

急性心肌梗死与中性粒细胞相关潜在生物标志物的机器学习分析

杨定燕¹, 余振球¹, 杨中愉²<https://doi.org/10.12307/2025.523>

投稿日期: 2024-04-17

采用日期: 2024-06-01

修回日期: 2024-07-13

在线日期: 2024-08-15

中图分类号:

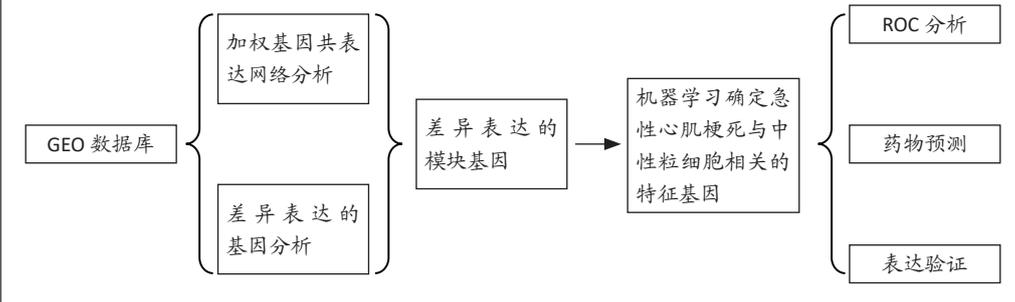
R452; R363; R542.22

文章编号:

2095-4344(2025)36-07909-12

文献标识码: A

文章快速阅读: 急性心肌梗死与中性粒细胞相关生物标志物的确定



文题释义:

急性心肌梗死: 是冠状动脉(指供应心脏自身血液的动脉, 分为左、右冠状动脉, 行走于心脏表面)完全闭塞, 从而导致心肌细胞急性/持续性缺血和缺氧所致的心肌坏死。

血清性生物标志物: 是指能被客观测量和评价, 反映疾病生理或病理过程, 以及对暴露或治疗干预措施产生生物学效应的指标。

摘要

背景: 精确的早期诊断和及时的再灌注治疗是挽救急性心肌梗死患者的生命并改善预后的重要前提条件。因此, 寻找能够早期诊断急性心肌梗死的理想生物标志物尤为重要。

目的: 拟通过生物信息学和机器学习分析急性心肌梗死与中性粒细胞相关的关键基因, 以探寻新的生物标志物。

方法: 基于GEO数据库和limma包鉴定急性心肌梗死的差异表达基因。使用反卷积算法探究免疫细胞浸润情况, 然后结合加权基因共表达网络分析(Weighted gene co-expression network analysis, WGCNA)、蛋白互作网络和机器学习筛选急性心肌梗死与中性粒细胞相关的特征基因, 并进行功能富集分析。用ROC曲线评估特征基因对急性心肌梗死的诊断价值。通过STITCH和Herb数据库筛查生物标志物的靶向药物。最后将在2023年3-6月于贵州医科大学附属医院内科首次诊断为急性心肌梗死的住院患者作为实验组, 同期心电图无缺血性改变、冠状动脉造影无狭窄的住院患者作为对照组, 收集两组患者外周血, 通过RT-qPCR验证基因在人外周血样本中的相对表达量。

结果与结论: ①共获得差异表达基因2 349个, 免疫浸润分析发现B cells memory, NK cells resting和Neutrophils等免疫细胞评分在疾病和正常组之间存在差异; ②使用WGCNA发现ME green和ME turquoise这两个基因模块与中性粒细胞与急性心肌梗死表现出最高的相关性; ③与差异表达基因相交后获得24个差异模块基因, 功能富集分析发现其与先天免疫反应、细菌的防御反应等多种过程相关; KEGG结果显示其主要与肿瘤坏死因子信号通路有关。④机器学习算法挖掘到的基因取交集后得到的特征基因为S100A12, PTCH1和LOC400499, 在GSE48060和GSE66360数据集中的ROC曲线下面积均大于0.7, 将其视为潜在的生物标志物。⑤基于STITCH和Herb数据库发现S100A12有11种靶向药物, PTCH1共发现6种靶向药物。⑥RT-qPCR结果显示, 与对照组相比, 急性心肌梗死患者中S100A12, PTCH1和LOC400499表达具有显著差异性($P < 0.05$)。⑦S100A12, PTCH1和LOC400499可能是急性心肌梗死潜在的诊断生物标志物, 但是其与急性心肌梗死相关的特异性尚需进一步研究, 其中S100A12可能是调控急性心肌梗死的潜在靶点。

关键词: 急性心肌梗死; 中性粒细胞; 生物标志物; 差异分析; WGCNA; 机器学习算法; 免疫浸润分析; S100A12; PTCH1; LOC400499

Machine learning-based analysis of neutrophil-associated potential biomarkers for acute myocardial infarction

Yang Dingyan¹, Yu Zhenqiu¹, Yang Zhongyu²¹Department of Hypertension, Affiliated Hospital of Guizhou Medical University, Guiyang 550004, Guizhou Province, China; ²North Sichuan Medical College, Nanchong 637100, Sichuan Province, China

Yang Dingyan, Master candidate, Department of Hypertension, Affiliated Hospital of Guizhou Medical University, Guiyang 550004, Guizhou Province, China

Corresponding author: Yu Zhenqiu, Chief physician, Department of Hypertension, Affiliated Hospital of Guizhou Medical University, Guiyang 550004, Guizhou Province, China

Abstract

BACKGROUND: Accurate early diagnosis and timely reperfusion therapy are important prerequisites for saving the lives and improving the prognosis of patients with acute myocardial infarction. Therefore, it is important to find ideal biomarkers for early diagnosis of acute myocardial infarction.

¹贵州医科大学附属医院高血压科, 贵州省贵阳市 550004; ²川北医学院, 四川省南充市 637100

第一作者: 杨定燕, 女, 1997年生, 贵州省人, 苗族, 贵州医科大学在读硕士, 主要从事心肌梗死相关研究。

通讯作者: 余振球, 主任医师, 贵州医科大学附属医院高血压科, 贵州省贵阳市 550004

<https://orcid.org/0009-0009-0095-9354> (杨定燕)

基金资助: 贵州省卫生健康委科学技术基金项目(gzkwkj2023-100), 项目负责人: 余振球

引用本文: 杨定燕, 余振球, 杨中愉. 急性心肌梗死与中性粒细胞相关潜在生物标志物的机器学习分析[J]. 中国组织工程研究, 2025, 29(36):7909-7920.



OBJECTIVE: To analyze key genes associated with neutrophils by acute myocardial infarction through bioinformatics and machine learning to explore new biomarkers.

METHODS: Differentially expressed genes were identified based on the Gene Expression Omnibus (GEO) database and Limma R package. Deconvolution algorithm was used to explore the immune cells infiltration level. Then, acute myocardial infarction and neutrophils-related biomarkers were screened by weighted gene co-expression network analysis (WGCNA), protein-protein interaction (PPI) networks, machine learning, and functional enrichment analysis. Receiver operating characteristic curve analysis was conducted to assess the diagnostic efficacy of biomarkers for acute myocardial infarction. Targeted drugs for biomarkers were screened through the STITCH and Herb database. Finally, the hospitalized patients who were first diagnosed with acute myocardial infarction in the Department of Cardiology of Affiliated Hospital of Guizhou Medical University from March to June 2023 were used as the experimental group, and the hospitalized patients who had no ischemic changes on electrocardiograms and no stenosis on coronary angiograms during the same period were used as the control group. Peripheral blood of the patients in the two groups was collected. The relative expressions of the genes were verified in the human peripheral blood samples by RT-qPCR.

RESULTS AND CONCLUSION: (1) A total of 2 349 differentially expressed genes were obtained, and immune infiltration analysis revealed differences in immune cell scores such as B cells memory, NK cells resting, and Neutrophils between the disease and normal groups. (2) Using WGCNA, two gene modules, ME green and ME turquoise, were found to exhibit the highest correlation with neutrophil fine with acute myocardial infarction. (3) Twenty-four differential module genes were obtained after intersecting with differentially expressed genes. Functional enrichment analysis revealed that they were associated with a variety of processes such as innate immune response and defense response to bacteria. KEGG results showed that they were mainly associated with the tumor necrosis factor signaling pathway. (4) The genes mined by the machine learning algorithm took the intersection to obtain three genes, namely, S100A12, PTCH1, and LOC400499, all of which were greater than 0.7 by the area under the receiver operating characteristic curve in both the GSE48060 and GSE66360 datasets. They were considered as potential biomarkers. (5) Based on the STITCH and Herb databases, 11 target drugs were found for S100A12 and a total of 6 target drugs were found for PTCH1. (6) RT-qPCR results showed that S100A12, PTCH1, and LOC400499 were significantly differentially expressed in acute myocardial infarction patients compared with controls ($P < 0.05$). (7) S100A12, PTCH1, and LOC400499 may be potential diagnostic biomarkers for acute myocardial infarction, but their specificity in relation to acute myocardial infarction needs to be further investigated, in which S100A12 may be a potential target for regulating acute myocardial infarction.

Key words: acute myocardial infarction; neutrophil; biomarker; differential analysis; WGCNA; machine learning algorithms; immune infiltration analysis; S100A12; PTCH1; LOC400499

Funding: Science and Technology Fund Project of Guizhou Provincial Health Commission, No. gzwkjz2023-100 (to YZQ)

How to cite this article: YANG DY, YU ZQ, YANG ZY. Machine learning-based analysis of neutrophil-associated potential biomarkers for acute myocardial infarction. *Zhongguo Zuzhi Gongcheng Yanjiu*. 2025;29(36):7909-7920.

