

基于机器学习算法鉴定哮喘的坏死性凋亡 相关生物标志物

王静¹, 刘晓菲², 曾荣³, 许长娟², 张锦涛², 董亮^{2,3}

(1. 吴忠市人民医院呼吸与危重症医学 I 科, 宁夏 吴忠 751199;

2. 山东第一医科大学第一附属医院(山东省千佛山医院)呼吸与危重症医学科 山东省呼吸疾病研究所, 山东 济南 250014;

3. 山东大学 山东省千佛山医院呼吸与危重症医学科, 山东 济南 250014)

摘要: **目的** 通过生物信息学分析的方法识别和验证哮喘中潜在的坏死性凋亡相关基因(necroptosis-related genes, NRGs)。**方法** 基因表达综合(Gene Expression Omnibus, GEO)数据库提供了基因表达谱数据集GSE76262, 并使用R软件筛选潜在的差异表达NRGs。对差异表达的NRGs进行蛋白质-蛋白质相互作用(protein-protein interaction, PPI)分析、基因本体论(gene ontology, GO)富集分析、京都基因与基因组百科全书(Kyoto encyclopedia of genes and genomes, KEGG)通路富集分析。通过最小绝对收缩和选择算子(least absolute shrinkage and selection operator, LASSO)逻辑回归和支持向量机递归特征消除(support vector machine-recursive feature elimination, SVM-RFE)算法分析上调的NRGs。共同特征基因被确定为潜在的诊断标志物, 并绘制受试者特征(receiver operator characteristic, ROC)曲线以验证其诊断效能。GSE137268进一步验证筛选出的特征基因的表达情况和诊断效能。使用在线工具预测可以靶向调控特征基因表达的微小RNAs(miRNAs)。**结果** 在118例哮喘患者和21名健康对照者中鉴定出33个差异表达的NRGs(13个上调和20个下调)。PPI结果提示20个差异表达的NRGs相互作用。GO和KEGG富集分析显示NRGs与多个信号通路、淋巴细胞激活、细胞凋亡和免疫调节等相关。LASSO和SVM-RFE筛选出7个上调的NRGs可作为潜在的诊断基因, 其ROC曲线显示出较高的诊断效率, 曲线下面积(area under the curve, AUC)高于0.7。经GSE137268验证, 7个特征基因与训练集的表达趋势相同(AUC>0.65)。预测hsa-miR-138-5p、hsa-miR-200b-3p和hsa-miR-30e-5p可调控哮喘患者NRGs的表达。**结论** BIRC3、HIF1A、FLOT1、NLRP3、RIPK2、GBE1和PELI1为哮喘的潜在生物标志物。Hsa-miR-138-5p、hsa-miR-200b-3p和hsa-miR-30e-5p分别是HIF1A、RIPK2和PELI1的上游调节因子。

关键词: 坏死性凋亡; 哮喘; 生物信息学分析; 机器学习; 生物标志物

中图分类号: R562

文献标志码: A

Identification of necroptosis-related biomarkers in asthma based on machine learning algorithms

WANG Jing¹, LIU Xiaofei², ZENG Rong³, XU Changjuan², ZHANG Jintao², DONG Liang^{2,3}

(1. Department of Respiratory I, Wuzhong People's Hospital, Wuzhong 751199, Ningxia, China;

2. Department of Respiratory, The First Affiliated Hospital of Shandong First Medical University & Shandong Provincial Qianfoshan Hospital, Shandong Institute of Respiratory, Jinan 250014, Shandong, China; 3. Department of Respiratory,

Shandong Provincial Qianfoshan Hospital, Shandong University, Jinan 250014, Shandong, China)

Abstract: Objective To identify and validate of potential necroptosis-related genes (NRGs) in asthma through bioinformatics analysis. **Methods** The Gene Expression Omnibus (GEO) database provided the gene expression profile dataset GSE76262, and R software was used to screen for potential differentially expressed NRGs. Protein-protein

interactions (PPI) analysis, gene ontology (GO) enrichment analysis, and Kyoto encyclopedia of genes and genomes (KEGG) pathway enrichment analysis were performed on the differentially expressed NRGs. The up-regulated NRGs were analyzed by least absolute shrinkage and selection operator (LASSO) logistic regression and support vector machine-recursive feature elimination (SVM-RFE) algorithms. Common signature genes were identified as potential diagnostic markers, and receiver operator characteristic (ROC) curves were drawn to verify their diagnostic efficacy. The expression and diagnostic efficacy of the screened signature genes were further confirmed by GSE137268. Online tools were used to predict microRNAs (miRNAs) that can target and regulate the expression of signature genes.

Results Thirty-three differentially expressed NRGs (13 up-regulated and 20 down-regulated) were identified in 118 asthma patients and 21 healthy controls. The PPI results indicated that 20 differentially expressed NRGs interacting. GO and KEGG enrichment analyses revealed enrichment items related to multiple signaling pathways, lymphocyte activation, apoptosis, and immune regulation. LASSO and SVM-RFE showed that seven up-regulated NRGs could be potential diagnosis genes. The ROC curves showed high diagnostic efficiency with the area under the curve (AUC) higher than 0.7. GSE137268 verified that seven signature genes showed the same expression trend as the training set (AUC > 0.65). Hsa-miR-138-5p, hsa-miR-200b-3p and hsa-miR-30e-5p were predicted to regulate the expression of necroptosis genes in asthmatic patients. **Conclusion** BIRC3, HIF1A, FLOT1, NLRP3, RIPK2, GBE1 and PELI1 can serve as potential biomarkers for asthma. Hsa-miR-138-5p, hsa-miR-200b-3p and hsa-miR-30e-5p are upstream regulators of HIF1A, RIPK2 and PELI1, respectively.

Key words: Necroptosis; Asthma; Bioinformatics analysis; Machine learning; Biomarkers

支气管哮喘(简称哮喘)是一种在发病年龄、相关危险因素、严重程度、合并症和治疗反应等方面具有异质性的慢性气道疾病^[1]。哮喘临床表现为随时间和强度变化的喘息、气短、胸闷和咳嗽等呼吸道症状以及伴有不同程度的呼气相气流受限,严重时气流受限可能会持续存在^[2]。越来越多的证据表明,哮喘的发病涉及多种病理机制,气道炎症是疾病发展的关键环节,可导致黏液产生、气道高反应性和气道壁重塑^[3]。

坏死性凋亡是细胞程序性死亡(programmed cell death, PCD)模式之一,是在细胞凋亡被阻断时激活的细胞自杀的过程^[4]。坏死性凋亡的重要形态特征是细胞膜破坏,重要介质是受体相互作用蛋白激酶(receptor interacting protein kinase, RIPK),混合谱系激酶结构域样蛋白(mixed lineage kinase domain-like protein, MLKL)是细胞坏死性凋亡的 executor。目前公认的坏死性凋亡机制由多种细胞因子或模式识别受体(pattern recognition receptors, PRR)触发,通过多种先天免疫信号通路以依赖或独立的方式激活 RIPK3。RIPK3 被磷酸化后,其下游的 MLKL 发生磷酸化而被活化。磷酸化的 MLKL 在细胞质膜上发生寡聚化,导致细胞膜变得不稳定或膜完整性遭到破坏,从而导致细胞坏死^[5-6]。坏死性凋亡的过程受多种机制调控,如凋亡抑制蛋白对 RIPK1 的泛素化、肿瘤坏死因子 α 诱导蛋白 3 (tumor necrosis factor alpha induced protein 3, TNFAIP3, 又称为锌指蛋白 A20)对 RIPK1 的去泛素化

