

基于单细胞测序数据分析晚期脓毒血症生物信息学特点

王润青¹, 杨刚¹, 何海珊¹, 尹艺洁¹, 李思佳¹, 黄叙琪¹, 廖健雄¹, 杨经文^{1,2}

(1. 广州医科大学附属清远医院(清远市人民医院)第六临床学院, 广东 清远 511518; 2. 广州医科大学附属清远医院(清远市人民医院)重症医学科, 广东 清远 511518)

摘要 **目的:** 基于单细胞转录组测序揭示晚期脓毒血症的生物信息学特点, 为脓毒血症的药物设计提供生物信息学参照。**方法:** 从GEO数据库获得数据集(GSE175453)的scRNA-seq数据集中的4例晚期脓毒血症血液样本, 采用Seurat包筛选每个细胞群体的差异表达基因。随后, 对其进行生物信息学分析, 包括基因本体论(GO)和基因组百科全书(KEGG)功能富集分析, 并对数据进行拟时序分析及细胞通讯分析。**结果:** 4例晚期脓毒血症血液样本共可得到19 117个细胞, 7个细胞群体。GO分析揭示差异基因主要涉及白细胞介导的免疫反应、分泌颗粒腔及免疫受体活性; KEGG富集分析的结果表明: 差异基因主要参与造血细胞系的生物学功能。单核细胞在晚期脓毒血症演变中起着重要作用, 拟时序结果显示, 单核细胞共具有3个分化命运, 1个分化节点。细胞通讯结果表明, 巨噬细胞迁移抑制因子(MIF)介导的信号通路在晚期脓毒血症细胞通讯中具有最高的发生率。**结论:** 本研究揭示了晚期脓毒血症各细胞群体以及差异基因参与的生物信息学过程, 揭示了MIF通路可能为治疗脓毒血症的潜在靶点。

关键词 单细胞测序; 生物信息学; 脓毒血症; 巨噬细胞迁移抑制因子

中图分类号: R459.7 **文献标识码:** A **文章编号:** 2095-9664(2024)04-0001-07

Analyzing the bioinformatics features of late-stage sepsis based on single-cell sequencing data

WANG Runqing¹, YANG Gang¹, HE Haishan¹, YIN Yijie¹, LI Sijia¹, HUANG Xuqi¹, LIAO Jianxiong¹, YANG Jingwen^{1,2}

(1. Affiliated Qingyuan Hospital, The Sixth Clinical Medical School, Guangzhou Medical University, Qingyuan People's Hospital, Qingyuan 511518, Guangdong, China; 2. Department of Critical Care Medicine, Affiliated Qingyuan Hospital, Guangzhou Medical University, Qingyuan People's Hospital, Qingyuan 511518, Guangdong, China)

Corresponding author: YANG Jingwen, Email: Ywenfeng2003@163.com

Abstract **Objective:** Using single-cell transcriptome sequencing, this study aimed to delineate the bioinformatics characteristics of late-stage sepsis, providing a reference for drug design targeting sepsis. **Methods:** The single-cell RNA sequencing data from four late-stage sepsis blood samples were retrieved from the GEO database (GSE175453). Differentially expressed genes in each cell cluster were identified using the Seurat package. Subsequent bioinformatics analyses included Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses, as well as pseudotime analysis and cell communication analysis. **Results:** In four late-stage sepsis blood samples, a total of 19 117 cells were categorized into 7 cell clusters. GO analysis revealed that differentially expressed genes are primarily involved in leukocyte-mediated immunity, secretion granule lumen, and immune receptor activity. KEGG enrichment analysis indicated these genes primarily participate in hematopoietic cell lineage functions. Monocytes play a significant role in the progression of late-stage sepsis, with pseudotime analysis showing three differentiation states and one differentiation node. Cell communication analysis highlighted that the macrophage migration inhibitory factor (MIF)-mediated signaling pathway has the highest occurrence probability in cell

communication during late-stage sepsis. **Conclusion:** This study elucidates the bioinformatics processes involving various cell types and differential gene involvement in late-stage sepsis, revealing the MIF pathway as a potential therapeutic target for treating sepsis.