0 引言 Introduction

急性心肌梗死是一个重大的公共卫生问题,通常由冠状动脉部分或完全闭塞导致心肌供血急剧减少或中断,使相应部位的心肌细胞持久的缺血缺氧而出现严重坏死,在全球有很高的致死率和致残率,危及大多数人的健康^[1-2]。精准的早期诊断和及时的心肌再灌注治疗能挽救濒临坏死的心肌细胞或限制梗死面积扩大,减轻梗死后的心肌重塑,提高急性心肌梗死患者的生存率,改善预后。目前,临床上常用肌钙蛋白、肌酸激酶同工酶作为急性心肌梗死患者早期诊断的主要血清生物标志物。急性心肌梗死后肌钙蛋白和肌酸激酶同工酶由坏死的心肌细胞释放到外周血中,但这个需要一定的时间,一般为三四个小时,这限制了更早期对急性心肌梗死的识别,特别是对于早期症状和心电图不典型的老齡或糖尿病等患者^[3-5]。此外,肌钙蛋白和肌酸激酶同工酶仅反映心肌受损的程度,不能很好地反映急性心肌梗死的发病机制和防治。

急性心肌梗死后,中性粒细胞作为最早被动员到梗死区域的白细胞,参与心肌修复和重塑的整个过程^[6]。早期中性粒细胞通过分泌基质金属蛋白酶、髓过氧化物酶、活性氧等炎症

递质,清除坏死的心肌细胞和基质碎片^[7-8]。随后,中性粒细胞诱导巨噬细胞和自身极化,激活成纤维细胞,抑制促炎因子产生,促进白细胞介素 10 和转化生长因子 β 等的产生来激活抗炎反应促进瘢痕的生成和血运重建,从而促进梗死区域的愈合修复^[9-11]。所以,早期中性粒细胞的适度活化是过渡到修复和重塑的必要条件,然而中性粒细胞的过度活化或消退延迟会导致心肌的不良重塑和功能障碍,最终导致心力衰竭,甚至死亡^[12-14]。按机制来说,中性粒细胞及其炎症递质的调控可能是急性心肌梗死治疗的理想方法。研究发现,炎症期 Dusp6 缺乏能抑制中性粒细胞毒性作用从而改善心脏修复,并且这种抑制作用对其他免疫细胞的影响很小^[15]。但是,除了白细胞介素 1 β 的靶向药物 Canakinuma 可以降低急性心肌梗死后患者主要心血管事件的发生率外,仍然很难将免疫调节和抗炎治疗策略转化为临床实践^[16]。因此,寻找新的急性心肌梗死与中性粒细胞相关的生物标志物可能不仅有助于急性心肌梗死的更早期识别,还可能有助于急性心肌梗死的防治。

微阵列分析能在短时间内发现影响疾病的基因,并且作为早期诊断的生物标志物^[17]。然而,微阵列分析在

重现性和灵敏度方面存在局限性。机器学习具有强大的计算能力,能够分析由数百到数千个样本和复杂的生物数据集,并从这些数据中提取有价值的生物学见解信息^[18]。因此,机器学习为解释不同类型的组学数据提供了创新方法,从而识别了新的生物标记物,有助于精确的疾病预测、患者分层和开发新的治疗方法。

1 资料和方法 Data and methods

1.1 设计 基因信息学试验,机器学习分析结合人外周血样本验证,独立样本 t 检验。

1.2 时间及地点 2023 年 6 月至 2024 年 3 月在贵州医科大学附属医院临床研究中心实验室完成。

1.3 资料

1.3.1 芯片数据来源和处理 从 GEO 数据库 (<https://www.ncbi.nlm.nih.gov>) 中下载急性心肌梗死的 RNA-seq 数据。选取 GSE48060 和 GSE66360 数据集。其中 GSE48060 数据集包含 31 例入院 48 h 内急性心肌梗死患者的外周血样本和 21 例对照样本,主要用于差异分析和模型的构建; GSE66360 数据集包含 49 例急性心肌梗死患者外周血样本和 50 例对照样本,用于差异基因筛选和模型的验证。

1.3.2 研究对象 选取 2023 年 3-6 月

在贵州医科大学附属医院心内科住院治疗的患者。将5例明确诊断为新发急性心肌梗死患者作为急性心肌梗死组；选取5例同期住院治疗的心电图无缺血性改变、冠状动脉造影无狭窄的患者为对照组；患者年龄均为30-60岁。急性心肌梗死的诊断符合第四版急性心肌梗死通用定义^[4]，即除需要心肌钙蛋白升高超过参考值上限第99百分位数外，同时还需至少一项心肌缺血证据：①心肌缺血的症状；②新发缺血性心电图改变；③新出现的病理性Q波；④影像学提示与缺血一致的新出现存活心肌的缺失或阶段性室壁运动异常；⑤冠状动脉造影或尸检证实冠状动脉血栓。研究方案于2023-05-16经贵州医科大学伦理审查批准(批准号：2022伦审第400号)，遵循《赫尔辛基宣言》原则。所有参与者均充分了解研究目的、研究过程和潜在风险，均已签署知情同意书。

1.4 方法

1.4.1 筛选急性心肌梗死与中性粒细胞相关的关键基因 原始数据下载后，使用R语言(https://www.r-project.org/)进行归一化处理。随后利用R语言limma包鉴定GSE48060数据集和GSE66360数据集中急性心肌梗死组和对照组中的差异表达基因，差异筛选条件为P-value < 0.5和|log₂FC| > 0.5。差异表达基因的火山图由R语言ggplot2包进行可视化处理。

免疫细胞浸润分析：利用反卷积算法计算GSE48060数据集中急性心肌梗死组和对照组的每个样本中22种免疫细胞比例，使用Wilcoxon检验评估每种免疫细胞的丰度的差异。

加权基因共表达网络分析(Weighted gene co-expression network analysis, WGCNA)分析鉴定与中性粒细胞相关的模块基因：为了得到急性心肌梗死与中性粒细胞最相关的目标基因，文章对GSE48060数据集进行WGCNA分析。首先使用R语言WGCNA包的goodSamplesGenes函数检验样本的基因是否需要过滤，如果没有缺失值则返回TURE。利用pickSoftThreshold功能进行软阈值的筛选，筛选依据无尺度网络重要参数，设置软阈值β(即power值)需满足

SFT.R.sq(即R²)大于0.85。构建基因层次聚类树，使用动态混合剪切算法对基因进行分配，得到模块基因，随后将相似性大的模块进行合并。模块识别后，并分析模块与急性心肌梗死和中性粒细胞之间的相关性，找到与之最相关的模块进行进一步分析。

差异表达筛选及功能富集分析：将GSE48060数据集和GSE66360数据集的差异表达基因与模块基因进行交集，交集取得的差异模块基因用作后续分析。随后使用R语言clusterProfiler包对差异模块基因进行KEGG和GO的富集分析，以确定差异模块基因的生物学功能和信号转导通路。

关键基因蛋白相互作用网络的构建：利用STRING数据库(https://cn.string-db.org/)对交集的差异模块基因进行蛋白互作网络(protein protein interaction, PPI)的构建，使用Cytoscape(https://cytoscape.org/)进行PPI网络的可视化和分析。

机器算法筛选候选生物标志物：采用随机森林模型，支持向量机模型和广义线性模型分别在GSE48060数据集中构建诊断模型，使用R语言caret包中的rfeControl函数筛选出特征基因，随后对3种方法筛选的特征基因取交集，得到3种算法中同时存在的基因作为候选生物标志物。

ROC曲线验证候选生物标志物的诊断能力：分别在GSE48060和GSE66360数据集中对候选生物标志物进行验证，确定生物标志物，并使用R语言pROC包绘制ROC曲线，曲线下的面积代表AUC值，一般AUC值越大，说明预测的较为准确。

单基因富集分析：以生物标志物为目的基因，分别计算所有基因和目的基因表达量的相关系数作为排序标准，以此来进行单基因富集分析。随后使用R语言clusterProfiler包进行富

集分析，筛选条件为P-adjust < 0.05。

药物预测：将确定的生物标志物输入STITCH数据库(https://ngdc.cnbc.ac.cn/databasecommons/database/id/208)和herb数据库(https://ngdc.cnbc.ac.cn/databasecommons/database/id/7435)获得基因对应的药物靶点信息，使用Cytoscape将其进行可视化展示。

1.4.2 样本收集和RT-qPCR验证结果 使用抗凝管采集患者外周血10 mL，2 000 r/min离心10 min，分离血浆，将其编号后保存于-80 °C冰箱。使用RNA分离试剂盒从外周血中提取总RNA(EK-5301，广州亿涛生物科技有限公司)。使用SweScript RT II First Strand cDNA Synthesis Kit试剂盒(G332，武汉塞维尔科技有限公司)对cDNA进行反转录。使用2xRealStar Fast SYBR qPCR Mix预混试剂盒(A301，北京康润诚业生物科技有限公司)配置反应体系，反应条件：预变性95 °C，2 min进行1个循环；95 °C变性15 s；60 °C退火/延伸30 s，进行40个循环。通过CPX 96Touch PCR仪(CFX96 Touch，美国Bio-Rad公司)测定mRNA表达量。使用β-actin为内参基因进行校正。按照2^{-ΔΔCt}计算基因相对表达水平，其中，ΔCt=(目的基因Ct-内参Ct)；ΔΔCt=(待测样品中目的基因ΔCt-参照样品中目的基因ΔCt)^[19]。所有引物由广州艾基生物科技有限公司合成，相关基因的引物序列见表1。

1.5 主要观察指标 急性心肌梗死与中性粒细胞相关的生物标志物S100A12, PTCH1和LOC400499在外周血中的相对表达。

1.6 统计学分析 使用Graphpad Prism 9.5统计学软件(GraphPad Software公司，美国)对RT-qPCR得到的相对表达量进行统计分析。计量资料符合正

表1 | 内参基因和生物标志物的引物序列

基因名称	引物序列(5'-3')	片段长度(bp)
S100A12	正向: CAC ATT CCT GTG CAT TGA GG	122
	反向: GGT GTC AAA ATG CCC CTT C	
PTCH1	正向: ACA AAC TCC TGG TGC AAA CC	207
	反向: CTT TGT CGT GGA CCC ATT C	
LOC400499	正向: CTG CTT CCG GGT GGT AGA C	291
	反向: TGA GTT TCC CGG CAT TCC AT	
β-actin	正向: CTC GCT TCG GCA GCA CA	96
	反向: AAC GCT TCA CGA ATT TGC GT	

态分布，以 $\bar{x} \pm s$ 表示。采用独立样本 t 检验比较两组间的差异。 $P < 0.05$ 为差异有显著性意义。

2 结果 Results

2.1 筛选差异表达的基因 在 GSE48060 数据集中得到疾病与正常相比显著差异表达基因共有 86 个，其中上调基因有 31 个，下调基因有 55 个，见图 1。在 GSE66360 数据集中得到疾病与正常相比显著差异表达基因共有 2 349 个，其中上调基因有 877 个，下调基因有 1 472 个，见图 2。

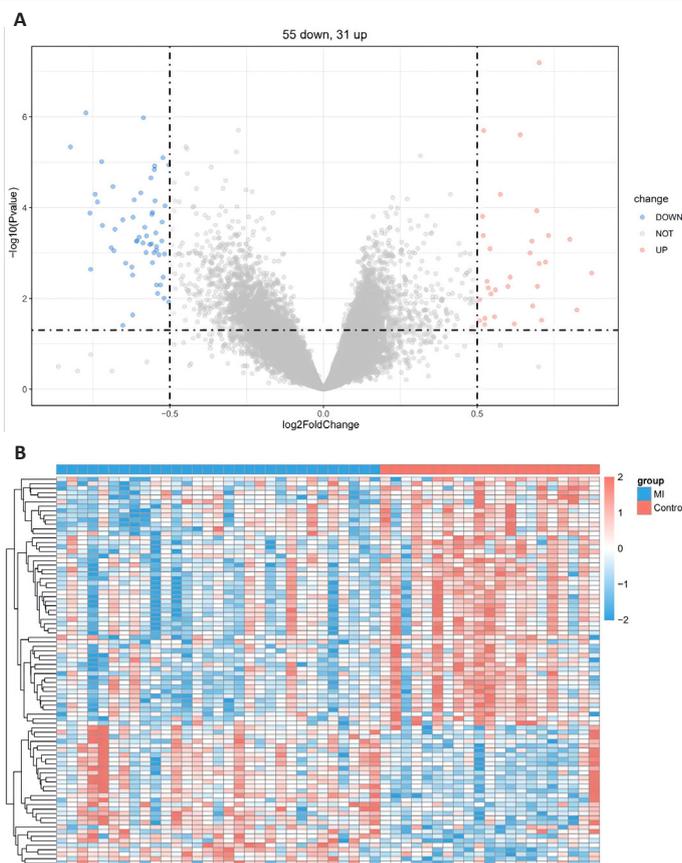
2.2 免疫浸润分析结果 图 3A 显示了每个样本中 20 种免疫细胞的丰度，M1 型巨噬细胞和活化的肥大细胞丰度为 0，故剔除。小提琴图表明记忆 B 细胞、静息的自然杀伤细胞和中性粒细胞的表达在急性心肌梗死组和对照组中存在显著差异，与对照组相比，中性粒细胞百分率显著升高，见图 3B。