以及 caspase-8 对由 RIPK1 和 RIPK3 组成的坏死小体的可变剪接等^[7-8]。与细胞凋亡不同,坏死性凋亡导致细胞内容物释放,引发一系列炎症反应,其在肺部疾病中的作用也已在一些研究中得到证实^[9-10]。然而,坏死性凋亡相关基因(necroptosis-related genes, NRGs)在哮喘中的表达和调控在很大程度上仍是未知的,需要进一步探索。

诱导痰是获取支气管上皮细胞标本的一种非侵入性且成熟的方法。痰基因表达谱可用于了解哮喘中的气道炎症^[11]。Kuo 等^[12]和 Baines 等^[13]分别使用诱导痰样本完成了哮喘患者和健康个体的基因表达谱数据集 GSE76262 和 GSE137268,并分析了哮喘患者和健康个体之间差异表达的基因。本研究拟使用这两个数据集的表达矩阵探索哮喘中潜在的 NRGs,为诊断和治疗哮喘提供潜在的生物标志物。

1 资料与方法

1.1 资料

从基因表达综合(Gene Expression Omnibus, GEO)数据库(<http://www.ncbi.nlm.nih.gov/geo/>)下载 GSE76262 和 GSE137268 的表达谱数据集。GSE76262 作为训练集,包括 118 例哮喘患者(哮喘组)和 21 名对照者(健康对照组)的诱导痰样本。GSE137268 作为验证集,包括 54 例哮喘患者和 15 名对照者的诱导痰样本。从人类基因数据库(<https://www.genecards.org/>)共检索到 614 个

NRGs。图 1 为研究的流程图。

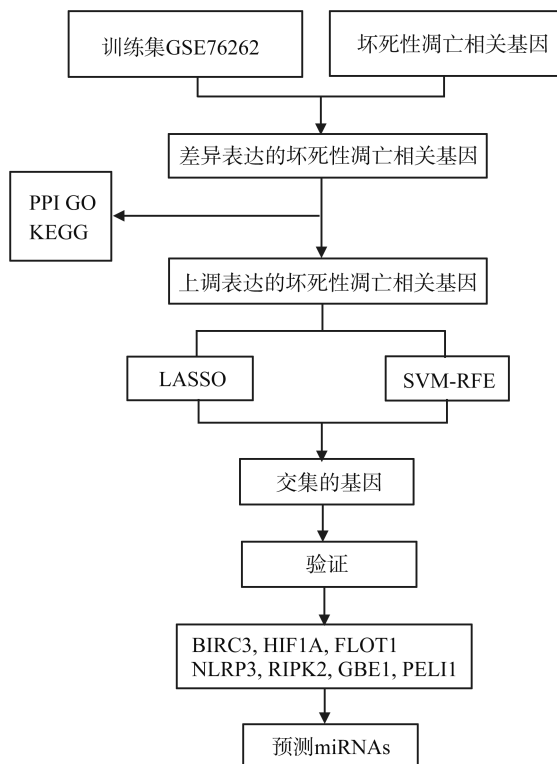


图 1 本研究的流程图

Figure 1 Flowchart of this study

1.2 方法

1.2.1 筛选差异表达的 NRGs

微阵列数据的标准化表达矩阵来自数据集 GSE76262, 使用注释文件对基因探针进行注释。R 软件的“limma”包(4.1.3 版)用于识别和提取差异表达 NRGs 的表达矩阵。从差异表达分析中获得 P 值, 通过使用控制错误发现率(false discovery rate, FDR)的多重假设检验进行校正, 以获得调整后的 P 值(Q 值)^[14]。差异表达基因的筛选标准为 $Q < 0.05$ 和 $|\log_2 FC| > 0.58$ ^[15]。差异表达的 NRGs 使用“pheatmap”和“ggplot2”包进行可视化, 以热图和火山图的形式呈现。

1.2.2 差异表达的 NRGs 的蛋白质-蛋白质相互作用 (protein-protein interaction, PPI) 网络分析

STRING 数据库 (<https://string-db.org/>) 是一个基于公共数据库和文献信息的蛋白质相互作用网络数据库, 通过对数据库及文献中存在的物理相互作用和功能关联进行收集并整合成 PPI 网络。通过 STRING 数据库构建差异表达 NRGs 的 PPI 网络, 过滤参数: 置信分数 > 0.4 , 隐藏无关联基因。Cytoscape 软件(版本 3.8.2)对 PPI 网络进行可视化。

1.2.3 NRGs 的功能分析

使用“clusterProfiler”包对差异表达的 NRGs 进

行基因本体论(gene ontology, GO)和京都基因与基因组百科全书(Kyoto encyclopedia of genes and genomes, KEGG)富集分析。 $Q < 0.05$ 被认为显著富集, 以气泡图的形式可视化。

1.2.4 机器学习算法筛选特征基因

基于 PPI 网络分析结果和 NRGs 的差异表达, 选择 GSE76262 中上调的 NRGs 表达矩阵进一步分析, 以筛选潜在的哮喘诊断标志物。分别使用“glmnet”包和“e1071”包进行最小绝对收缩和选择算子(least absolute shrinkage and selection operator, LASSO)回归和支持向量机递归特征消除(support vector machine-recursive feature elimination, SVM-RFE)算法分析。“Venn”包用于获取上述两种算法得到的相交特征基因作为标记基因。使用“pROC”包绘制特征基因的受试者特征(receiver operator characteristic, ROC)曲线。

1.2.5 疾病特征基因的验证

在 GSE137268 中使用“limma”和“ggpubr”包对筛选出的特征基因进行验证, 以箱线图显示它们在哮喘组和健康对照组中的表达水平, 对表达具有显著差异($P < 0.05$)的基因进一步绘制 ROC 曲线以明确诊断效能。

1.2.6 预测特征基因的 miRNA

通过4个常用的在线预测工具,即 ENCORI (<https://starbase.sysu.edu.cn/>)、miRBD (<http://mirdb.org/>)、RNAInter (<https://www.rna-society.org/>)和 Targetscan (<https://www.targetscan.org/>),寻找可以靶向和调控特征基因表达的 miRNA。气道活检样本的 microRNA 表达谱来自 GSE142237 中的哮喘患者和健康对照者^[16]。差异表达的 miRNA 的筛选标准为 $P < 0.05$ 和 $|\log_2 FC| > 0.58$ 。将4个数据库中预测的相同 miRNAs 进行重叠并标记,再与哮喘患者差异表达的 miRNAs 进行标记,得

到相同的 miRNAs 即为能够靶向调控哮喘患者 NRGs 的 miRNAs。

2 结果

2.1 NRGs 在哮喘中的差异表达分析

共鉴定出33个差异表达的 NRGs,包括13个显著上调基因和20个显著下调基因(表1),以火山图(图2A)和热图(图2B)呈现。其中,表达显著上调($|\log_2 FC| > 1$)的 NRGs 为 TNFAIP3、NLRP3、BIRC3 和 PELI1。

表1 哮喘组与健康对照组中33个差异表达的 NRGs

Table 1 The 33 differentially expressed NRGs in asthma samples compared to healthy samples