Key words single-cell RNA-sequencing; Bioinformatics; sepsis; macrophage migration-inhibitory factors

脓毒血症是一种极其容易危及生命,具有高致残率、致死率的器官功能障碍综合征,其致残率、致死率极高,病理生理学特点复杂,病情变化快,2017年,全球每100 000人口即有677.5例脓毒血症患者,且以每年9%的速度增长,大约2%的患者以及高达75%的重症监护病房患者发生脓毒症,其中30%发展为脓毒性休克^[1]。大量研究表明:脓毒血症发病是由多种免疫细胞、免疫介质、细胞因子共同参与的复杂的动态病理过程,脓毒血症早期以细胞因子风暴为代表的过度免疫反应为主,而在脓毒症后期逐渐发展为免疫抑制、持续性炎症和分解代谢综合征,因此,脓毒血症晚期的治疗十分困难^[2],尽管当前已有大量用于脓毒血症临床评估的分子标记物,然而,这些分子标记物的特异性仍不尽理想。脓毒血症的发生发展与外周血单个核细胞,包括T细胞、单核细胞、NK细胞以及淋巴细胞等密切相关,在疾病早期,外周血免疫细胞可识别PAMPs并释放大量细胞因子,引发“免疫因子风暴”,而晚期脓毒血症多种外周血免疫细胞功能受损,且炎症因子大量释放入血,表现为显著的免疫抑制。实际上,脓毒血症的发病机制极其复杂,了解疾病发生发展中的分子机制是充分了解生物学标记物和进行特异性治疗的必要前提^[3],因此,本文主要基于单细胞测序探讨晚期脓毒血症细胞群和相关差异基因的生物过程,旨在为了解脓毒血症的生物学机制和药物靶点设计提供参考。

1 资料与方法

1.1 研究资料

运用NCBI's Gene Expression Omnibus(GEO)数据库(<https://www.ncbi.nlm.nih.gov/geo/>)。

1.2 研究方法

1.2.1 晚期脓毒血症的scRNA-seq数据下载 以“sepsis”和“single cell”为检索词,在GEO数据库下载单细胞转录组表达数据,共纳入来源于GEO数据库编号为GSE175453^[4]的数据集的4例晚期脓毒血症患者血液样本的scRNA-seq数据。

1.2.2 scRNA-seq的数据整合与细胞注释 使用R

语言(4.2.3)Seurat包读取10×数据并导入程序,分别得到各个样本的单细胞转录组表达矩阵,使用Harmony包对各个样本数据进行整合,并采用Seurat包去除批次效应,将基因表达数>200且线粒体基因比例<15%的数据纳入研究。运用Scale法对数据进行归一化处理,随后进行主成分分析降维(PCA)以及UMAP可视化降维,采用SingleR包对细胞进行自动化注释后,使用plot1cell包进行可视化。

1.2.3 各细胞群体差异基因的富集分析 采用Seurat包的FindAllMarkers函数筛选每个细胞群体最显著的50个差异基因(differentially expressed genes, DEGs),使用clusterProfiler、enrichplot程序包并分别对其对进行基因本体论(genetic ontology, GO)分析,注释其基因的分子功能(molecular function, MF)、生物过程(biological process, BP)和细胞成分(cellular component, CC),在3个层面分别展示了前5的富集GO功能;进行基因组百科全书(kyoto encyclopedia of genes and genomes, KEGG)对差异表达基因进行富集分析,取显著富集的前15条通路进行可视化。

1.2.4 晚期脓毒血症各细胞群体的拟时序分析 选取注释后的细胞群体,使用Monocle3包进行不同亚群细胞的拟时序轨迹拟合分析,且使用graph_test函数对数据集内的所有基因进行空间差异分析,计算莫兰指数(Moran's I),Moran's I>0表明该基因具有空间正相关性,且Moran's I越接近于1,则认为其空间正相关性越大。使用Monocle包对相应细胞进行拟时序轨迹分析,以探索细胞的分化轨迹。

1.2.5 晚期脓毒血症细胞群体间的细胞通讯分析 使用Cellchat包内置的CellChatDB数据库进行受体-配体比对与计算,从而分别对细胞间的通讯频率与强度、细胞间通讯的概率、相关通路的网络集中度以及信号转导关键基因进行可视化分析。

1.3 统计分析

所有数据均使用R软件4.2.3版本进行分析,Wilcoxon秩和检验用于分析各细胞群体的差异基因,Fisher精确检验用于差异基因的GO、KEGG富集分析的显著性分析,置换检验用于各细胞群体受体-

