制诊断 ROC 曲线和计算 AUC 值,并通过 R 包 ggplot2 绘制这两个数据集中独立预后基因表达值小提琴图。

2 结果

2.1 免疫浸润分析 免疫浸润分析显示,脓毒症组和正常组的免疫细胞浸润具有差异。在脓毒症组中,浆细胞(plasma cells)、活化 CD4⁺记忆 T 细胞(T cells CD4 memory activated)、调节性 T 细胞(T cells regulatory)、单核细胞(monocyte)、M0 型巨噬细胞(macrophage M0)和中性粒细胞(neutrophils)的浸润程度上升,而初始 B 细胞(B cells naive)、CD8⁺T 细胞(T cells CD8)、初始 CD4⁺T 细胞(T cells CD4 naive)、静息 CD4⁺记忆 T 细胞(T cells CD4 memory resting)和静息自然杀伤细胞(NK cells resting)的浸润程度下降,见图 1。



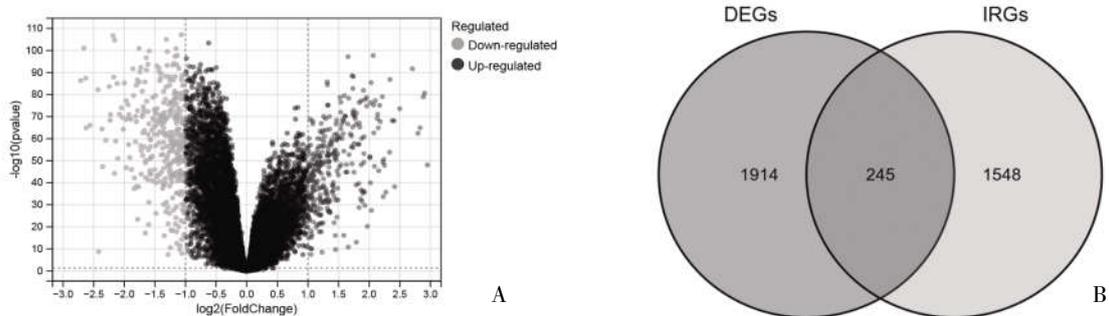
注:A:GSE134347 数据集中各标本的免疫细胞浸润比例;B:脓毒症组和正常组免疫细胞浸润比例; **P*<0.05, ***P*<0.01, ****P*<0.001

图 1 免疫浸润分析

2.2 差异分析 共获得 911 个上调基因和 1248 个下调基因,见图 2A。通过与免疫相关基因取交集,共获得差异表达的免疫相关基因 245 个,见图 2B。

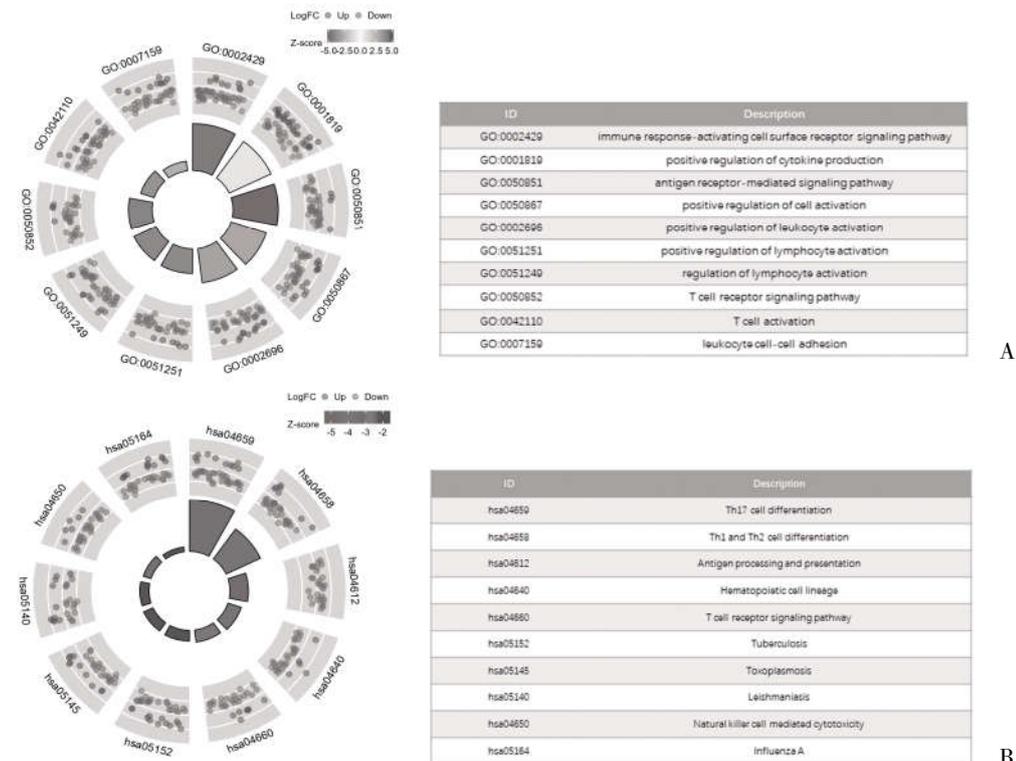
2.3 生物学功能富集分析 对筛选的 245 个免疫相关基因进行生物学富集分析:GO 富集显示上述基因主要富集于免疫反应激活细胞表面受体信号通路、细胞因子产生的正调控、抗原受体介导的信号通路、细胞活化的正调控、白细胞活化的正调控、淋巴细胞活化的正调控、淋巴细胞活化的调节、T 细胞受体信号通路、T 细胞活化、白细胞-细胞粘附,见图 3A;KEGG 信号通路富集分析显示,上述基因主要富集于 Th17 细胞分化、Th1 和 Th2 细胞分化、抗原处理和呈递、造血细胞系、T 细胞受体信号通路、肺结核、弓形虫病、利什曼病、自然杀伤细胞介导的细胞毒性和甲型流感信号通路,见图 3B。