基因	$\log_2 FC$	差异表达	P	Q
TNFAIP3	1.070 678	上调	5.11×10^{-5}	2.05×10^{-3}
NLRP3	1.069 095	上调	6.51×10^{-5}	2.09×10^{-3}
BIRC3	1.006 980	上调	5.26×10^{-6}	8.08×10^{-4}
PELI1	1.006 188	上调	2.50×10^{-5}	1.50×10^{-3}
TNIP1	0.734 968	上调	6.16×10^{-4}	6.30×10^{-3}
FLOT1	0.729 068	上调	1.13×10^{-5}	9.07×10^{-4}
SLC25A37	0.709 180	上调	4.07×10^{-3}	1.78×10^{-2}
CCL2	0.691 406	上调	1.40×10^{-2}	4.27×10^{-2}
HIF1A	0.661 928	上调	7.17×10^{-6}	8.08×10^{-4}
NFKBIA	0.634 767	上调	2.81×10^{-4}	4.82×10^{-3}
TUBA4A	0.628 316	上调	4.75×10^{-4}	6.01×10^{-3}
GBE1	0.609 018	上调	9.85×10^{-3}	3.38×10^{-2}
RIPK2	0.590 607	上调	2.04×10^{-4}	4.27×10^{-3}
CARD6	-0.593 34	下调	2.03×10^{-3}	1.27×10^{-2}
CRTAM	-0.604 12	下调	6.72×10^{-3}	2.54×10^{-2}
DAPK1	-0.606 45	下调	7.56×10^{-4}	6.93×10^{-3}
RBMX	-0.608 95	下调	6.36×10^{-5}	2.09×10^{-3}
RPS27L	-0.610 66	下调	3.22×10^{-3}	1.54×10^{-2}
CCT6A	-0.612 95	下调	3.22×10^{-4}	4.82×10^{-3}
SRP14	-0.621 22	下调	2.92×10^{-5}	1.50×10^{-3}
YWHAG	-0.634 87	下调	3.10×10^{-4}	4.82×10^{-3}
PDCD6IP	-0.638 24	下调	2.38×10^{-4}	4.62×10^{-3}
NRIP1	-0.639 64	下调	2.05×10^{-3}	1.27×10^{-2}
CD74	-0.641 22	下调	3.03×10^{-4}	4.82×10^{-3}
PYCARD	-0.663 69	下调	4.99×10^{-5}	2.05×10^{-3}
GLTP	-0.676 30	下调	1.67×10^{-6}	4.70×10^{-4}
DYNLL1	-0.683 64	下调	2.57×10^{-4}	4.82×10^{-3}
PRDX1	-0.722 75	下调	2.71×10^{-4}	4.82×10^{-3}
NDUFA4	-0.753 07	下调	6.69×10^{-5}	2.09×10^{-3}
AXL	-0.820 12	下调	7.08×10^{-4}	6.93×10^{-3}
TNFSF12	-0.841 52	下调	2.71×10^{-5}	1.50×10^{-3}
TRPC6	-0.922 77	下调	1.01×10^{-5}	9.07×10^{-4}
MYO6	-0.980 31	下调	3.30×10^{-7}	1.86×10^{-4}

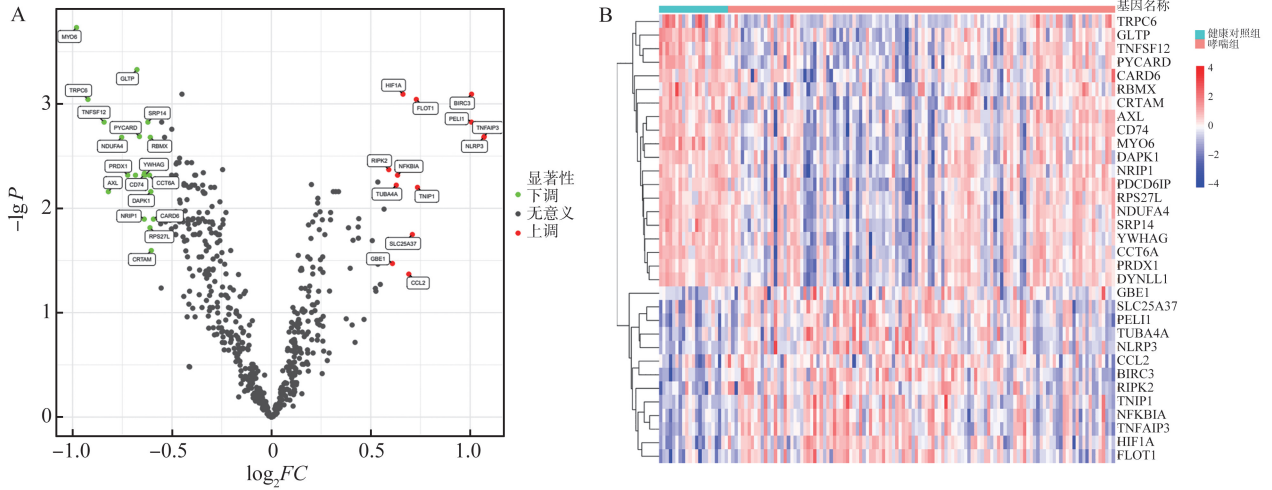


图2 哮喘组和健康对照组中差异表达的NRGs
 A: 差异表达的NRGs火山图; B: 33个差异表达NRGs的热图。
 Figure 2 Differentially expressed NRGs in asthma and healthy samples
 A: Volcano plot of differentially expressed NRGs; B: Heatmap of 33 differentially expressed NRGs.

2.2 差异表达的NRGs的PPI网络分析

33个NRGs中有20个存在相互作用(图3A)。其中,5个上调基因NFKBIA、TNFAIP3、NLRP3、

CCL2和RIPK2位于PPI网络的核心,每个基因的交互次数如图3B,PPI网络核心的基因均存在4种以上的相互作用。

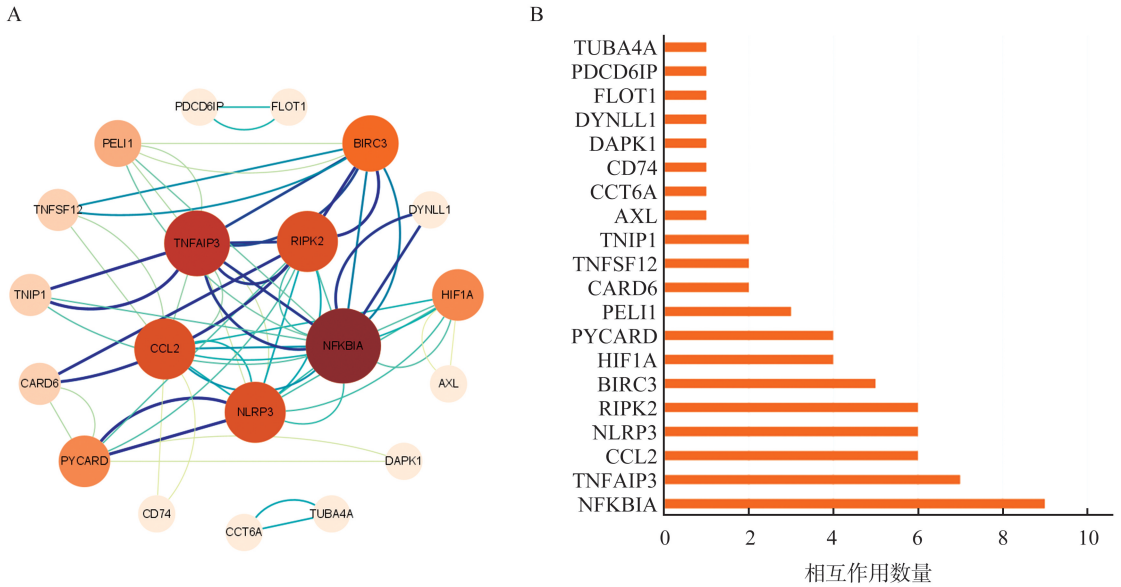


图3 33个差异表达的NRGs的PPI分析
 A: 存在PPI的差异表达NRGs; B: 每个差异表达的NRGs的相互作用数量。
 Figure 3 PPI analysis of 33 differentially expressed NRGs
 A: PPI of differentially expressed NRGs; B: The number of interactions for each differentially expressed NRGs.