2.3 WGCNA 分析结果 文章对 GSE48060 的样本进行聚类，在去除异常值后，绘制了一个样本聚类树，见图 4。文章将软阈值设为 6 ($R^2=0.85$) 用以构建无标度网络，见图 5。聚集了 11 个模块，见图 6，模块过多，将相似性大的模块合并后共有 7 个模块，见图 7。计算并绘制每个模块与急性心肌梗死和中性粒细胞之间的相关性，见图 8。结果表明，ME green 和 ME turquoise 与急性心肌梗死和中性粒细胞相关性最高 ($P < 0.05$)，被认为是关键模块进行后续分析。

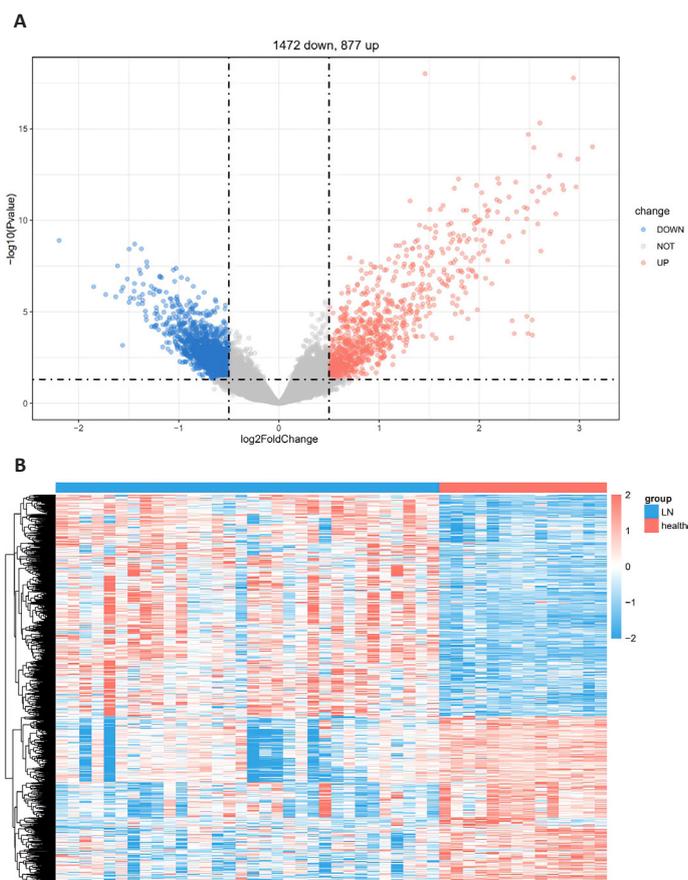
2.4 差异模块基因筛选及功能富集分析结果

2.4.1 差异模块基因筛选结果 使用 Ven 图展示 GSE48060 数据集、GSE66360 数据集差异基因与模块基因的交集，结果显示上调组存在 20 个交集基因，下调组存在 4 个交集基因，见图 9，将这 24 个基因作为后续分析的潜在基因。

2.4.2 模块差异基因功能富集分析结果 分别对 24 个差异模块基因进行基于 KEGG 和 GO 的富集分析，文章以 $P.adjust < 0.05$ 的条件对富集到的通路进行筛选，最终结果显示 24 个 mRNA 共富集到 355 个 GO 条目，如图 10 展示，Top 前 30 条的条目，包括生物学过程、细胞组分和分子功能。

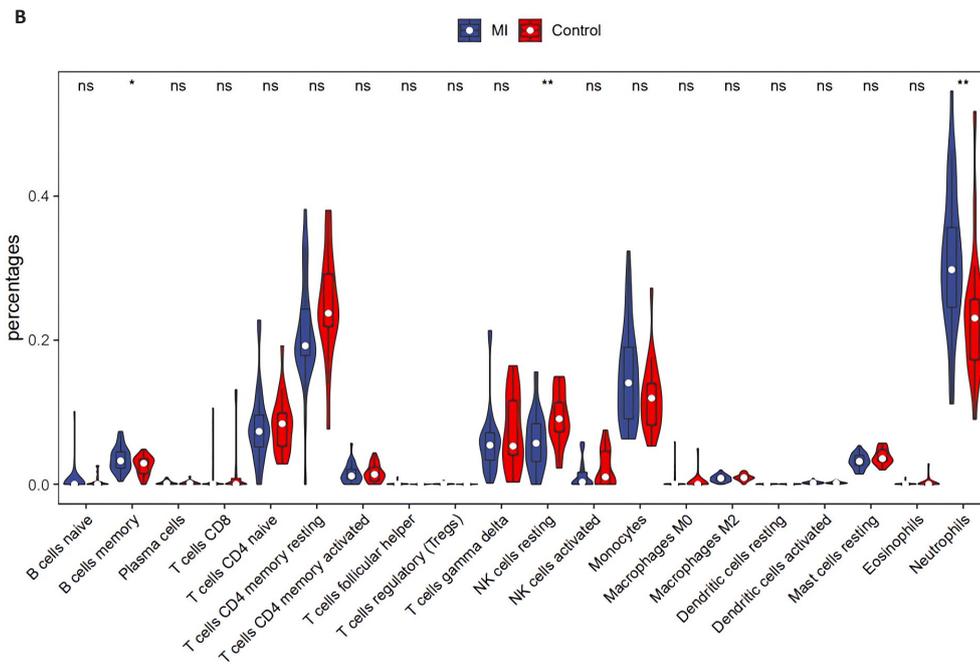
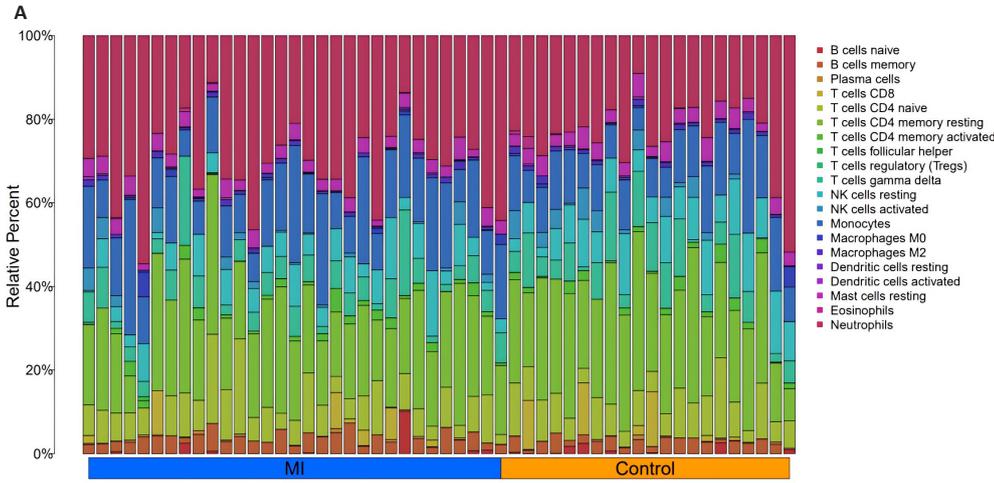


图注：图 A 为 GSE48060 数据集中差异表达基因的火山图，横坐标表示差异倍数，纵坐标表示差异的显著性；红色的点表示上调基因，蓝色的点表示下调基因；B 为急性心肌梗死和健康样本中 86 个差异表达基因的热图，每行表示一个基因在样品中的表达量，每列表示一个样品中基因的表达量。每个小方格表示每个样本的不同基因标准化之后的表达量，其颜色表示含量多少，表达量越高颜色越红，越少越蓝。



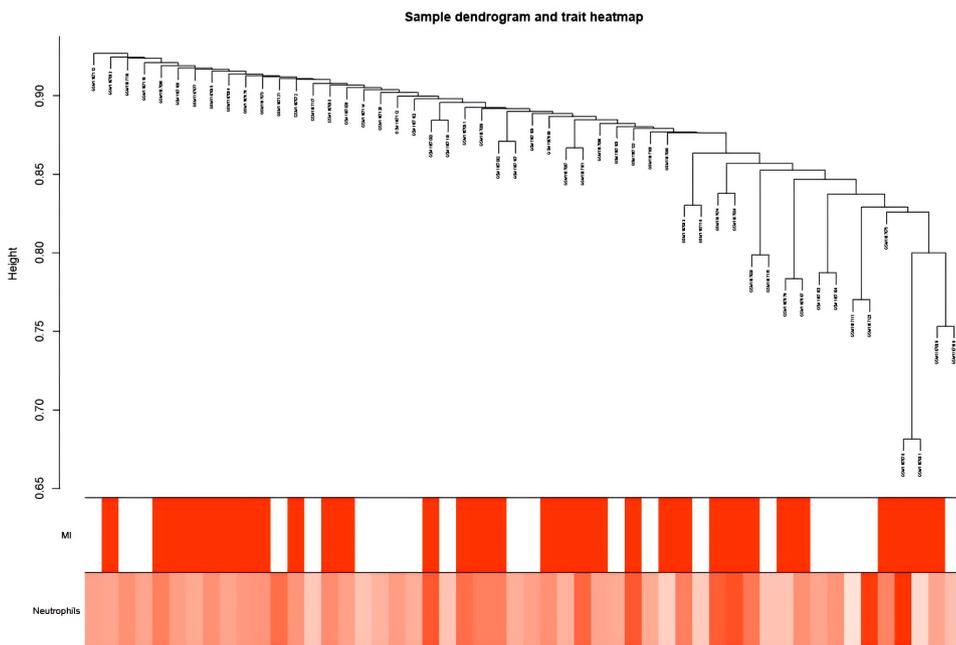
图注：图 A 为 GSE66030 数据集中差异表达基因的火山图，横坐标表示差异倍数，纵坐标表示差异的显著性；红色的点表示上调基因，蓝色的点表示下调基因；B 为急性心肌梗死和健康样本中 2 349 个差异表达基因的热图，每行表示一个基因在样品中的表达量，每列表示一个样品中基因的表达量。每个小方格表示每个样本的不同基因标准化之后的表达量，其颜色表示含量多少，表达量越高颜色越红，越少越蓝。

图 2 | 急性心肌梗死和健康样本中差异表达基因 (GSE66030 数据集)