配体对的差异性分析, $P < 0.05$ 认为差异有统计学意义。

2 结果

2.1 晚期脓毒血症的单细胞图谱和标志基因

经过数据处理后共得到 19 117 个细胞, 7 个主要的细胞亚群(图 1A)。应用热图展示各细胞亚群的高表达基因(图 1B)。

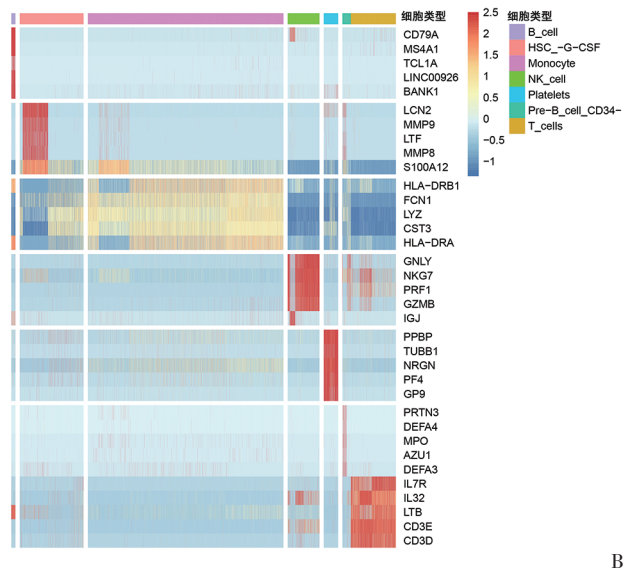
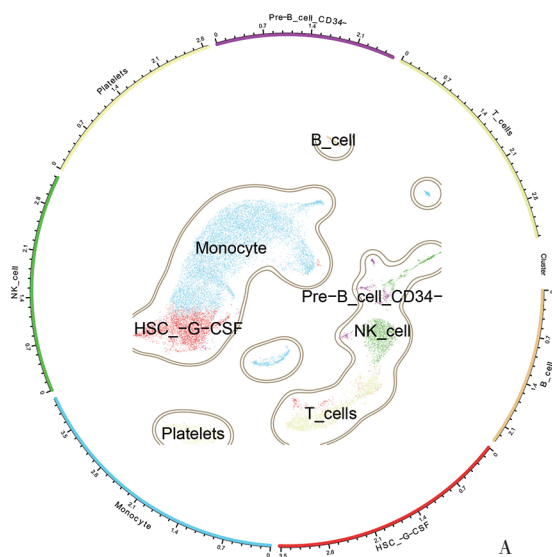
2.2 晚期脓毒血症各细胞群体差异基因的富集

分析

在 BP 中, DEGs 主要与白细胞介导的免疫反应等功能密切相关; 在 CC 中, 主要与分泌颗粒腔等密切相关; 在 MF 中, 主要与免疫受体活性密切相关(图 2A)。而 KEGG 富集分析的结果表明: 差异基因主要参与造血细胞系的生物学功能等(图 2B)。

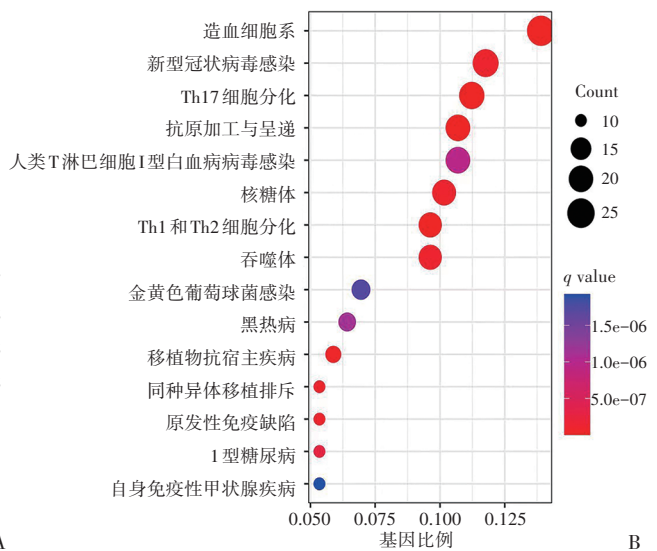
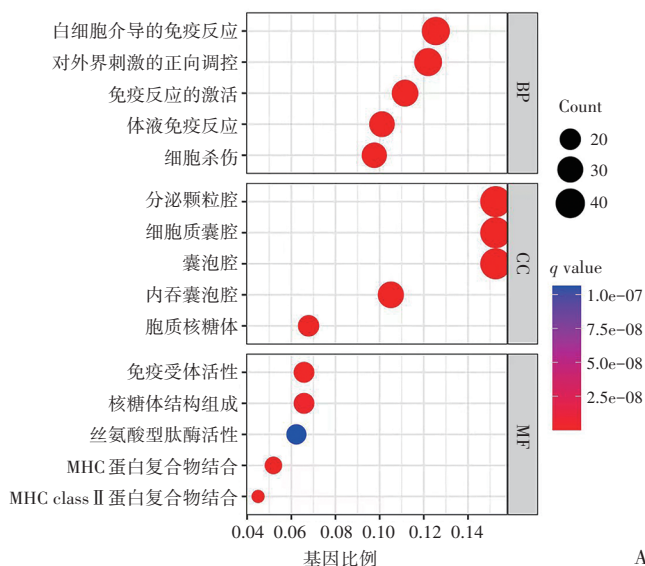
2.3 晚期脓毒血症的拟时序分析