2.3 差异表达的NRGs的GO和KEGG富集分析

GO富集分析的生物学过程主要涉及多个信号通路,如I-kappaB激酶/NF-kappaB信号通路、Toll样受体信号通路、核苷酸结合寡聚结构域(nucleotide binding oligomerization domain, NOD)-受体信号通路、细胞因子介导的信号通路和免疫反应调节信号通路。生物过程也对刺激作出反应,如脂多糖、其他细菌来源的分子和生物刺激。

这些差异表达的NRGs也参与了淋巴细胞的激活和调节以及细胞因子产生的调节。细胞成分主要参与形成炎性体复合物和色素颗粒。分子功能主要与凋亡过程中肽酶激活剂的活性有关(图4A)。在KEGG通路富集分析中,差异表达的NRG主要参与NOD样受体信号通路、NF-kappa B信号通路、坏死性凋亡和微生物感染(图4B)。

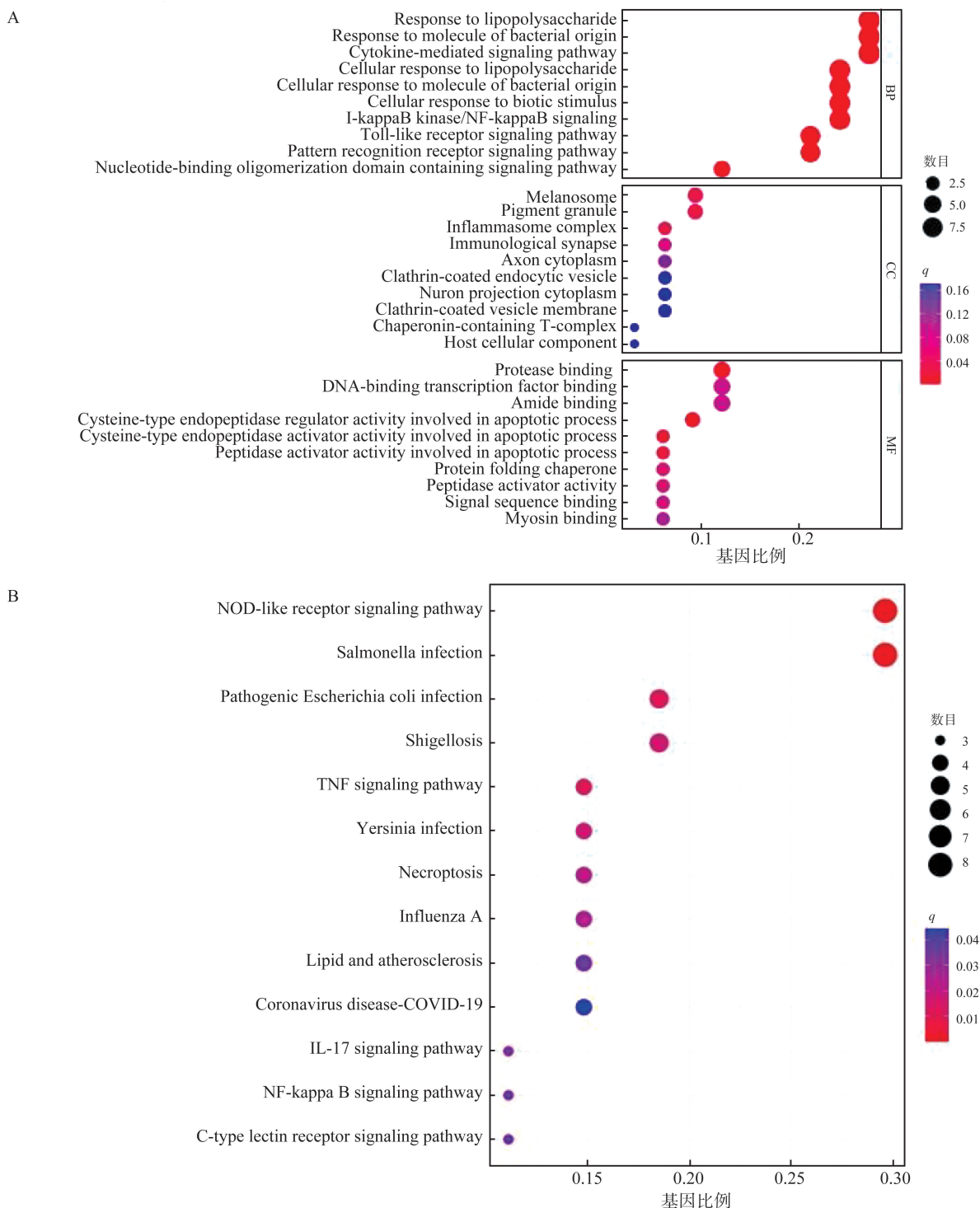


图4 33个差异表达NRG的GO(A)和KEGG(B)富集分析结果

纵坐标是功能或通路的名称,横坐标是基因的比例,圆圈的大小代表每个功能或通路中富集的基因数量,圆圈的颜色代表富集显著性。

Figure 4 GO (A) and KEGG (B) enrichment analysis results of 33 differentially expressed NRGs

The ordinate is the name of the function or pathway, the abscissa is the ratio of genes, the size of the circle represents the number of genes enriched in each function or pathway, and the circles' color represents the significance of the enrichment.

2.4 疾病诊断标志物的筛选

基于PPI网络中的核心基因和差异表达分析中的前5位基因以上调的NRGs为主,选择上调的NRGs进行机器学习以筛选哮喘标志物。通过LASSO回归算法(图5A)鉴定出7个特征基因,通

过SVM-RFE算法鉴定出13个特征基因(图5B)。将上述两种算法得到的基因进行比较并对重叠的基因进行标记。最后,获得7个具有诊断价值的疾病标志物(图5C),包括BIRC3、HIF1A、FLOT1、PELI1、NLRP3、RIPK2和GBE1。在GSE76262中绘