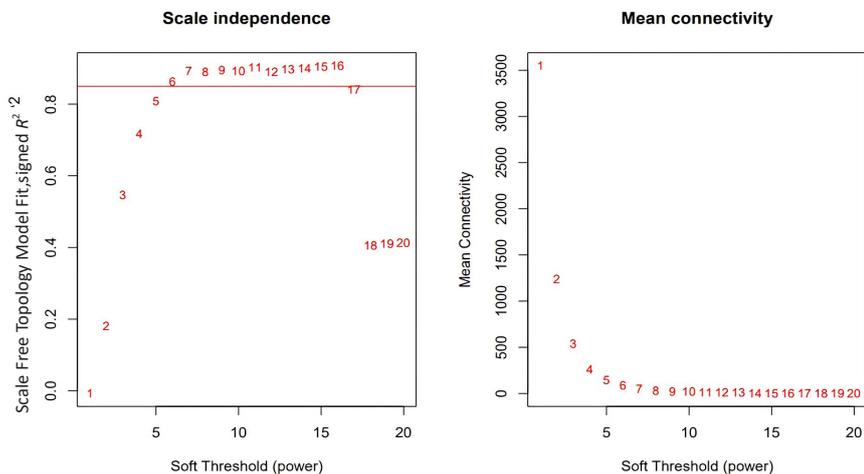
图注：图 A 为反卷积算法得到的免疫细胞柱状堆叠图，每一列表示每一例患者中不同免疫细胞评分高低情况，不同颜色表示不同类型的免疫细胞；图 B 为免疫细胞在正常疾病组中的免疫评分小提琴图展示，* $P < 0.05$ ，** $P < 0.01$ ，^{ns} $P > 0.05$ 。MI 为急性心肌梗死组，control 为对照组。B cell naive 为 B 幼稚细胞，B cell memory 为 B 记忆细胞，Plasma cells 为浆细胞，T cells CD8 为 CD8 T 细胞，T cells CD4 naive 为 CD4 T 幼稚细胞，T cells CD4 memory resting 为静息的 CD4 T 记忆细胞；T cells follicular helper 为 T 细胞滤泡辅助细胞；T cells regulatory (Tregs) 为调节性 T 细胞；T cells gamma delta 为 $\gamma\delta$ T 细胞；NK cells resting 为静息的 NK 细胞；NK cells activated 为活化的 NK 细胞；Monocytes 为单核细胞；Macrophages M0 为 M0 型巨噬细胞；Macrophages M2 为 M2 型巨噬细胞；Dendritic cells resting 为静息的树突状细胞；Dendritic cells activated 为活化的树突状细胞；Mast cells resting 为静息的肥大细胞；Eosinophils 为嗜酸性粒细胞；Neutrophils 为中性粒细胞。

图 3 | 免疫浸润分析结果



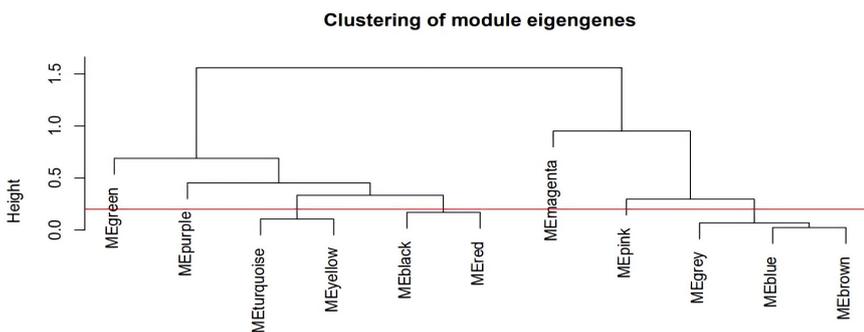
图注：此图分为两部分，上半部分表示样本聚类树形图，纵坐标为高度，高度越高基因表达越不相似；下半部分为表型性状热图，颜色越深表示表型与样本的基因表达量关系越密切。Sample dendrogram and trait heatmap 为样本树状图和表型性状热图；MI 为心肌梗死；Neutrophils 为中性粒细胞。

图 4 | 样本和形状树形图



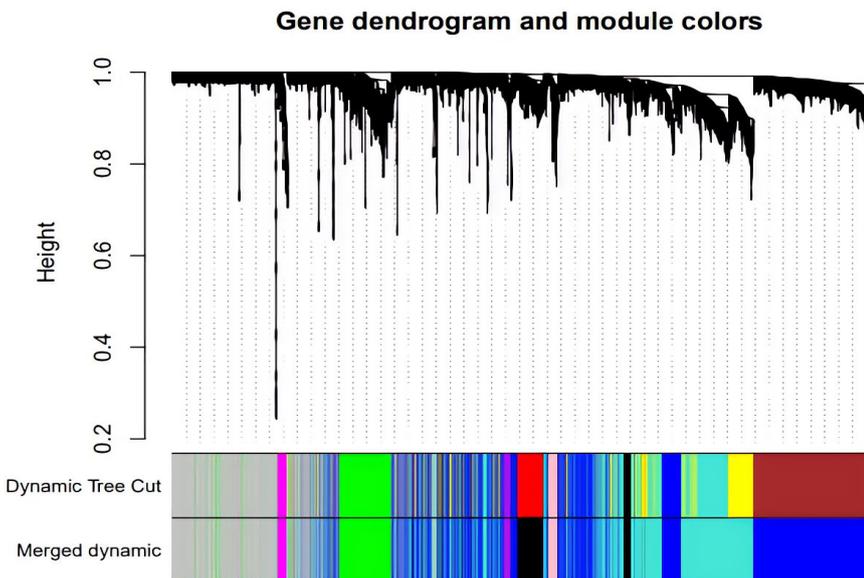
图注：两张图横轴均代表权重参数 power 值，左图纵轴表示 signed R^2 ，相关系数的平方越高，说明该网络越逼近无尺度分布；右图纵轴表示对应的基因模块中所有基因邻接函数的均值。

图 5 | 无尺度软阈值的筛选结果



图注：纵坐标为高度，高度越高，数字越大，基因表达越不相似，树状图不同的分支代表不同的模块基因。Clustering of module eigengenes 为模块基因聚类；Height 为高度；MEgreen 为绿色模块基因；MEpurple 为紫色模块基因；MEturquoise 为蓝绿色模块基因；MEyellow 为黄色模块基因；MEblack 为黑色模块基因；MERed 为红色模块基因；MEagenta 为品红色模块基因；MEpink 为粉色模块基因；MEgrey 为灰色模块基因；MEblue 为蓝色模块基因；MEbrown 为棕色模块基因。

图 6 | 聚类模块树形图



图注：图的上半部分是一个层次聚类图，每个叶子结点都是一条竖线，代表一个基因；纵坐标为高度，高度越高基因表达越不相似；下半部分为基因模块，不同的颜色代表不同的模块，灰色的模块为未归类于任何共表达模块的基因组成。Height 为高度；Dynamic Tree Cut 为动态树剪切；Merged dynamic 为动态合并相似模块后的颜色；Gene dendrogram and module color 为基因树状图和模块颜色。

图 7 | 模块识别与合并分析结果

生物学过程主要富集于炎症反应、对细菌的防御反应、中性粒细胞介导的免疫反应等方面；细胞组分主要富集于内吞囊泡膜、三级颗粒及特定颗粒流明等方面；分子功能富集于丝氨酸水解酶活性、丝氨酸型肽酶活性、RAGE 受体结合等方面。KEGG 富集结果显示富集到 7 条 KEGG 通路，如图 11 所示，肿瘤坏死因子信号通路、神经活性配体 - 受体相互作用及雌激素信号通路显著富集。

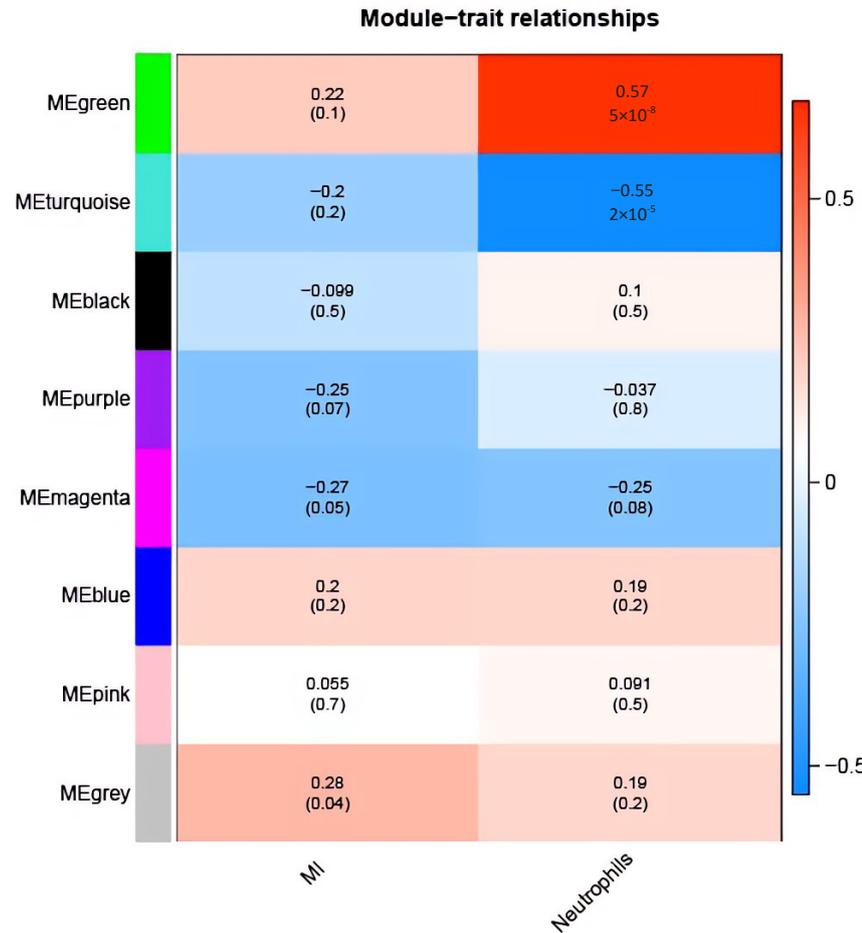
2.4.3 模块差异基因蛋白相互作用网络的构建 为了解这 24 个基因的关联，文章构建了一个 PPI 网络。该网络由包含 21 个节点，76 条边组成，见图 12。其中 S100A12 与 16 种蛋白具有相互作用，见图 13，通过 Cytohubba 筛选 24 个差异模块基因中的核心基因，获得了 MMP9，S100A12，VNN1 和 AQP9 为核心基因。

2.5 机器学习构建诊断模型

2.5.1 机器学习 如图 14A-C 所示，使用支持向量模型算法得到 14 个特征基因，分别为 S100A12，PTCH1，DSC2，ADM，MGAM，LOC400499，HP，ANXA3，MMP9，SOCS3，SLPI，MCEMP1，SULT1B1，TDRD9，使用广义线性模型算法得到 8 个特征基因，分别为 PTCH1，MGAM，SULT1B1，LOC400499，S100A12，GZMA，HP 和 SOCS3，使用随机森林模型算法得到 16 个特征基因，分别为 LOC400499，ASGR2，PTCH1，PTGDR，ANXA3，SULT1B1，SOCS3，ARG1，GZMA，SLPI，FCGR1B，MCEMP1，DSC2，S100A12，B3GNT5 和 A2M_AS1。