拟时序结果表明: 随拟时间进展, CST3、LYZ、S100A6、S100A8、S100A9、TYROBP 表达量逐渐增



注: A: 4 例晚期脓毒血症血液样本的细胞组成; B: 7 种细胞群体的标志基因热图。

图 1 单细胞图谱与标志性基因热图



注: A: GO 富集分析中前 5 显著的生物过程、细胞成分以及分子功能; B: KEGG 富集分析的前 15 条显著的富集通路。

图 2 差异基因的 GO 以及 KEGG 富集分析

加,IFITM1、IL7R、RPL10逐渐降低,基因的表达与拟时间的进展呈现较强的共变趋势(图3A)。Monocle3的拟时序轨迹分析提示单核细胞在演变过程中有分支形成,提示单核细胞在脓毒血症晚期演变中起着重要作用,Monocle包的可视化结果表明,单核细胞具有3种分化状态以及1个分化节点,单核细胞由分化状态1分别向分化状态2和分化状态3分化(图3B、图3C)。

2.4 晚期脓毒血症的细胞通讯分析

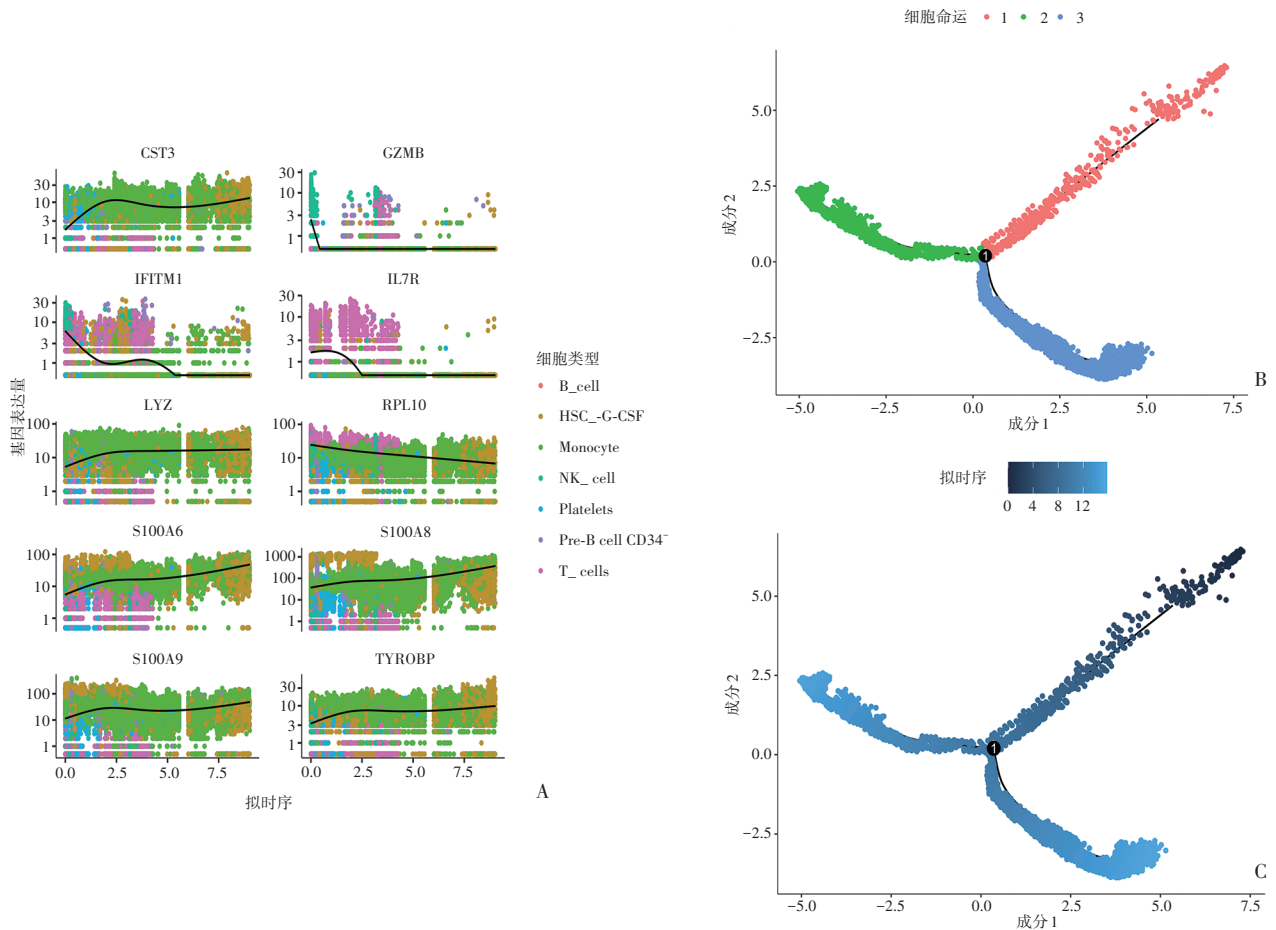
Cellchat包计算结果表明,单核细胞和T细胞、单核细胞与B细胞、T细胞和NK细胞间具有显著的交互强度与频率(图4A);MIF(巨噬细胞迁移抑制因子)介导的信号通路发生的概率最高,且主要集中于B细胞-B细胞、B细胞-单核细胞、T细胞-T细胞、T细胞-单核细胞及单核细胞-单核细胞通讯中(图4B),B细胞、T细胞和单核细胞在MIF信号通路的产生和传导中起着最为关键的作用(图4C、图4D)。