制 7 个标志基因的 ROC 曲线,观察到重叠基因的诊 断效率较高,AUC 均高于 0.7(图 5D)。

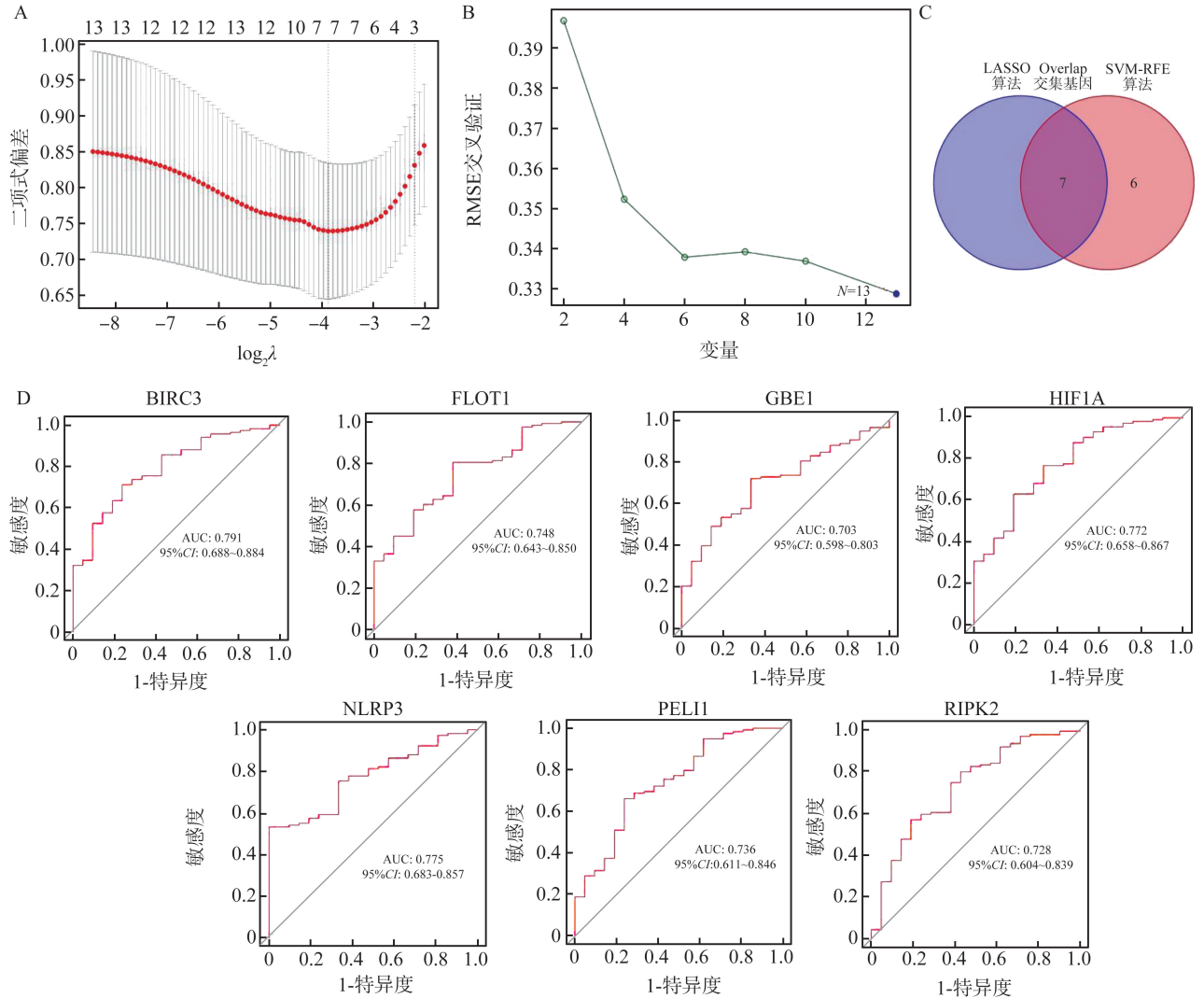


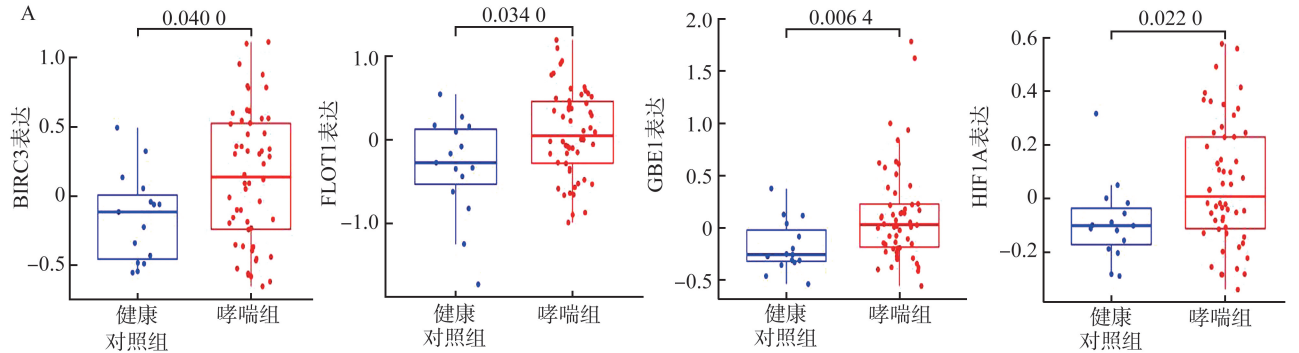
图 5 疾病诊断标记物的筛选
 A: LASSO 回归得到 7 个特征基因; B: SVM-RFE 得到 13 个特征基因; C: 维恩图显示两种算法得到的 7 个共同特征基因; D: ROC 曲线显示训练集中 7 个交集基因的诊断效率。

Figure 5 Screening of disease diagnostic markers
 A: LASSO regression yielded 7 signature genes; B: SVM-RFE yielded 13 signature genes; C: Venn diagram showed 7 common signature genes obtained by the two algorithms; D: ROC curve showed the diagnostic efficiency of 7 intersection genes in the training set.

2.5 诊断标志物的验证

以 GSE137268 作为验证集,对 7 个 NRGs 作为哮喘标志物的诊断效能进行验证。BIRC3、HIF1A、

FLOT1、NLRP3、RIPK2、GBE1 和 PELI1 均通过验证,在哮喘组中的表达水平高于对照组 ($P < 0.05$),且 ROC 曲线显示 AUC 均高于 0.6。见图 6。