2.4 构建风险预后模型 对 245 个免疫相关基因进行 LASSO 回归分析筛选预后基因,见图 4A;共获得 DEFA4、IL4R、TFRC、CCL5、CAMP、CX3CR1、PRKCA、CD28 和 LTB 9 个免疫相关预后基因,LASSO 回归轨迹见图 4B。进一步通过多因素 COX 回归筛选独立预后基因,DEFA4 为危险因素 ($HR>1$),而 CAMP、CX3CR1 和 PRKCA 为保护因素 ($HR<1$),见图 4C。基于以上 4 个基因的表达值与 COX 回归系数构建风险预后模型,计算 GSE65682 中每个脓毒症患者的风险分数:风险分数= $0.2300 \times DEFA4 - 0.2894 \times CAMP - 0.2958 \times CX3CR1 - 0.2388 \times PRKCA$ 。进一步按照风险分数中位数将 GSE65682 中的脓毒症患者分为高风险组和低风险组,比较两组 28 d 生存差异,结果显示高风险组 28 d 生存率低于低风险组 ($P=2.7e-7$),见图 5。



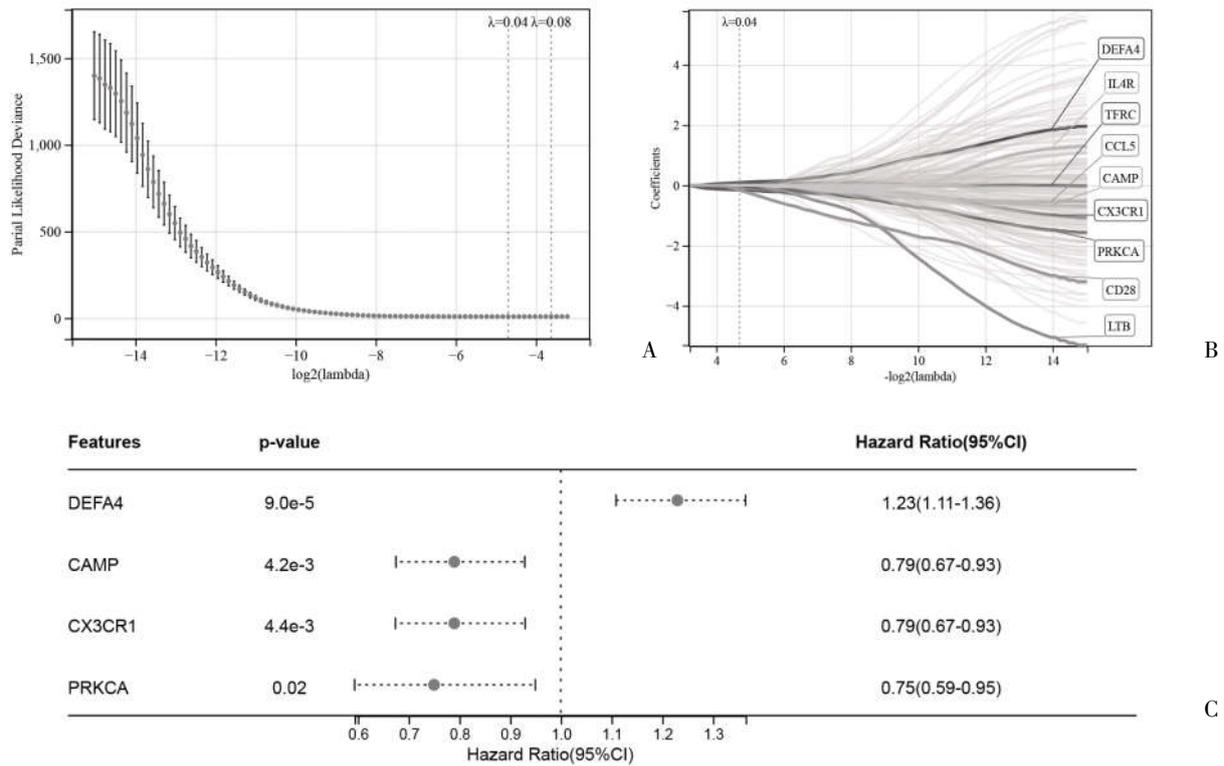
注:A:差异表达基因火山图;B:差异表达基因与免疫相关基因的韦恩图

图 2 差异基因分析



注:A:GO 富集分析;B:KEGG 富集分析

图 3 生物学功能富集分析



注 A:LASSO 回归系数筛选;B:LASSO 回归轨迹图;C:多因素 COX 回归森林图

图 4 LASSO 回归和多因素 COX 回归分析

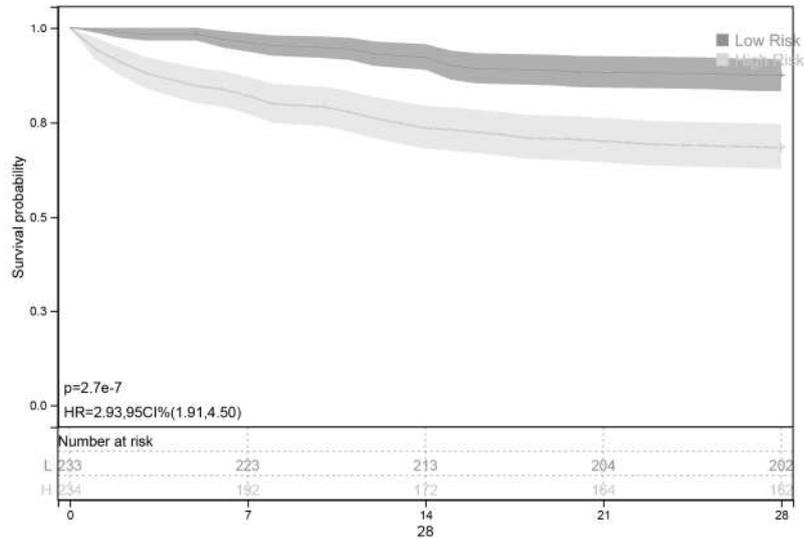
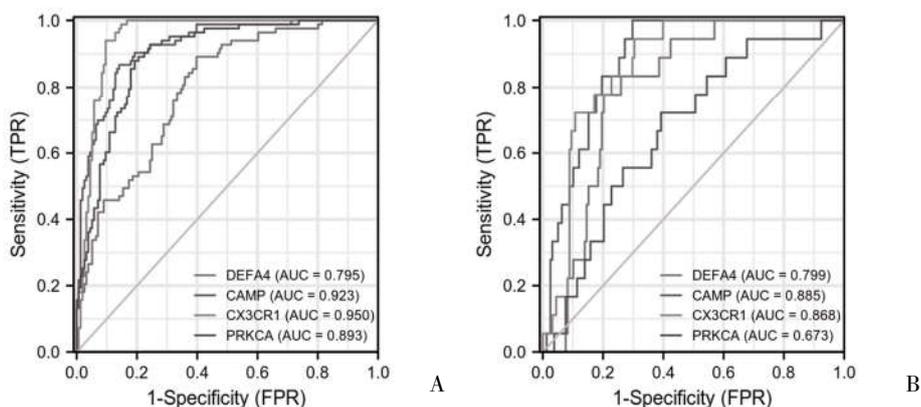