3 讨论

脓毒血症是由宿主对感染因素的免疫反应调节引起的综合征,进入脓毒症晚期后,可继而演变为多器官功能障碍综合征,甚至引起死亡^[5]。

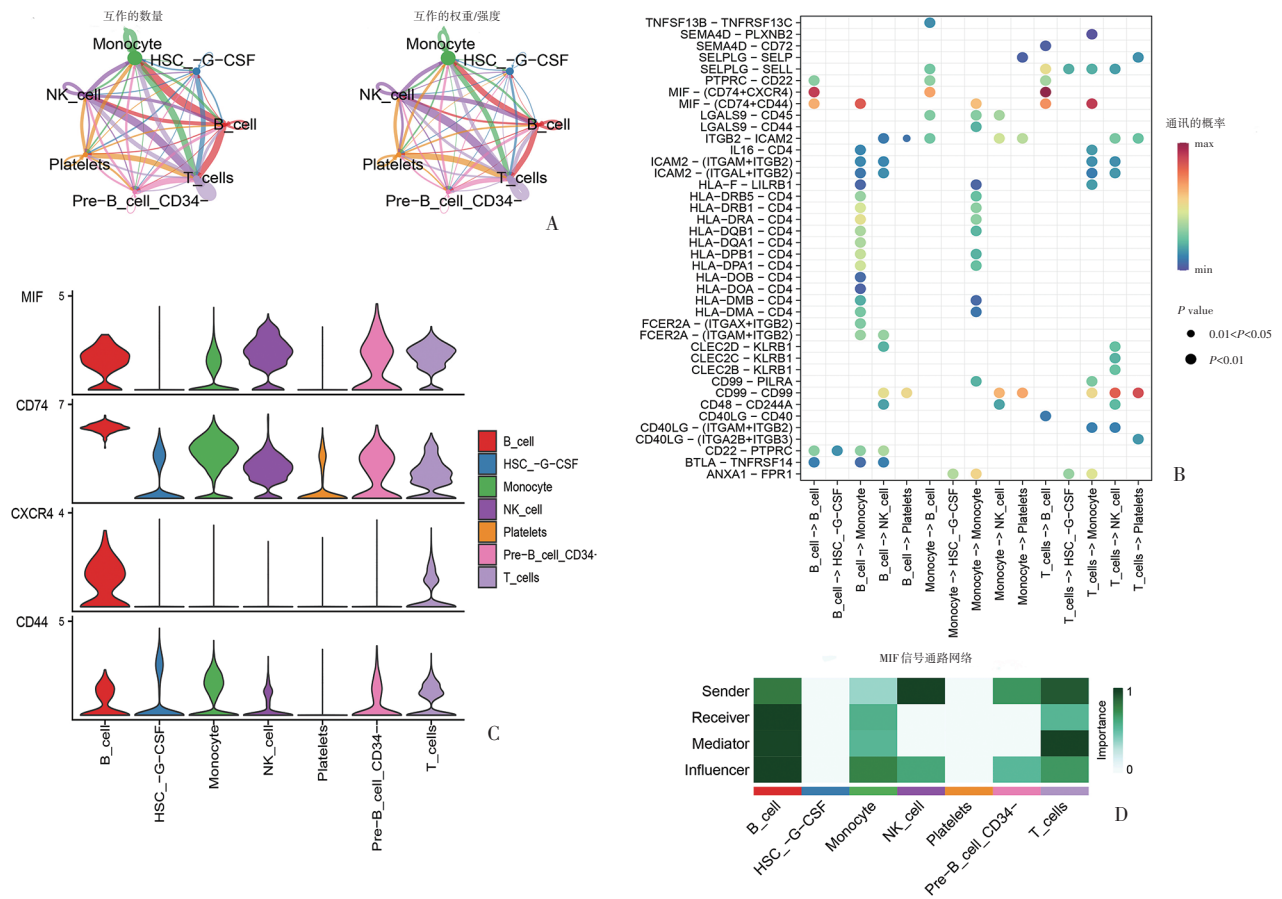
脓毒血症缺乏临床特异性、具有复杂的病理生理学特点^[6],目前,已有学者将脓毒血症归纳为:全身反应综合征。在脓毒症早期,病原体触发免疫系统产生过度炎症反应^[7];而进入脓毒症后期,则表现为显著的免疫抑制、器官衰竭^[8-10]。尽管当前已有170多种生物标记物用于脓毒血症的临床评估^[11],然而学界对于这些分子标记物的认识仍不充足,相关分子机制仍未明确,而生物信息学分析为正确认识脓毒血症的分子机制、设计针对脓毒血症的靶点指明了前景^[12]。