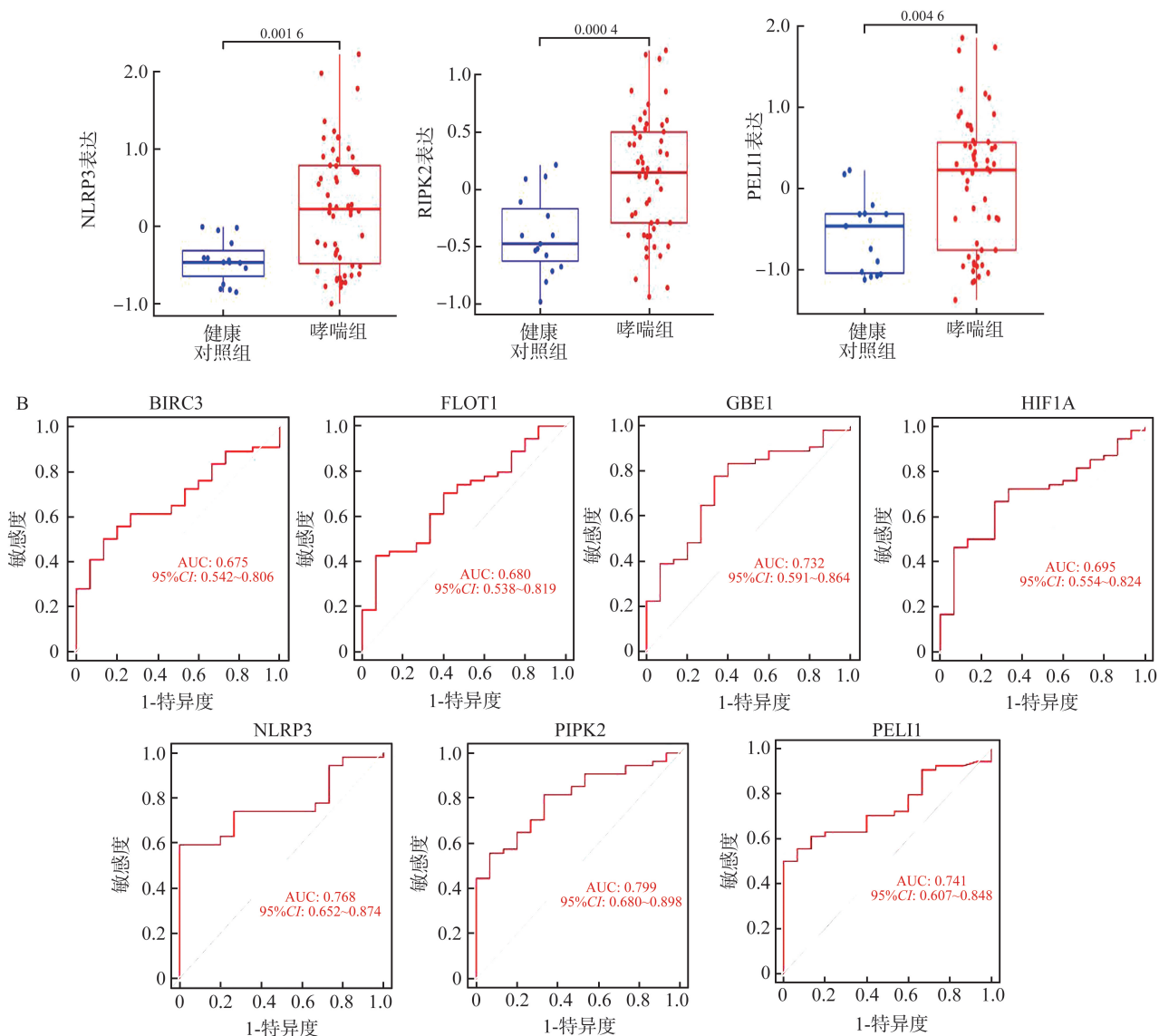


图6 疾病诊断标志物的验证

A: 验证集中7个特征基因的差异表达; B: 7个特征基因的ROC曲线。

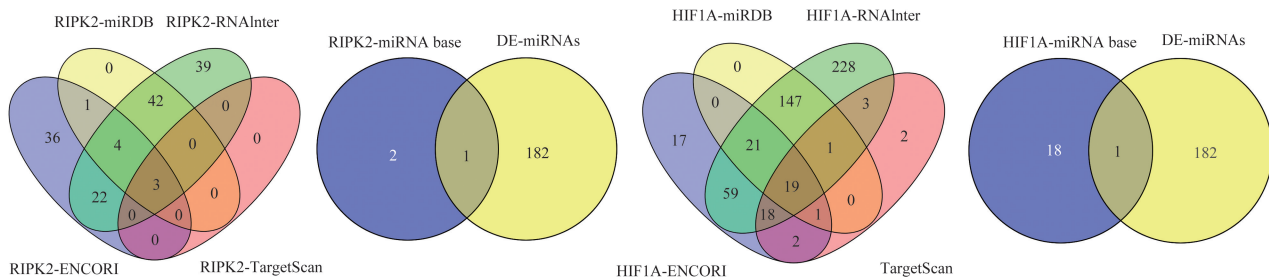
Figure 6 Validation of disease diagnostic markers

A: Differential expression of 7 signature genes in the validation set; B: ROC curves of 7 signature genes.

2.6 miRNA 和靶基因的预测

在线工具以7个经过验证的特征基因作为靶基因来预测其上游的miRNA, 将预测到的miRNA与哮喘患者中差异表达的miRNA进行比较, 最后得到3个miRNA可以靶向调控哮喘患者的NRGs的表达, 分别是 hsa-miR-138-5p、hsa-miR-

200b-3p 和 hsa-miR-30e-5p。见图7。在哮喘患者中 miRNA 与其靶基因的表达趋势相反, 即 hsa-miR-138-5p(下调)/HIF1A(上调)、hsa-miR-200b-3p(下调)/RIPK2(上调)和 hsa-miR-30e-5p(下调)/PELI1(上调)。



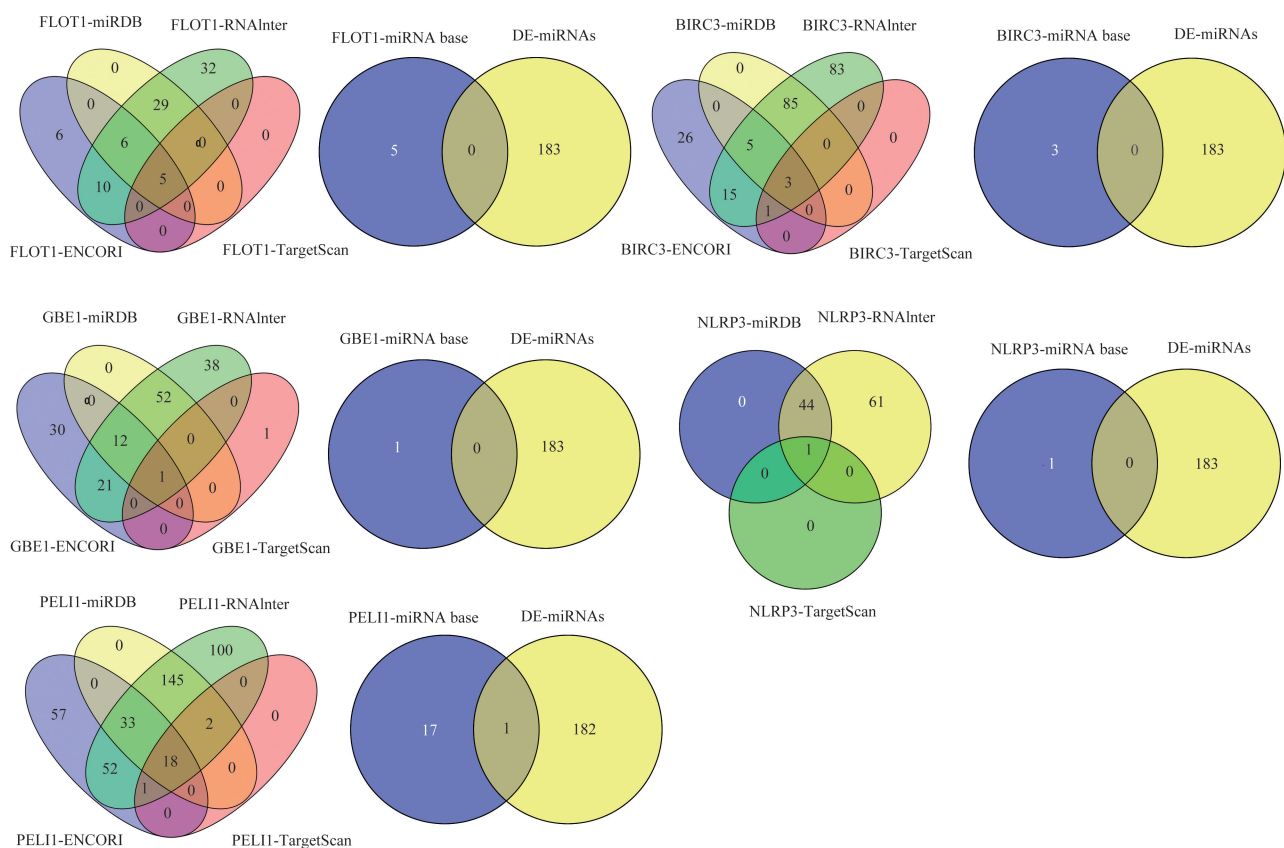


图7 预测调控靶基因表达的 miRNA

Figure 7 Predicting miRNAs regulated the expression of target genes

3 讨论

细胞凋亡最初被认为是 PCD 的唯一机制。随后,研究者通过实验分析出细胞死亡的全过程,发现了多种 PCD 通路,如坏死性凋亡、细胞焦亡和铁死亡。其中,坏死性凋亡是在细胞凋亡被抑制的条件下被激活的细胞防御通路,涉及多种细胞信号通路的激活^[6,17]。越来越多的证据表明,坏死性凋亡可能通过影响气道炎症参与哮喘的发病机制。Lu 等^[18]发现,在香烟烟雾诱导的慢性阻塞性肺病模型中,与坏死性凋亡相关的 mRNA 和蛋白质水平增加,而抑制坏死性凋亡可以减少香烟烟雾引起的气道炎症、气道重塑和肺气肿。Oikonomou 等^[19]发现,气道上皮细胞在屋尘螨刺激下,气道屏障表面发生坏死性凋亡和免疫稳态失衡,这加剧了气道炎症反应,促使哮喘恶化。然而,需要更深入的研究来了解坏死性凋亡在哮喘发病机制中的作用。本研究基于从人类基因数据库中获得 614 个 NRGs,重新分析了哮喘患者和健康对照者诱导痰的基因表达矩阵,通过研究哮喘患者 NRG 的表达来寻找哮喘的潜在发病机制和诊断标志物。