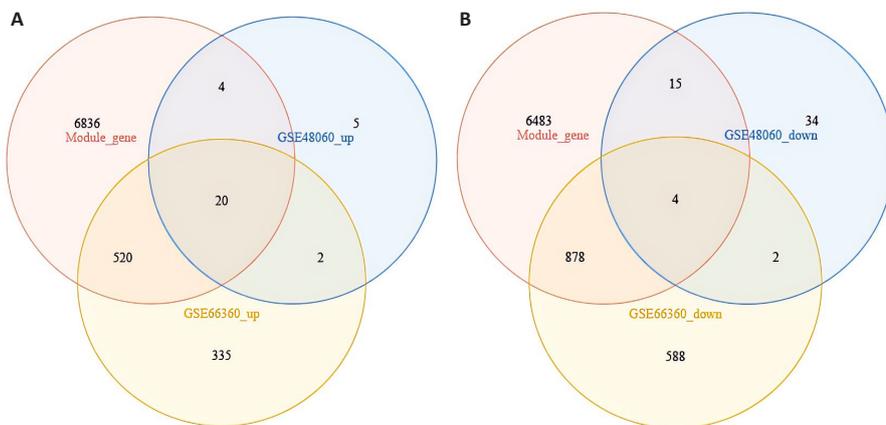
2.5.2 特征基因取交集 将支持向量模型算法得到 14 个特征基因，广义线性模型算法得到 8 个特征基因，随机森林模型算法得到 16 个特征基因进行取交集后，得到的 4 个基因在 3 种机器学习算法中同时存在的基因分别为 S100A12，PTCH1，LOC400499 和 SULT1B1，见图 14D，将其定为候选的生物标志物。

2.5.3 ROC 曲线验证关键核心基因的诊断能力 在 GES48060 数据集中分析急性心肌梗死和对照样本之间与中性粒细胞相关的候选标志物的表达，并构建 ROC 曲线。图 15A 中所有基



图注：纵坐标为不同模块，横坐标为表型性状，颜色条带表示相关性范围；每一个方块表示模块基因和性状的相关性，红色表示正相关，蓝色表示负相关，颜色越深相关性越强。方块内的数字表示相关性和显著性。MI为心肌梗死；Neutrophils为中性粒细胞；Module-trait relationships为模块与表型特征关联图谱；MEgreen为绿色模块基因；MEturquoise为蓝绿色模块基因；MEblack为黑色模块基因；MEpurple为紫色模块基因；MEagenta为品红色模块基因；MEblue为蓝色模块基因；MEpink为粉色模块基因；MEgrey为灰色模块基因。

图8 | 模块基因和临床性状相关性分析结果图



图注：图A为差异表达的上调基因的交集图；B为差异表达的下调基因的交集图。Module-gene为模块基因，up上调基因，down为下调基因。模块基因以粉色表示，GSE48060数据集中的差异表达基因以蓝色表示，GSE66360数据集中的差异表达基因以橘色表示，三者交集部分基因代表在差异表达的模块基因、GSE48060和GSE66360数据集中均存在。

图9 | 模块基因和差异表达基因的交集图

因的AUC值都大于0.78，4个基因构建模型的AUC值为0.92，说明这4个基因对于区分疾病和对照样本具有一定的准确性。

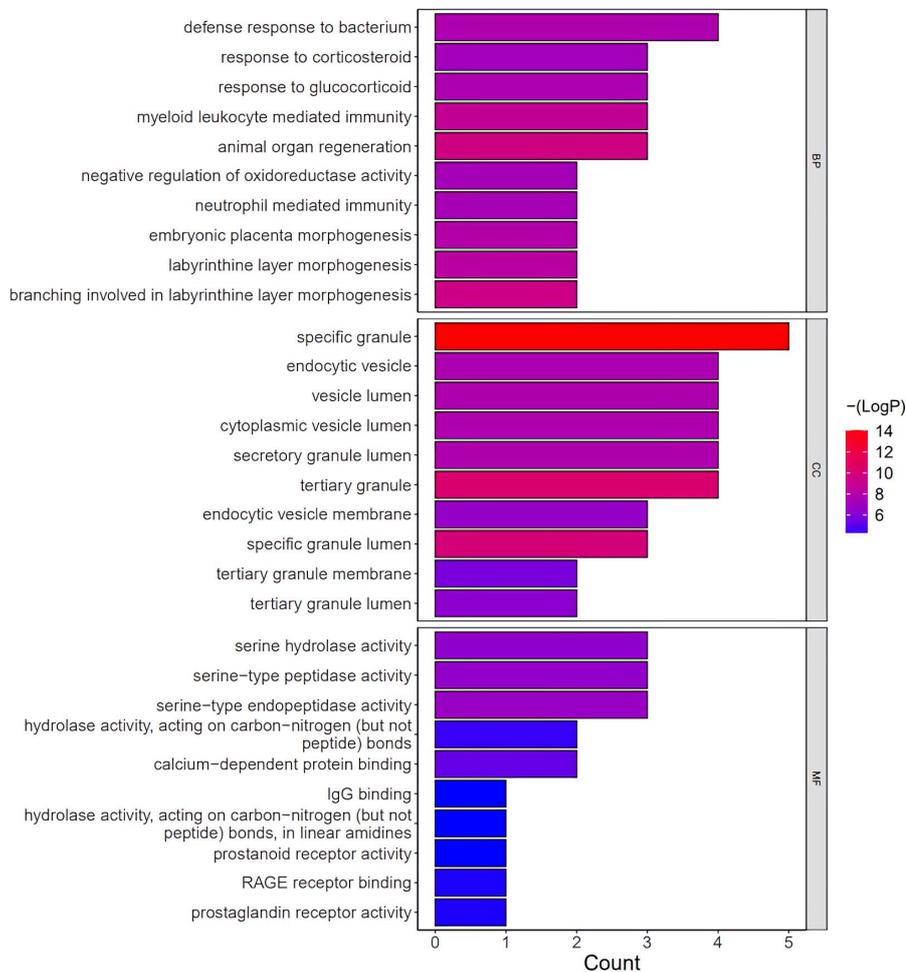
2.5.4 模型基因ROC外部验证结果在GSE66360数据集同样绘制4个基因的ROC曲线，结果见图15B，结果显示4个基因中S100A12，PTCH1和LOC400499基因的AUC值均大于0.7，综合模型的AUC值为0.95，说明S100A12，PTCH1和LOC400499这3个基因效果较为可靠，将这3个基因视为急性心肌梗死的潜在诊断生物标志物。

2.6 药物预测结果 S100A12基因输入STITCH数据库 (<http://stitch.embl.de/>) 后获得对应的药物靶点信息有10种，PTCH1基因在STITCH数据库结果得到1个药物。随后将PTCH1基因输入到herb数据库 (<http://herb.ac.cn/>) 中得到5个药物，S100A12基因在herb数据库结果得到1个药物。使用Cytoscape软件进行靶点-药物网络构建，如图16所示，Paquinimod, Amlexanox, quercitin及17-bate-estradiol等可能具有治疗急性心肌梗死的作用。

2.7 RT-qPCR验证关键基因表达结果 共收集了5例急性心肌梗死患者和5例对照组的外周血浆，为了验证数据集的可靠性，采用RT-qPCR来验证临床样本中上述3个基因的表达水平。结果显示，与正常对照组相比，急性心肌梗死患者中S100A12的表达上调 ($P < 0.05$)，见图17，PTCH1和LOC400499表达显著下调 ($P < 0.05$)，见图17。

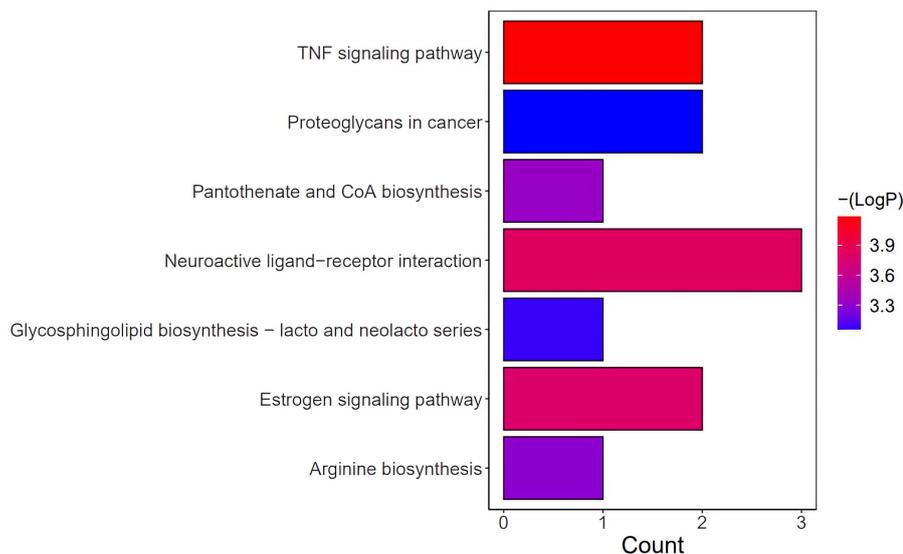
3 讨论 Discussion

急性心肌梗死发病后临床医师一般根据典型的临床症状、心电图和传统的生物标志物肌钙蛋白及肌酸激酶同工酶水平进行诊断，随后进行溶栓或经皮冠状动脉介入术等再灌注治疗^[20]。精准的早期诊断和及早治疗对于挽救患者生命和改善预后十分重要。而传统的生物标志物在外周血升高从而被检测到的时间为三四个小时，无法更早期识别急性心肌梗死患者。因此，寻找新的更早期诊断生物标志物非常重要。