当前针对脓毒血症的生物信息学研究多集中于Bulk RNA-seq层次,然而其分辨率较低,且掩盖了细胞群内转录组的异质性,无法对单细胞层次的



注:A:莫兰指数显著的基因沿拟时序演变图;B,C:单核细胞拟时序轨迹图。

图3 基因沿拟时序演变图以及单核细胞拟时序轨迹图



注:A:7种细胞群体的细胞互作频率与强度;B:细胞通讯中不同受体-配体对和信号通路的发生概率;C:MIF通路相关基因在7个细胞群体的表达;D:MIF通路的网络集中度。

图4 细胞通讯分析

基因进行定量分析^[13-15]。本研究立足于单细胞测序,对单个细胞中的mRNA进行基因定量以及差异分析、通路和功能的富集分析,在一定程度上减少样本量少以及细胞异质性强的局限性,有助于理解晚期脓毒血症演变过程中的表达调控机制。

关于晚期脓毒血症的单细胞测序,基于GEO数据库GSE175453数据集的原始研究已进行部分的探索,该研究指出:晚期脓毒血症的非髓系细胞的转录组谱表现为显著的免疫抑制性以及低水平的促炎特性,且感染的微生物类型可显著改变晚期脓毒血症的淋巴细胞转录组谱系,提示对脓毒血症患者进行个体化免疫治疗的必要性^[4]。然而上述研究局限于关于脓毒血症血液样本的免疫谱系研究,对晚期脓毒血症中免疫细胞生物学过程缺乏了解,仍需要进一步的功能研究来阐明差异基因的具体作用。

本研究相较于原始研究的新颖之处在于:首

先,通过GO和KEGG富集分析明确了各细胞群基因的差异基因参与的生物信息学过程,为正确了解晚期脓毒血症生物学特点提供参考;其次,通过拟时序分析揭示了与细胞分化密切相关的相应基因,揭示了单核细胞的分化轨迹;且通过对细胞通讯的生物信息学探索,揭示了脓毒血症相关MIF信号通路的显著活跃,为脓毒血症的药物设计以及个性化治疗提供生物信息学依据。

本研究对GEO数据库中的GSE175453数据集的共4例scRNA-seq血液样本进行生物信息学分析,可得晚期脓毒血症患者血液样本存在19117个细胞,7个主要的细胞亚群,单核细胞群具有最多的细胞数量,共10303个细胞。

对7个细胞亚群各50个差异基因进行GO和KEGG富集分析。GO富集分析结果表明:在BP中,差异基因主要参与白细胞介导的免疫反应、对外界刺激的正向调控、免疫反应的激活、体液免疫反应、

细胞杀伤;在CC中,差异基因与分泌颗粒腔、细胞质囊腔、囊泡腔、内吞囊泡腔、胞质核糖体密切相关;MF中,差异基因与免疫受体活性、核糖体结构组成、丝氨酸型肽酶活性相关。而KEGG富集分析表明:差异基因主要与造血细胞系、新型冠状病毒感染、Th17细胞分化、抗原加工与呈递、人类T淋巴细胞I型白血病病毒感染等通路激活密切相关。