Kuo 等^[12]和 Baines 等^[13]通过微阵列分析,对哮喘患者和健康个体的诱导痰进行基因表达分析。Kuo 等通过层次聚类分析的方法,分析了差异表达基因和特定基因,并确定了 3 个与转录组相关的簇。Baines 等发现,富集于 P38 信号通路中的基因与严重哮喘和中性粒细胞气道炎症相关。本研究从坏死性凋亡的角度,对 Kuo 等完成的基因表达矩阵进行了二次分析,共鉴定出 33 个潜在的哮喘 NRGs,其中 13 个上调基因、20 个下调基因。为了进一步探索这些差异表达的 NRGs 在哮喘中的潜在作用,进行了基因功能研究,包括 PPI、GO 和 KEGG 分析。PPI 结果显示,有 20 个差异表达的 NRGs 相互作用。5 个上调基因 (NFKBIA、TNFAIP3、NLRP3、CCL2 和 RIPK2) 位于 PPI 网络的核心,与 BIRC3、HIF1A、PYCARD、CARD6、TNIP1 和 DYNLL1 等基因密切相关。

GO 富集分析结果显示,这些差异表达的 NRGs 参与了机体对脂多糖、其他细菌来源的分子和生物刺激做出的反应。它们还参与 I-kappaB 激酶/NF-kappaB 信号通路、Toll 样受体信号通路、NOD 样受体信号通路、细胞因子介导的信号通路和免疫反应调节信号通路。此外,差异表达的 NRGs 参与了以

淋巴细胞为主的细胞活化、调节细胞因子产生等生物学过程。NLRP3 和 YCARD 参与了炎症小体复合物和色素颗粒等细胞成分的形成。BIRC3 和 PYCARD 的分子功能与细胞凋亡过程中肽酶激活剂的活性有关。KEGG 通路富集分析结果显示, BIRC3、PYCARD、TNFAIP3、NLRP3、RIPK2、NFKBIA、CARD6 和 CCL2 等 8 个基因在 NOD 样受体信号通路中显著富集, 其他通路包括 NF- κ B 信号通路、坏死性凋亡、IL-17 信号通路和微生物感染。结合 GO 和 KEGG 的分析结果, 我们推测这些差异表达的基因可能通过活化和调节淋巴细胞产生细胞因子参与到 NOD 样受体、肿瘤坏死因子受体、NF-kappaB 等信号通路中, 诱导了细胞发生凋亡、坏死性凋亡和其他程序性死亡, 以对机体受到的各种刺激做出反应。Miller 等^[20]发现, 当气道上皮细胞对屋尘螨刺激作出反应时, RIPK2 被激活并在 CCL2 的产生和过敏性气道炎症中发挥重要作用。Ait Yahia 等^[21]后续研究发现, 屋尘螨诱导的实验性哮喘通过 NOD 样受体-RIPK2 通路加重过敏性气道炎症。相应地, 阻断 NOD1 受体或 RIPK2 可以减轻气道炎症和气道阻力。这些已有的研究成果为本研究的推测提供了理论依据。

LASSO 逻辑回归和 SVM-RFE 算法等机器学习方法可以鉴定潜在的哮喘生物标志物。LASSO 逻辑回归通过在分类误差的最小值处找到 λ 来确定变量^[22]。SVM-RFE 算法通过减去 SVM 生成的特征向量来找到最优变量^[23]。本研究最终将 BIRC3、HIF1A、FLOT1、NLRP3、RIPK2、GBE1 和 PELI1 确定为坏死性凋亡相关的哮喘潜在标志物, 并且进一步验证结果表明这个策略是可行的。

据报道, 这些经过验证的特征基因均与呼吸系统疾病有关。RIPK2 已被证实可以影响过敏性气道炎症的发生与发展^[20-21]。Du 等^[24]发现, 哮喘患者诱导痰中 BIRC3 显著升高, 与气道嗜酸性粒细胞、外周血过敏性炎症、2 型细胞因子、气道阻塞呈正相关。Huerta-Yepey 等^[25]发现 HIF 直接参与过敏性气道炎症的发展, 并首次证明 HIF1A 在哮喘和鼻炎患者中增加。Li 等^[26]发现 FLOT1 在哮喘气道平滑肌细胞中的表达显著增加, 而在莫西沙星治疗后显著降低。NLRP3 是一种细胞内 PRR, 可识别病原体相关分子模式、损伤相关分子模式和干扰细胞质稳态的分子。NLRP3 炎症小体对于提供有效的免疫反应至关重要。然而, 炎症小体介导的反应的异常激活会导致过度的炎症和组织损伤, 影响包括哮喘在内的慢性炎症性疾病发病机制^[27]。越来越多的

证据表明, NLRP3 炎症小体在重症哮喘、类固醇抵抗性哮喘、肥胖相关性哮喘、屋尘螨诱发的哮喘和中性粒细胞性哮喘的发病机制中起重要作用^[28-30]。GBE1 在肺部疾病中的相关研究较少, 主要是参与肿瘤微环境的免疫调节^[31-32]。Wei 等^[33]研究表明, PELI1 是 IL-1 信号通路中的信号转导子, 可作为儿童哮喘发病机制和潜在治疗靶点的生物标志物。以往的研究和本研究的生物信息学分析结果提示, 坏死性凋亡可能通过促进气道炎症参与哮喘的发生发展, 一些 NRGs 具有作为哮喘诊断标志物的潜力。

MiRNA 是一类内源性非编码小 RNA 分子, 它们通过直接降解靶 mRNA 或通过 mRNA 结合阻断蛋白质翻译从而调控基因表达, 参与许多生物过程^[34]。诸多研究表明, miRNAs 参与哮喘的发病机制。ENCORI 是一个专门针对 RNA 相互作用的综合型数据库, 提供了全面的 miRNA-mRNA 相互作用网络^[35]。TargetScan 是一个专门分析哺乳动物 miRNA 靶基因的数据库, 利用结合位点的保守性进行预测, 对于大部分保守的结合位点而言, 其准确性较好^[36]。MiRDB 是一个通过机器学习方法对预测的靶基因和 GO 数据库的综合分析来预测 miRNA 的在线数据库^[37]。RNAInter 从文献及数个 RNA 相关数据库中获取的总条目超过 4 000 万条, 是多种类型 RNA 相互作用的数据库, 可同时预测多种与 RNA 相互作用的分子^[38]。本研究通过 4 个常用的数据库预测了可能调控特征基因表达的 miRNAs, 得到了 3 对调控关系, 即 hsa-miR-138-5p/HIF1A、hsa-miR-200b-3p/RIPK2 和 hsa-miR-30e-5p/PELI1, 且 miRNA 和靶基因表达趋势相反。Tang 等^[39]研究发现, 抑制 miR-138-5p 可以缓解 OVA 诱导的过敏性鼻炎和哮喘综合征中的过敏症状并减少炎症因子。Liu 等^[40]的研究表明, 哮喘患者气道上皮细胞中 miR-200b-3p 的表达下调, 且通过 miR-200b-3p/FSTL1 轴调节哮喘气道重塑和炎症表型。目前缺乏 hsa-miR-30e-5p 在肺部疾病中的研究, 一项研究发现 hsa-miR-30e-5p 可能是自身免疫疾病中系统性红斑狼疮的生物标志物^[41]。MiRNA 与特征基因之间的靶向调控关系为哮喘和坏死性凋亡的机制和潜在的治疗靶点提供了更多信息。