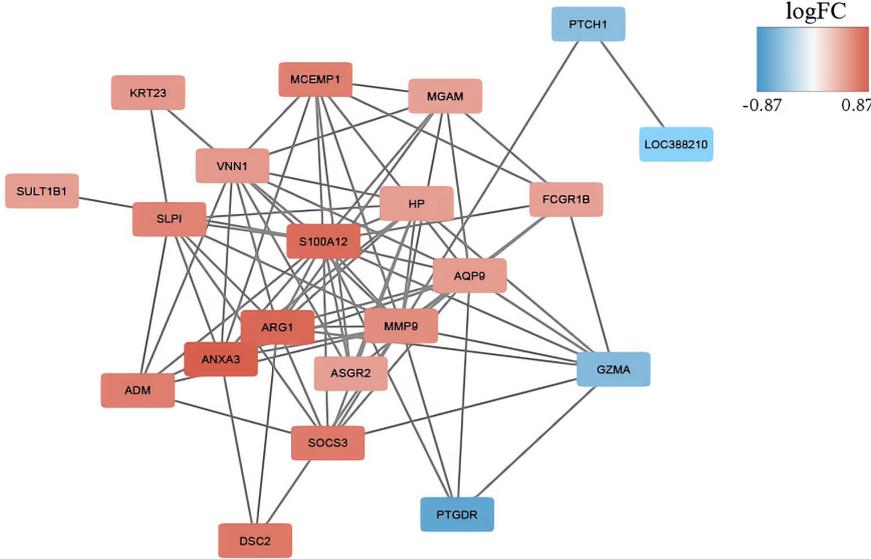
图注：纵坐标表示 GO 富集前 30 条分析结果，横坐标表示基因数目；BP 代表生物过程；CC 代表细胞成分；MF 代表分子功能。颜色表征差异的显著性。defense response to bacterium 为对细菌的防御反应；response to corticosteroid 为对皮质类固醇的反应；response to glucocorticoid 为对糖皮质激素的反应；myeloid leukocyte mediated immunity 为髓系白细胞介导免疫；animal organ regeneration 为动物器官再生；negative regulation of oxidoreductase activity 为氧化还原酶活性的负调控；neutrophil mediated immunity 为中性粒细胞介导免疫；embryonic placenta morphogenesis 为胚胎胎盘形态发生；labyrinthine layer morphogenesis 为迷路层形态发生；branching involved in labyrinthine layer morphogenesis 为分支参与迷路的形态发生；specific granule 为特定颗粒；endocytic vesicle 为内吞囊泡；vesicle lumen 为囊泡腔；cytoplasmic vesicle lumen 为细胞质泡腔；secretory granule lumen 为分泌颗粒腔；tertiary granule 为三级颗粒；endocytic vesicle membrane 为胞吞囊泡膜；specific granule lumen 为特定颗粒内腔；tertiary granule membrane 为三级颗粒膜；tertiary granule lumen 为三级颗粒管腔；serine hydrolase activity 为丝氨酸水解酶活性；serine-type peptidase activity 为丝氨酸型肽酶活性；serine-type endopeptidase activity 为丝氨酸型内肽酶活性；hydrolase activity, acting on carbon-nitrogen (but not peptide) bonds 为水解酶活性，作用于碳氮键（而不是肽键）；calcium-dependent protein binding 为钙依赖性蛋白结合；IgG binding 为 IgG 结合；hydrolase activity, acting on carbon-nitrogen (but not peptide) bonds, in linear amidines 为水解酶活性，作用于线性脲的碳氮键（而不是肽键）；prostanoid receptor activity 为前列腺素受体活性；RAGE receptor binding 为 RAGE 受体结合；prostaglandin receptor activity 为前列腺素受体活性。

图 10 | 对 24 个差异模块基因进行 GO 富集分析结果图



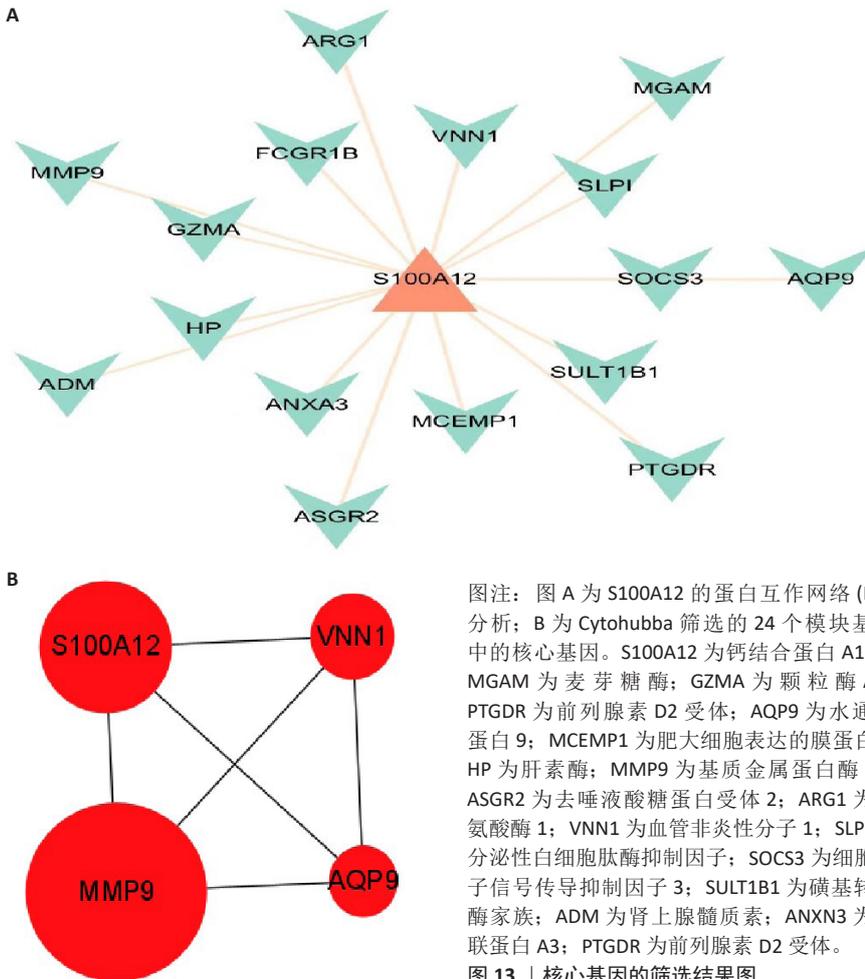
图注：横坐标表示基因数目，纵坐标表示显著富集的通路，颜色表征差异的显著性。TNF signaling pathway 为肿瘤坏死因子信号通路，Proteoglycans in cancer 为蛋白聚糖，Pantothenate and CoA biosynthesis 为泛酸盐和辅酶 A 生物合成；Neuroactive ligand-receptor interaction 为神经活性配体-受体相互作用，Glycosphingolipid biosynthesis-lacto and neolacto series 为鞘糖脂生物合成-乳酸和新乳酸系列，Estrogen signaling pathway 为雌激素信号通道，Arginine biosynthesis 为精氨酸生物合成。

图 11 | 对 24 个差异模块基因进行 KEGG 富集分析结果图



图注: LogFC 为对数差异倍数。S100A12 为钙结合蛋白 A12; PTCH1 为修补 1; LOC382210 为 THO 复合体 2; MGAM 为麦芽糖酶; FCGR1B 为 FC γ 受体 1B; GZMA 为颗粒酶 A; PTGDR 为前列腺素 D2 受体; AQP9 为水通道蛋白 9; MCEMP1 为肥大细胞表达的膜蛋白; HP 为肝素酶; MMP9 为基质金属蛋白酶 9; ASGR2 为去唾液酸糖蛋白受体 2; ARG1 为精氨酸酶 1; VNN1 为血管非炎性分子 1; KRT23 为角蛋白 23; SLPI 为分泌性白细胞肽酶抑制因子; SOCS3 为细胞因子信号传导抑制因子 3; DSC2 为桥粒糖蛋白 2; ANXA3 为膜联蛋白 A3; SULT1B1 为磺基转移酶家族; ADM 为肾上腺髓质素。

图 12 | 蛋白互作网络 (PPI) 分析结果图

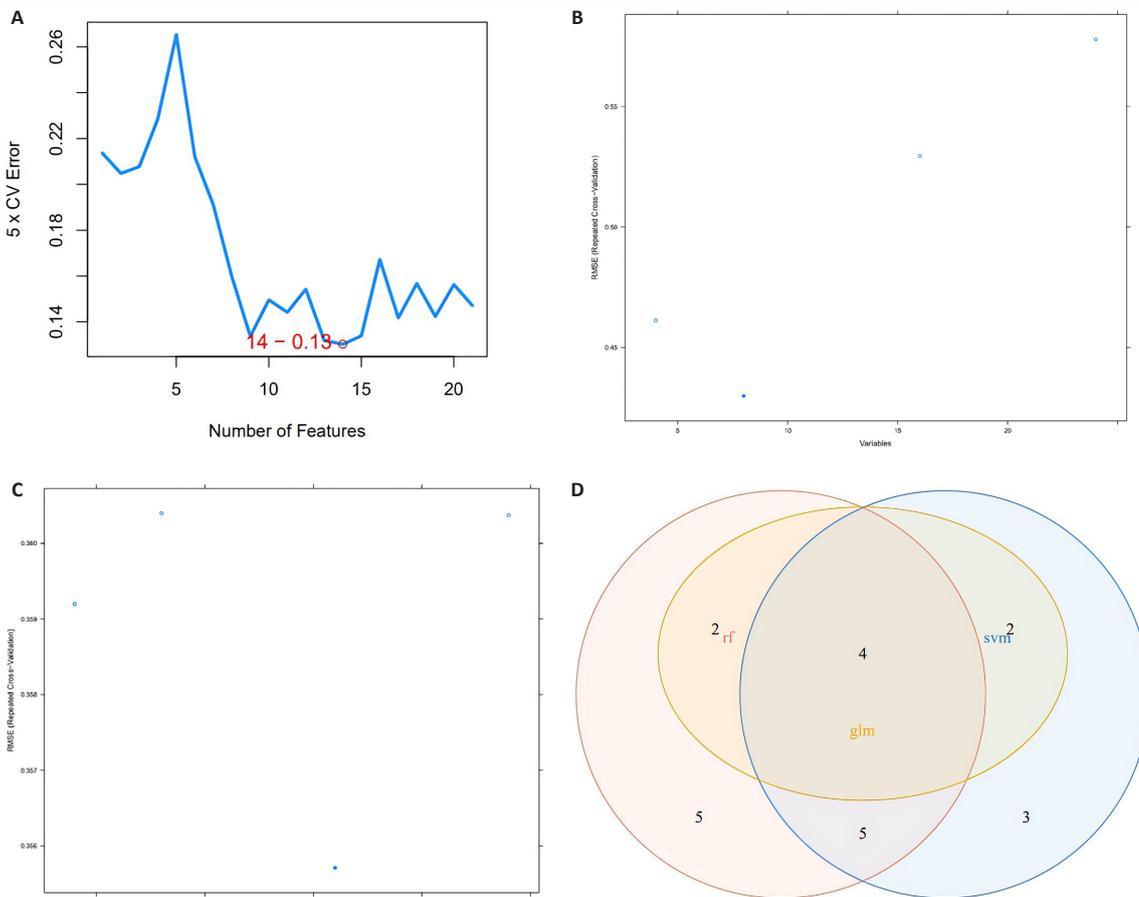


图注: 图 A 为 S100A12 的蛋白互作网络 (PPI) 分析; B 为 Cytohubba 筛选的 24 个模块基因中的核心基因。S100A12 为钙结合蛋白 A12; MGAM 为麦芽糖酶; GZMA 为颗粒酶 A; PTGDR 为前列腺素 D2 受体; AQP9 为水通道蛋白 9; MCEMP1 为肥大细胞表达的膜蛋白; HP 为肝素酶; MMP9 为基质金属蛋白酶 9; ASGR2 为去唾液酸糖蛋白受体 2; ARG1 为精氨酸酶 1; VNN1 为血管非炎性分子 1; SLPI 为分泌性白细胞肽酶抑制因子; SOCS3 为细胞因子信号传导抑制因子 3; SULT1B1 为磺基转移酶家族; ADM 为肾上腺髓质素; ANXA3 为膜联蛋白 A3; PTGDR 为前列腺素 D2 受体。