图 5 生存分析

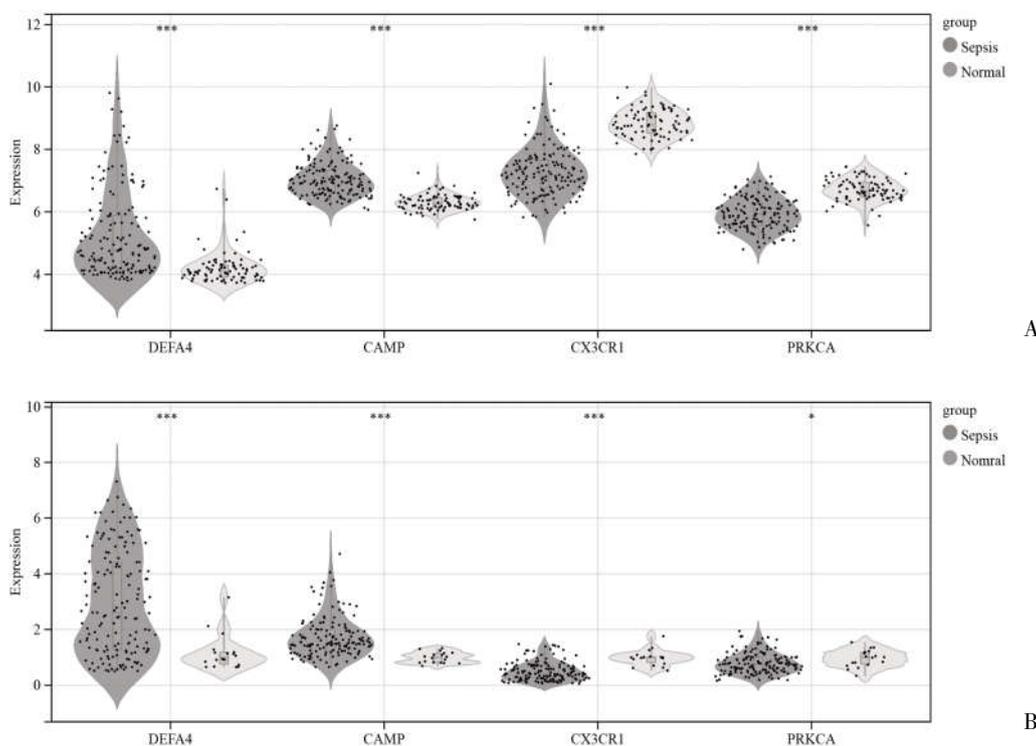
2.5 诊断 ROC 曲线及验证 ROC 曲线显示,DEFA4 (AUC=0.795)和 PRKCA(AUC=0.893)展示出一定的诊断准确性,而 CAMP (AUC=0.923)和 CX3CR1 (AUC=0.950)则具有较高的诊断准确性,见图 6A。为了进一步验证上述结果,利用 GSE13904 验证集绘制 DEFA4、CAMP、CX3CR1 和 PRKCA 的 ROC 曲线,结果显示 DEFA4、CAMP 和 CX3CR1 展示出一

定的诊断准确性,而 PRKCA 则具有较低的准确性,见图 6B。进一步根据 DEFA4、CAMP、CX3CR1 和 PRKCA 在 GSE134347 训练集与 GSE13904 验证集中的表达值,分别绘制表达值小提琴,结果显示 DEFA4 和 CAMP 的表达在脓毒症组上调 ($P < 0.05$),CX3CR1 和 PRKCA 的表达则在正常组中上调 ($P < 0.05$),见图 7。



注:A:GSE134347 训练集 ROC 曲线;B:GSE13904 验证集 ROC 曲线

图 6 该模型诊断脓毒症的 ROC 曲线



注:A:GSE134347 训练集基因表达小提琴图;B:GSE13904 验证集基因表达小提琴图

图 7 DEFA4、CAMP、CX3CR1 和 PRKCA 基因表达水平

3 讨论

在过去的 30 年里,脓毒症的定义和管理有很大的发展^[8]。这得益于对临床综合征的理解、血流动力学监测工具的进步和抢救措施改进,但脓毒症仍是危重患者发病和死亡的主要原因之一^[9]。多数患者在脓毒症的初始高炎症阶段可存活下来,但会进展到后期的免疫抑制阶段,其中 30% 的患者死于继发感染,免疫反应则可能在脓症患者死亡率中起主要作用^[10]。因此,探索免疫相关基因在脓毒症中的作用,筛选免疫相关基因的脓毒症诊断生物标志物和构建预后风险模型对脓毒症的早期诊断和预后评估