拟时序结果表明:CST3、LYZ、S100A6、S100A8、S100A9、TYROBP、IFITM1、IL7R、RPL10、GZMB具有较显著的空间正相关性,且随拟时间进展,CST3、LYZ、S100A6、S100A8、S100A9、TYROBP表达量逐渐增加,IFITM1、IL7R、RPL10逐渐降低,基因的表达与拟时间的进展呈现较强的共变趋势,且样本中的单核细胞在晚期脓毒血症演变中起着关键作用,在拟时序发展过程中,单核细胞存在3种分化状态,由状态1向状态2、3过渡。

细胞通讯结果表明:单核细胞与T细胞、单核细胞与B细胞以及T细胞与NK细胞间具有最显著的联系,在7个细胞亚群的信号通路中,MIF通路的发生概率最高,且B细胞、T细胞、单核细胞在MIF介导的通路信号的产生和传导中起着关键的作用。在本研究中,我们观察到了MIF通路在脓毒血症的显著活跃,且MIF相关基因,包括MIF、CD74、CXCR4、CD44在多种免疫细胞的相对高表达,提示MIF通路在晚期脓毒血症的病理生理变化中至关重要,这与前人所做的脓毒血症相关MIF信号通路的研究结果相符,因而本研究一定程度上验证了MIF确在脓毒血症中介导信号通路中起重要作用。

MIF作为一种前炎性因子,可以结合并激活受体CD74/CD44、CXCR2、CXCR4、CXCR7等,在多种人类疾病发生发展起着关键作用^[16-17]。MIF可抑制巨噬细胞的迁移,诱导其局部浸润;而糖皮质激素,作为脓毒血症抗感染治疗的常用药物,可通过以下机制调控抗炎反应:糖皮质激素通过作用于膜联蛋白A1(Annexin 1),促进其表达上调,而Annexin 1作为一种抗炎蛋白,可特异性作用于磷脂酶2 α (cPLA-2 α),抑制其激活,进一步抑制花生四烯酸释放和前列腺素的转化过程,然而,MIF可抑制上述通路的激活,抑制糖皮质激素的抗炎效应^[18]。在多个小鼠脓毒休克模型中,MIF的增加导致全身性免疫反应的发生以及死亡率的显著提高,而采用抗MIF抗体可以减少炎症因子的产生,减少组织以及器官损伤,显著降低死亡率^[19]。细胞因子CD74、CD44、CXCR4

在MIF通路中起着关键作用,其中,CD74本身即可作用于蛋白激酶,放大炎症信号;此外,MIF、CD74、CD44、SRC酪氨酸激酶间可形成信号复合物^[17],CXCR4也为MIF潜在的功能性受体,在脓毒血症的病理生理过程中,MIF因子作用于CD74、CD44、CXCR4等细胞因子,促进MIF通路激活,从而抑制巨噬细胞的迁移,且将白细胞募集至炎症病变部位,导致大量IL-1 β 、TNF- α 等炎症因子“级联”性释放,导致炎症部位缺血坏死,可诱发脓毒血症相关性急性呼吸窘迫综合征(ARDS)、脓毒血症相关性肝损伤的发生与发展^[21]。基于上述研究基础,部分学者已尝试开发靶向MIF的通路药物,为脓毒血症的防治指明了前景^[22]。

本研究亦存在以下不足:由于单细胞测序价格昂贵,因此公共数据库相关的开源数据较少,因此研究受到测序样本量的限制,仍有待更大测序样本进行验证,以期探索更为完整的脓毒血症生物信息学过程,且关于MIF通路相关分子开展的临床研究仍较少,其疗效和生物标记物意义仍需要进一步观察、判断。

参考文献

- [1] 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.
- [2] Zhang YY, Ning BT. Signaling pathways and intervention therapies in sepsis [J]. *Signal Transduct Target Ther*, 2021, 6(1): 407.
- [3] Van Der Poll T, Van De Veerdonk FL, Scicluna BP, et al. The immunopathology of sepsis and potential therapeutic targets [J]. *Nat Rev Immunol*, 2017, 17(7): 407-420.
- [4] Darden DB, Dong X, Brusko MA, et al. A Novel Single Cell RNA-seq Analysis of Non-Myeloid Circulating Cells in Late Sepsis [J]. *Front Immunol*, 2021, 12: 696536.
- [5] Evans L, Rhodes A, Alhazzani W, et al. Surviving sepsis campaign: international guidelines for management of sepsis and septic shock 2021 [J]. *Intensive Care Med*, 2021, 47(11): 1181-1247.
- [6] Leligowicz A, Matthay MA. Heterogeneity in sepsis: new biological evidence with clinical applications [J]. *Crit Care*, 2019, 23(1): 80.
- [7] Van Der Poll T, Shankar-Hari M, Wiersinga WJ. The immunology of sepsis [J]. *Immunity*, 2021, 54(11): 2450-2464.