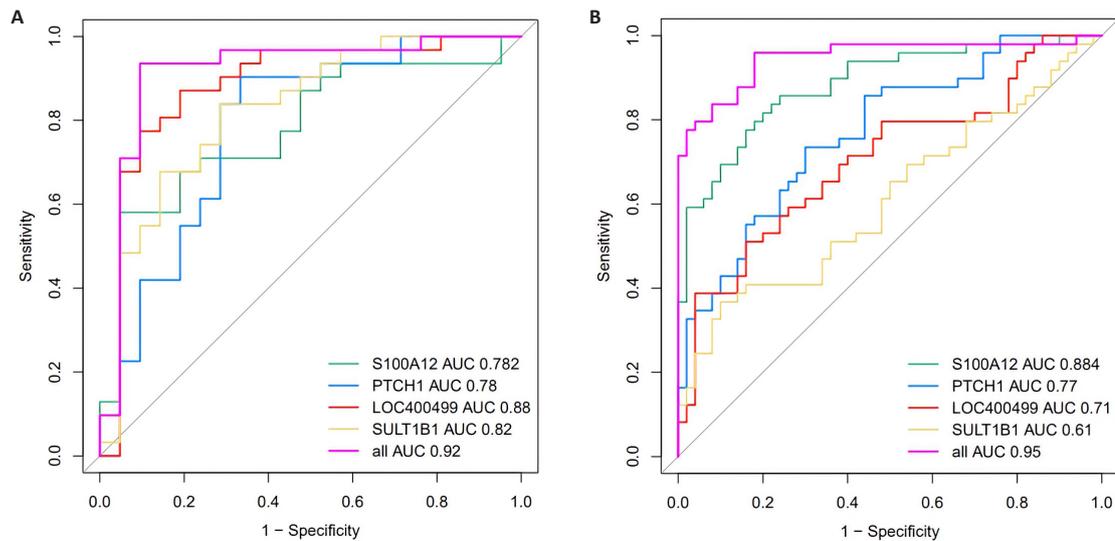
图 13 | 核心基因的筛选结果图

中性粒细胞是一种多形核细胞,被认为是先天免疫反应的主要参与者,与其分泌的炎症递质和蛋白酶在急性心肌梗死急性期和修复期中发挥着十分重要的作用。因此,文章使用了一系列生物信息学和机器学习的分析方法进行急性心肌梗死与中性粒细胞相关生物标志物的筛选,以期得到可能识别更早期急性心肌梗死的生物标志物以及新的与中性粒细胞相关的基因靶点。最终将 S100A12, PTCH1 和 LOC400499 视为诊断急性心肌梗死的潜在生物标志物。为了进一步探讨这 3 个基因在急性心肌梗死中的临床意义,文章收集了急性心肌梗死患者经皮冠状动脉介入术后的外周血,采用 RT-qPCR 检测这 3 个基因的表达水平,发现这 3 个基因在对照组和急性心肌梗死组中表达具有显著差异,可能是急性心肌梗死潜在的诊断生物标志物。随后,文章通过 Cytohubba 筛选 24 个差异模块基因中的核心基因,结合 3 种机器学习算法得到的生物标志物,文章推测,差异表达显著的 S100A12 可能在急性心肌梗死中起关键调控作用,有可能成为急性心肌梗死的潜在调控靶点。

S100A12 是一种在正常条件下主要存在于中性粒细胞的胞浆内的 S100 蛋白家族成员。当细胞受损和/或活化, S100A12 会被转移到胞膜然后分泌到胞外,进入血液循环并招募更多的中性粒细胞、单核细胞等到达炎症部位并活化,完成起始和调节炎症反应及诱导细胞凋亡等过程的发生^[21-22]。因此, S100A12 也被认为是促炎因子。大量的研究表明, S100A12 在急性心肌梗死中也发挥着十分重要的作用。在 ST 段抬高型心肌梗死患者中 S100A12 显著升高,胸痛后 2 h 内达到峰值^[23]。一项经过 52 周严格生活方式控制的研究中, 63 例患有心血管疾病或具有 2 种及以上心血管疾病危险因素参与者外周血中, S100A12 是减少显著的基因之一^[24], 这表明 S100A12 的表达是可调节的。一项利用 AC16 细胞构建缺血再灌注损伤模型的研究中发现 S100A12 表达水平大幅上调。随后,此研究进行 S100A12 敲除,发现敲除 S100A12 不仅能通过提高细胞活力和



图注：图 A 为支持向量模型算法筛选特征基因；B 为广义线性模型算法筛选特征基因；C 为随机森林模型算法筛选特征基因。图 A-C 横坐标均表示特征基因的个数，纵坐标均表示泛化误差；折线图线的趋势代表特征基因个数与泛化误差的关系。图 D 为支持向量模型、广义线性模型和随机森林模型算法的特征基因交集确定候选生物标志物。rf 代表随机森林模型；svm 代表支持向量模型；glm 代表广义线性模型。
图 14 | 通过机器学习方法识别候选生物标志物



图注：图 A, B 分别为 GSE48060 数据集、GSE66360 数据集中 S100A12, PTCH1, LOC400499, SULT1B1 4 个基因的 ROC 曲线分析。横坐标表示 1-特异性，纵坐标表示敏感度。

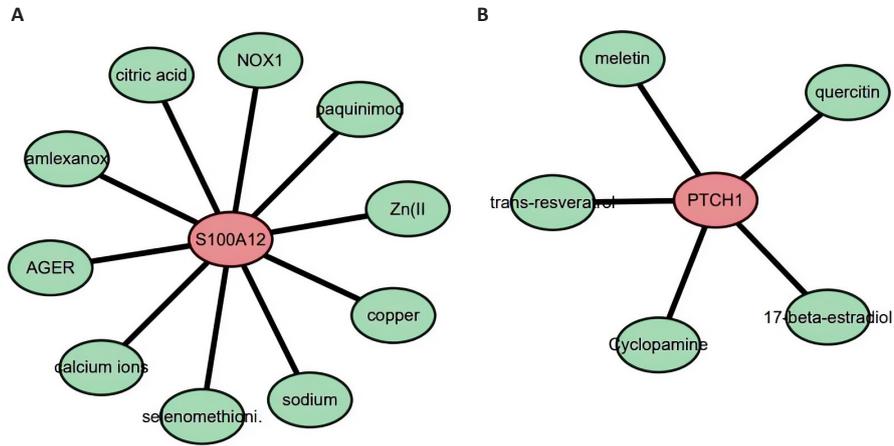
图 15 | ROC 曲线分析结果

减少乳酸脱氢酶的释放来抑制缺血再灌注孙损伤诱导的 AC16 细胞损伤，还能减少细胞凋亡以及一些炎性细胞因子的产生，恢复氧化-抗氧化因子的平衡，提示 S100A12 可能是治疗缺血再灌注损伤的潜在靶点^[25]。

PTCH1 除了在多种癌症中表达，

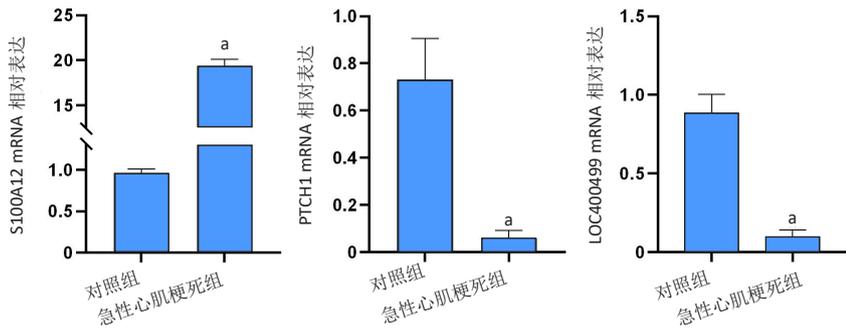
如肺癌、乳腺癌、前列腺癌、卵巢癌、结肠癌、脑癌、肾上腺皮质癌和黑色素瘤，在心肌细胞和血管平滑肌细胞也有表达，是 Sonic Hedgehog (Shh) 配体的受体^[26-28]。PTCH1 与 Shh 结合后激活下游信号分子参与冠状脉管系统和心脏功能的稳定维持。在急性心

肌梗死的大鼠中，阻断 PTCH1 的表达导致梗死面积增加、心功能障碍加剧，而过表达 PTCH1 能抑制心肌细胞凋亡，改善心功能，这可能是通过增强边界区脉管系统和最大限度减少组织缺氧来实现的，这些研究也表明，PTCH1 可能是急性心肌梗死的潜在治



图注：图 A 为可能靶向调控 S100A12 的药物：amlexanox 为氨来占诺；calcium ions 为钙离子；sodium 为钠；copper 为铜；paquinimod 为帕奎尼莫；crtric acid 为柠檬酸；AGER 为晚期糖基化终产物；NOX1 为氧化酶 1；seenomethicni。图 B 为可能靶向调控 PTCH1 的药物：quercitin 为槲皮素；17-bate-estradiol 为 17β 雌二醇；trans-resverao 为白藜芦醇；cyclopamine 为氯丙咪嗪；meletin 为槲皮素。

图 16 | 靶向药物预测结果



图注：与对照组相比，^a $P < 0.05$ 。

图 17 | 关键基因表达的 RT-qPCR 分析结果

疗靶点^[29-30]。LOC400499 在人体多种组织中低表达，然而很少有研究关注 LOC400499 在心血管疾病方面的作用，在急性心肌梗死中的生物学作用仍不清楚。因此，需进一步研究来验证文章的发现。

文章还预测了基因的靶向药物，以期辅助急性心肌梗死的治疗。通过靶点预测帕奎莫德、氨来占诺、槲皮素和 17β 雌二醇等可能具有治疗急性心肌梗死的作用。帕奎莫德也称为 ABR-215757，在动脉粥样硬化中通过影响 S100A12 表达减少病变，减轻炎症细胞浸润，缩小坏死核心以及降低钙化和弹性纤维降解程度^[31]。另外，氨来占诺和槲皮素对急性心肌梗死后大鼠的心功能发挥保护作用^[32-33]，但这些基因靶向药物还需进一步研究证实，以期急性心肌梗死治疗带来最大效益。

然而，文章存在一定的局限性。

首先，样本总数相对较少，这可能会限制差异基因的发现，降低研究的准确性。其次，数据集外周血的收集时间是在紧急经皮冠状动脉介入术前，而文章的血浆样本是在紧急经皮冠状动脉介入术后立即收集的，不同的时间和干预因素可能会对 S100A12 等蛋白水平产生影响。文章对 S100A12、PTCH1 和 LOC400499 mRNA 相对表达的测定，其时效性和成本性及可及性，尚需进一步进行大规模、前瞻性、国际多中心研究进一步验证。

综上所述，通过生物信息学分析和机器学习方法确认了 3 个与急性心肌梗死和中性粒细胞相关的基因，分别为 S100A12、PTCH1 和 LOC400499，可能是更早期诊断急性心肌梗死的生物标志物，但是其与急性心肌梗死相关的特异性尚需进一步研究，其中 S100A12 可能是调控急性心肌梗死的潜在靶点。

作者贡献：文章设计者为余振球和杨定燕。资料收集和数据分析师为杨定燕和杨中愉。杨定燕撰写论文。余振球审核。

利益冲突：文章的全部作者声明，在课题研究过程和文章撰写过程中不存在利益冲突。

开放获取声明：这是一篇开放获取文章，根据《知识共享许可协议》“署名-非商业性使用-相同方式共享 4.0”条款，在合理引用的情况下，允许他人以非商业性目的基于原文内容编辑、调整和扩展，同时允许任何用户阅读、下载、拷贝、传递、打印、检索、超级链接该文献，并为之建立索引，用作软件的输入数据或其它任何合法用途。