具有重要意义。

本研究对 GEO 数据集的脓毒症组和正常组进行免疫浸润分析,结果显示脓毒症组与正常组外周血中免疫细胞浸润程度具有差异,这说明免疫调节在脓毒症中起重要作用。为了进一步探索免疫调节在脓毒症中的作用,本研究对脓毒症组和正常组进行差异分析,并从 ImmPort 数据库下载免疫相关基因与差异基因取交集,共获得 245 个免疫相关差异表达基因。生物功能富集分析显示 245 个基因主要富集于 Th17 细胞分化、Th1 和 Th2 细胞分化、肺结核以及弓形虫病等免疫和感染相关信号通路。多项

研究表明,这些信号通路与脓毒症密切相关。如 Mishra R 等^[11]报道肺结核性脓毒症休克是一种罕见的疾病,但其具有非常高的死亡率,这与败血症流行病学普遍预期相反。Souza MC 等^[12]报道,慢性弓形虫感染可加剧继发性多微生物脓毒症。Th 细胞作为适应性免疫的一部分,参与了脓毒症的免疫调节。脓症患者中 Th17 和 Th1 细胞的活化失调,以及 Th17/Treg 比例失衡与脓毒症预后相关^[13,14]。Xue M 等^[15]报道,社区获得性严重脓毒症患发作时, Th2/Th1 比例显著上升, Th2 占主导地位,其持续动态增加与 ICU 获得性感染和 28 d 死亡率有关。

为了进一步筛选免疫相关脓毒症预后基因,本研究对上述 245 个基因进行 LASSO 回归和多因素 COX 回归,结果显示 DEFA4、CAMP、CX3CR1 和 PRKCA 可以作为脓毒症的独立预后基因。DEFA4 又名 HNP4(人中性粒细胞 α 防御素 4),编码一种先天免疫阳离子抗菌肽,对革兰氏阴性菌具有较强的抗菌效果,也能保护人外周血单核细胞免受 HIV-1 感染^[16]。CAMP 编码一种 Cathelicidin 抗菌肽 LL-17,是呼吸道内宿主防御细菌感染中的重要元素^[17],并且可以抑制 2 型登革热病毒的感染^[18]。CX3CR1 是趋化因子 fractalkine 的受体, CX3CR1+ 细胞可以阻止沙门氏菌的侵袭^[19],并在控制可能导致组织损伤的异常肠道炎症方面具有重要作用^[20]。PRKCA 编码蛋白激酶 C 属于丝氨酸和苏氨酸特异性蛋白激酶家族,蛋白激酶 C 参与调节先天免疫和适应性免疫,参与 2 型先天淋巴细胞的活化和 Th2 细胞的活化与增殖^[21]。基于上述 4 个基因构建脓毒症的预后风险模型,结果显示高风险组 28 d 生存率低于低风险组。由此推测 DEFA4、CAMP、CX3CR1 和 PRKCA 在脓毒症的免疫调节中起到关键作用,且与预后紧密相关。ROC 曲线显示, DEFA4、CAMP 和 CX3CR1 在训练集和验证集中均表现出一定的诊断能力(AUC>0.75),而 PRKCA 在验证集中 AUC 为 0.673,说明 DEFA4、CAMP、CX3CR1 和 PRKCA 可以成为脓毒症免疫相关生物标志物。本研究结果提示, DEFA4、CAMP、CX3CR1 和 PRKCA 可能在脓毒症的免疫调节中起到重要的作用,并且与预后紧密相关。探索脓毒症的早期诊断生物标志物以及构建预后模型有助于给予脓症患者及时有效的治疗,同时提高患者生存率。

综上所述,本研究基于 4 个免疫相关的独立预

后基因 DEFA4、CAMP、CX3CR1 和 PRKCA,构建了脓毒症预后风险模型,其在脓毒症预后预测中具有一定的应用价值,并且上述 4 个基因都具有一定的脓毒症诊断能力。

参考文献:

- [1]Fernando SM,Rochwerg B,Seely AJE.Clinical implications of the Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3)[J].CMAJ,2018,190(36):E1058-E1059.
- [2]Rudd KE,Johnson SC,Agesa KM,et al.Global, regional, and national sepsis incidence and mortality, 1990-2017: analysis for the Global Burden of Disease Study [J].Lancet,2020,395(10219): 200-211.
- [3]Venkatesh B,Finfer S,Cohen J,et al.Adjunctive Glucocorticoid Therapy in Patients with Septic Shock [J].N Engl J Med, 2018,378(9):797-808.
- [4]Barichello T,Sayana P,Giridharan VV,et al.Long -Term Cognitive Outcomes After Sepsis: a Translational Systematic Review[J].Mol Neurobiol,2019,56(1):186-251.
- [5]Hamidzadeh K,Christensen SM,Dalby E,et al.Macrophages and the Recovery from Acute and Chronic Inflammation [J].Annu Rev Physiol,2017,79:567-592.
- [6]Cao C,Yu M,Chai Y.Pathological alteration and therapeutic implications of sepsis -induced immune cell apoptosis [J].Cell Death Dis,2019,10(10):782.
- [7]Jensen IJ,Li X,Mcgonagill PW,et al.Sepsis leads to lasting changes in phenotype and function of memory CD8 T cells[J].Elife,2021,10:e70989.
- [8]Gyawali B,Ramakrishna K,Dhamoon AS.Sepsis: The evolution in definition, pathophysiology, and management [J].SAGE Open Med,2019,7:2050312119835043.
- [9]Cecconi M,Evans L,Levy M,et al.Sepsis and septic shock [J].Lancet,2018,392(10141):75-87.
- [10]Brady J,Horie S,Laffey JG.Role of the adaptive immune response in sepsis[J].Intensive Care Med Exp,2020,8(Suppl 1):20.
- [11]Mishra R,Patel HK,Singasani R,et al.Tuberculosis septic shock, an elusive pathophysiology and hurdles in management: A case report and review of literature[J].World J Crit Care Med, 2019,8(5):72-81.
- [12]Souza MC,Fonseca DM,Kanashiro A,et al.Chronic Toxoplasma gondii Infection Exacerbates Secondary Polymicrobial Sepsis[J].Front Cell Infect Microbiol,2017,7:116.
- [13]Coakley JD,Breen EP,Moreno-Olivera A,et al.Dysregulated T helper type 1 (Th1) and Th17 responses in elderly hospitalised patients with infection and sepsis [J].PLoS One,2019,14 (10): e0224276.

(下转第 13 页)

(上接第 6 页)

- [14]Guo J,Tao W,Tang D,et al.Th17/regulatory T cell imbalance in sepsis patients with multiple organ dysfunction syndrome: attenuated by high-volume hemofiltration [J].Int J Artif Organs,2017,40(11):607-614.
- [15]Xue M,Xie J,Liu L,et al.Early and dynamic alterations of Th2/Th1 in previously immunocompetent patients with community-acquired severe sepsis:a prospective observational study [J].J Transl Med,2019,17(1):57.
- [16]Hu H,Di B,Tolbert WD,et al.Systematic mutational analysis of human neutrophil alpha-defensin HNP4[J].Biochim Biophys Acta Biomembr,2019,1861(4):835-844.
- [17]Majewski K,Zelechowska P,Brzezinska-Blaszczyk E.Circulating cathelicidin LL-37 in adult patients with pulmonary infectious diseases[J].Clin Invest Med,2017,40(1):E34-E39.
- [18]Alagarasu K,Patil PS,Shil P,et al.In-vitro effect of human

- cathelicidin antimicrobial peptide LL-37 on dengue virus type 2 [J].Peptides,2017,92:23-30.
- [19]Man AL,Gicheva N,Regoli M,et al.CX3CR1+ Cell-Mediated Salmonella Exclusion Protects the Intestinal Mucosa during the Initial Stage of Infection[J].J Immunol,2017,198(1):335-343.
- [20]Marelli G,Belgiovine C,Mantovani A,et al.Non-redundant role of the chemokine receptor CX3CR1 in the anti-inflammatory function of gut macrophages [J].Immunobiology,2017,222(2):463-472.
- [21]Madouri F,Chenuet P,Beuraud C,et al.Protein kinase C theta controls type 2 innate lymphoid cell and TH2 responses to house dust mite allergen[J].J Allergy Clin Immunol,2017,139(5):1650-1666.

收稿日期:2022-01-16;修回日期:2022-02-18

编辑/成森