- [8] Liu D, Huang SY, Sun JH, et al. Sepsis - induced immunosuppression: mechanisms, diagnosis and current treatment options[J]. *Mil Med Res*, 2022, 9(1): 56.
- [9] Wood AJ, Vassallo AM, Ruchaud-Sparagano MH, et al. C5a impairs phagosomal maturation in the neutrophil through phosphoproteomic remodeling [J]. *JCI Insight*, 2020, 5(15):e137029.
- [10] 黄鹤,田昭涛,黎檀实. 脓毒血症中固有免疫细胞的调节机制研究进展[J]. *中国免疫学杂志*, 2016, 32(4): 576-583.
- [11] Pierrakos C, Velissaris D, Bisdorff M, et al. Biomarkers of sepsis: time for a reappraisal [J]. *Crit Care*, 2020, 24(1): 287.
- [12] Reyes M, Filbin MR, Bhattacharyya RP, et al. An immune-cell signature of bacterial sepsis [J]. *Nat Med*, 2020, 26(3): 333-340.
- [13] Yao RQ, Li ZX, Wang LX, et al. Single-cell transcriptome profiling of the immune space - time landscape reveals dendritic cell regulatory program in polymicrobial sepsis [J]. *Theranostics*, 2022, 12(10): 4606-4628.
- [14] Lei Y, Tang R, Xu J, et al. Applications of single - cell sequencing in cancer research: progress and perspectives [J]. *J Hematol Oncol*, 2021, 14(1): 91.
- [15] Grün D, Van Oudenaarden A. Design and Analysis of Single-Cell Sequencing Experiments [J]. *Cell*, 2015, 163(4): 799-810.
- [16] Kang I, Bucala R. The immunobiology of MIF: function, genetics and prospects for precision medicine [J]. *Nat Rev Rheumatol*, 2019, 15(7): 427-437.
- [17] Sumaiya K, Langford D, Natarajaseenivasan K, et al. Macrophage migration inhibitory factor (MIF) : A multifaceted cytokine regulated by genetic and physiological strategies [J]. *Pharmacol Ther*, 2022, 233: 108024.
- [18] Sun Y, Wang Y, Li JH, et al. Macrophage migration inhibitory factor counter-regulates dexamethasone-induced annexin 1 expression and influences the release of eicosanoids in murine macrophages [J]. *Immunology*, 2013, 140(2): 250-258.
- [19] Tilstam PV, Schulte W, Holowka T, et al. MIF but not MIF-2 recruits inflammatory macrophages in an experimental polymicrobial sepsis model [J]. *J Clin Invest*, 2021, 131(23): e127171.
- [20] Grieb G, Kim BS, Simons D, et al. MIF and CD74 - suitability as clinical biomarkers [J]. *Mini Rev Med Chem*, 2014, 14(14): 1125-1131.
- [21] Bloom J, Sun S, Al - Abed Y. MIF, a controversial cytokine: a review of structural features, challenges, and opportunities for drug development [J]. *Expert Opin Ther Targets*, 2016, 20(12): 1463-1475.
- [22] Fox RJ, Coffey CS, Conwit R, et al. Phase 2 Trial of Ibudilast in Progressive Multiple Sclerosis [J]. *N Engl J Med*, 2018, 379(9): 846-855.

(收稿日期:2023-11-02)

(本文编辑:欧阳菁)

· 读者 · 作者 · 编者 ·

医学论文中阿拉伯数字的使用规则

本刊关于阿拉伯数字的使用规则如下:(1)凡是可以使用阿拉伯数字而且很得体的地方均应使用阿拉伯数字。(2)公历世纪、年代、年、月、日和时刻必须使用阿拉伯数字。(3)日期可采用全数字式写法,例如:1999-02-01或1999 02 01或19990201,本刊采用如下格式:1999年2月1日,年份用4位数表示,不能简写,如1999年不能写成99年。(4)日的时间表示按以下写法,如下午3时20分50秒,写成15:20:50。(5)计量和计数单位的数字一律使用阿拉伯数字。(6)引文标注中的版次、卷、期、页码等使用阿拉伯数字。(7)小数点前或后超过4位数(含4位),应从小数点起向左或向右每3位空半个阿拉伯数字的空隙(1/4汉字),不用千分撇分节法,年份、部队代号、仪器型号、标准号等非计量数字不分节。