版权转让：文章出版前全体作者与编辑部签署了文章版权转让协议。

出版规范：该文章撰写遵守国际医学期刊编辑委员会《学术研究实验与报告和医学期刊编辑与发表的推荐规范》。文章出版前已经过专业反剽窃文献检测系统进行 3 次查重。文章经小同行外审专家双盲外审，同行评议认为文章符合期刊发表宗旨。

4 参考文献 References

- [1] MEHTA LS, BECKIE TM, DEVON HA, et al. Acute myocardial infarction in women: a scientific statement from the american heart association. *Circulation*. 2016;133(9):916-947.
- [2] PAGIDIPATI NJ, GAZIANO TA. Estimating deaths from cardiovascular disease: a review of global methodologies of mortality measurement. *Circulation*. 2013; 127(6):749-756.
- [3] ALAOUR B, LIEW F, KAIER TE. Cardiac Troponin-diagnostic problems and impact on cardiovascular disease. *Ann Med*. 2018; 50(8):655-665.
- [4] THYGESSEN K, ALPERT JS, JAFFE AS, et al. Fourth universal definition of myocardial infarction (2018). *Circulation*. 2018;138(20):e618-e651.
- [5] CLERICO A, ZANINOTTO M, PLEBANI M. Rapid rule-in and rule-out protocols of acute myocardial infarction using hs-cTnl and hs-cTnT methods. *Clin Chem Lab Med*. 2024;62(2):213-217.
- [6] PRABHU SD, FRANGOGIANNIS NG. The biological basis for cardiac repair after myocardial infarction: from inflammation to fibrosis. *Circ Res*. 2016;119(1):91-112.
- [7] EL KM, SHI H, VUONG S, et al. Nitroxides mitigate neutrophil-mediated damage to the myocardium after experimental myocardial infarction in rats. *Int J Mol Sci*. 2020;21(20):7650.

- [8] DELEON-PENNELL KY, TIAN Y, ZHANG B, et al. CD36 Is a Matrix metalloproteinase-9 substrate that stimulates neutrophil apoptosis and removal during cardiac remodeling. *Circ Cardiovasc Genet.* 2016; 9(1):14-25.
- [9] HORCKMANS M, RING L, DUCHENE J, et al. Neutrophils orchestrate post-myocardial infarction healing by polarizing macrophages towards a reparative phenotype. *Eur Heart J.* 2017;38(3):187-197.
- [10] CURAJ A, SCHUMACHER D, RUSU M, et al. Neutrophils modulate fibroblast function and promote healing and scar formation after murine myocardial infarction. *Int J Mol Sci.* 2020;21(10):3685.
- [11] SREEJIT G, ABDEL-LATIF A, ATHMANATHAN B, et al. Neutrophil-derived s100a8/a9 amplify granulopoiesis after myocardial infarction. *Circulation.* 2020;141(13):1080-1094.
- [12] CHRISTIA P, BUJAK M, GONZALEZ-QUESADA C, et al. Systematic characterization of myocardial inflammation, repair, and remodeling in a mouse model of reperfused myocardial infarction. *J Histochem Cytochem.* 2013;61(8):555-570.
- [13] SWIRSKI FK, NAHRENDORF M. Leukocyte behavior in atherosclerosis, myocardial infarction, and heart failure. *Science.* 2013;339(6116):161-166.
- [14] KOLOGRIVOVA I, SHATOLKINA M, SUSLOVA T, et al. Cells of the immune system in cardiac remodeling: main players in resolution of inflammation and repair after myocardial infarction. *Front Immunol.* 2021;12:664457.
- [15] ZHOU X, ZHANG C, WU X, et al. Dusp6 deficiency attenuates neutrophil-mediated cardiac damage in the acute inflammatory phase of myocardial infarction. *Nat Commun.* 2022;13(1):6672.
- [16] RIDKER PM, MACFADYEN JG, GLYNN RJ, et al. Inhibition of interleukin-1beta by canakinumab and cardiovascular outcomes in patients with chronic kidney disease. *J Am Coll Cardiol.* 2018;71(21):2405-2414.
- [17] ZHENG PF, ZOU QC, CHEN LZ, et al. Identifying patterns of immune related cells and genes in the peripheral blood of acute myocardial infarction patients using a small cohort. *J Transl Med.* 2022;20(1):321.
- [18] TORUN FM, VIRREIRA WS, DOLL S, et al. Transparent exploration of machine learning for biomarker discovery from proteomics and omics data. *J Proteome Res.* 2023;22(2):359-367.
- [19] MIAO M, CAO S, TIAN Y, et al. Potential diagnostic biomarkers: 6 cuproptosis- and ferroptosis-related genes linking immune infiltration in acute myocardial infarction. *Genes Immun.* 2023;24(4):159-170.
- [20] ALGOET M, JANSSENS S, HIMMELREICH U, et al. Myocardial ischemia-reperfusion injury and the influence of inflammation. *Trends Cardiovasc Med.* 2023;33(6):357-366.
- [21] CARVALHO A, LU J, FRANCIS JD, et al. S100A12 in digestive diseases and health: a scoping review. *Gastroenterol Res Pract.* 2020;2020:2868373.
- [22] ZHAI H, HUANG L, GONG Y, et al. Human plasma transcriptome implicates dysregulated s100a12 expression: a strong, early-stage prognostic factor in st-segment elevated myocardial infarction: bioinformatics analysis and experimental verification. *Front Cardiovasc Med.* 2022;9:874436.
- [23] ZHANG X, CHENG M, GAO N, et al. Utility of S100A12 as an early biomarker in patients with ST-segment elevation myocardial infarction. *Front Cardiovasc Med.* 2021;8:747511.
- [24] ELLSWORTH DL, CROFT DJ, WEYANDT J, et al. Intensive cardiovascular risk reduction induces sustainable changes in expression of genes and pathways important to vascular function. *Circ Cardiovasc Genet.* 2014;7(2):151-160.
- [25] LI Q, DENG G, GAO Y. S100 calcium-binding protein A12 knockdown ameliorates hypoxia-reoxygenation-induced inflammation and apoptosis in human cardiomyocytes by regulating caspase-4-mediated non-classical pyroptosis. *Gen Physiol Biophys.* 2022;41(4):287-297.
- [26] HASANOVIC A, MUS-VETEAU I. Targeting the multidrug transporter ptch1 potentiates chemotherapy efficiency. *Cells.* 2018;7(8):107.
- [27] DOHENY D, MANORE SG, WONG GL, et al. Hedgehog signaling and truncated GLI1 in cancer. *Cells.* 2020;9(9):2114.
- [28] POLA R, LING LE, SILVER M, et al. The morphogen Sonic hedgehog is an indirect angiogenic agent upregulating two families of angiogenic growth factors. *Nat Med.* 2001;7(6):706-711.
- [29] LI SH, ZHANG YY, SUN YL, et al. Inhibition of microRNA-802-5p inhibits myocardial apoptosis after myocardial infarction via Sonic Hedgehog signaling pathway by targeting PTCH1. *Eur Rev Med Pharmacol Sci.* 2021;25(1):326-334.
- [30] LAVINE KJ, KOVACS A, ORNITZ DM. Hedgehog signaling is critical for maintenance of the adult coronary vasculature in mice. *J Clin Invest.* 2008; 118(7):2404-2414.
- [31] YAN L, BJORK P, BUTUC R, et al. Beneficial effects of quinoline-3-carboxamide (ABR-215757) on atherosclerotic plaque morphology in S100A12 transgenic ApoE null mice. *Atherosclerosis.* 2013;228(1):69-79.
- [32] ALBADRANI GM, BINMOWYNA MN, BIN-JUMAH MN, et al. Quercetin prevents myocardial infarction adverse remodeling in rats by attenuating TGF-beta1/Smad3 signaling: different mechanisms of action. *Saudi J Biol Sci.* 2021;28(5):2772-2782.
- [33] MO C, HAN H, TANG X, et al. Protein kinase TBK1/IKKepsilon inhibitor Amlexanox improves cardiac function after acute myocardial infarction in rats. *Panminerva Med.* 2023;65(3):343-350.

(责任编辑: WJ, ZN, QY, ZLJ)

急性心肌梗死与中性粒细胞相关潜在生物标志物的机器学习分析

文章特色分析

一、文章重要性

1. 临床诊断需求迫切

急性心肌梗死是全球范围内致死率极高的心血管急症，早期诊断对挽救生命至关重要。目前依赖的肌钙蛋白等生物标志物存在时间滞后性，无法满足“超早期”诊断需求。本研究致力于寻找更早、更特异的新型生物标志物，具有明确的临床转化价值。

2. 免疫机制深入挖掘

研究聚焦于中性粒细胞在急性心肌梗死过程中的关键作用，揭示其不仅是炎症反应的执行者，更参与心肌修复与重塑，为理解心肌梗死的免疫病理机制提供了新视角。

二、创新性特色

1. 多组学与机器学习融合分析

研究整合了：

- 差异表达分析
- 免疫细胞浸润分析
- WGCNA 共表达网络构建
- 蛋白互作网络分析
- 三种机器学习算法（SVM、GLM、随机森林）筛选特征基因

这种多层次、多方法的整合分析，显著提升了生物标志物筛选的可靠性和解释力。

2. 发现新型候选生物标志物

最终锁定三个基因：S100A12、PTCH1、LOC400499，并在独立数据集和临床样本中验证其表达差异与诊断效能（AUC > 0.7），具备成为新型诊断标志物的潜力。

3. 药物靶点预测

通过 STITCH 和 Herb 数据库预测了 S100A12 和 PTCH1 的潜在靶向药物（如帕奎尼莫、氨来咕诺、槲皮素等），为后续药物研发提供了方向。

4. 实验验证增强说服力

通过 RT-qPCR 在临床样本中验证了三个基因在急性心肌梗死患者外周血中的表达变化，增强了研究结果的临床相关性。

三、对学科的启示

1. 推动“免疫-心血管”交叉研究

本研究将免疫细胞（中性粒细胞）与心血管疾病（急性心肌梗死）紧密结合，强调了免疫调节在心血管疾病防治中的重要性，为“免疫心脏病学”提供了实证支持。

2. 展示机器学习在生物标志物挖掘中的优势

研究展示了机器学习在处理高维、复杂生物数据方面的强大能力，为今后在其它疾病中开展类似研究提供了方法论参考。

3. 促进精准医疗与个性化治疗

通过识别特异性基因标志物并预测靶向药物，本研究为急性心肌梗死的早期诊断和靶向治疗提供了新思路，符合精准医学的发展趋势。

4. 强调多中心、大样本验证的必要性

作者也指出研究的局限性（样本量小、时间点单一），提示未来需开展更大规模、前瞻性、多中心的研究以进一步验证这些标志物的临床价值。

总结

本研究通过生物信息学+机器学习+实验验证的系统性策略，成功筛选并初步验证了 S100A12、PTCH1、LOC400499 作为急性心肌梗死早期诊断与治疗的潜在生物标志物。其创新性在于方法的整合性与结果的转化潜力，对推动急性心肌梗死的早期诊断、机制理解及靶向治疗具有重要学术与临床意义。