本研究存在的局限性: ①数据集的样本和临床信息数量有限, 未来需要更多样本, 根据严重程度和炎症表型对哮喘进行亚组分析; ② $\log_2 FC$ 截止标准不够高, 导致哮喘组和健康对照组之间一些关键基因的表达存在细微差异; ③研究是对先前发布的数据集进行的二次挖掘和分析, 虽然结果与之前的一

些研究一致,但可靠性需要进一步的实验验证。

综上所述,本研究通过生物信息学的方法鉴定出 BIRC3、HIF1A、FLOT1、NLRP3、RIPK2、GBE1 和 PELI1 等 7 个坏死性凋亡相关基因可作为哮喘的潜在生物标志物。未来需要通过更多样本的研究验证这些基因作为诊断哮喘生物标志物的可靠性及最佳诊断阈值,还需通过基础实验探究这些基因在哮喘气道炎症的发生、发展中的作用。

参考文献:

- [1] Global Initiative for Asthma. Global strategy for asthma management and prevention [EB/OL]. (2024-01-20) [2024-05-07]. <https://ginasthma.org/2024-report/>.
- [2] Reddel HK, Bacharier LB, Bateman ED, et al. Global initiative for asthma strategy 2021: executive summary and rationale for key changes [J]. *Am J Respir Crit Care Med*, 2022, 205(1): 17-35.
- [3] Hammad H, Lambrecht BN. The basic immunology of asthma[J]. *Cell*, 2021, 184(6): 1469-1485.
- [4] Garnish SE, Tovey Crutchfield EC, Murphy JM, et al. Add necroptosis to your asthma action plan[J]. *Immunol Cell Biol*, 2021, 99(8): 800-802.
- [5] Vanden Berghe T, Linkermann A, Jouan-Lanhouet S, et al. Regulated necrosis: the expanding network of non-apoptotic cell death pathways [J]. *Nat Rev Mol Cell Biol*, 2014, 15(2): 135-147.
- [6] Pasparakis M, Vandenabeele P. Necroptosis and its role in inflammation[J]. *Nature*, 2015, 517(7534): 311-320.
- [7] Green DR. The coming decade of cell death research: five riddles[J]. *Cell*, 2019, 177(5): 1094-1107.
- [8] Seo J, Nam YW, Kim S, et al. Necroptosis molecular mechanisms: recent findings regarding novel necroptosis regulators[J]. *Exp Mol Med*, 2021, 53(6): 1007-1017.
- [9] Faust H, Mangalmurti NS. Collateral damage: necroptosis in the development of lung injury[J]. *Am J Physiol Lung Cell Mol Physiol*, 2020, 318(2): L215-L225.
- [10] Zhang T, Yin CR, Boyd DF, et al. Influenza virus Z-RNAs induce ZBP1-mediated necroptosis [J]. *Cell*, 2020, 180(6): 1115-1129.
- [11] Baines KJ, Simpson JL, Wood LG, et al. Transcriptional phenotypes of asthma defined by gene expression profiling of induced sputum samples [J]. *J Allergy Clin Immunol*, 2011, 127(1): 153-160.
- [12] Kuo CS, Pavlidis S, Loza M, et al. T-helper cell type 2 (Th2) and non-Th2 molecular phenotypes of asthma using sputum transcriptomics in U-BIOPRED [J]. *Eur Respir J*, 2017, 49(2): 1602135. doi: 10.1183/13993003.02135-2016.
- [13] Baines KJ, Fricker M, McDonald VM, et al. Sputum transcriptomics implicates increased p38 signalling activity in severe asthma [J]. *Respirology*, 2020, 25(7): 709-718.
- [14] Reiner A, Yekutieli D, Benjamini Y. Identifying differentially expressed genes using false discovery rate controlling procedures [J]. *Bioinformatics*, 2003, 19(3): 368-375.
- [15] Sun SL, Shen YH, Wang J, et al. Identification and validation of autophagy-related genes in chronic obstructive pulmonary disease [J]. *Int J Chron Obstruct Pulmon Dis*, 2021, 16: 67-78. doi:10.2147/COPD.S288428.
- [16] Chen D, Wu WL, Yi LL, et al. A potential circRNA-miRNA-mRNA regulatory network in asthmatic airway epithelial cells identified by integrated analysis of microarray datasets [J]. *Front Mol Biosci*, 2021, 8: 703307. doi:10.3389/fmolb.2021.703307.
- [17] Khan I, Yousif A, Chesnokov M, et al. A decade of cell death studies: breathing new life into necroptosis [J]. *Pharmacol Ther*, 2021, 220: 107717. doi: 10.1016/j.pharmthera.2020.107717.
- [18] Lu Z, van Eeckhoutte HP, Liu G, et al. Necroptosis signaling promotes inflammation, airway remodeling, and emphysema in chronic obstructive pulmonary disease [J]. *Am J Respir Crit Care Med*, 2021, 204(6): 667-681.
- [19] Oikonomou N, Schuijjs MJ, Chatzigiagkos A, et al. Airway epithelial cell necroptosis contributes to asthma exacerbation in a mouse model of house dust mite-induced allergic inflammation [J]. *Mucosal Immunol*, 2021, 14(5): 1160-1171.
- [20] Miller MH, Shehat MG, Alcedo KP, et al. Frontline Science: RIP2 promotes house dust mite-induced allergic airway inflammation [J]. *J Leukoc Biol*, 2018, 104(3): 447-459.
- [21] Ait Yahia S, Audoussert C, Alvarez-Simon D, et al. NOD1 sensing of house dust mite-derived microbiota promotes allergic experimental asthma [J]. *J Allergy Clin Immunol*, 2021, 148(2): 394-406.
- [22] Robert T. Regression shrinkage and selection via the lasso [J]. *J R Stat Soc Ser B Methodol*, 2018, 58(1): 267-288.
- [23] Guyon I, Weston J, Barnhill S, et al. Gene selection for cancer classification using support vector machines [J]. *Mach Learn*, 2002, 46(1): 389-422.
- [24] Du LJ, Xu CY, Zeng ZM, et al. Exploration of induced sputum BIRC3 levels and clinical implications in asthma [J]. *BMC Pulm Med*, 2022, 22(1): 86. doi: 10.1186/s12890-022-01887-2.
- [25] Huerta-Yepez S, Baay-Guzman GJ, Bebenek IG, et al. Hypoxia inducible factor promotes murine allergic air-

- way inflammation and is increased in asthma and rhinitis [J]. *Allergy*, 2011, 66(7): 909-918.
- [26] Li HT, Ye C, Zhou M, et al. Moxifloxacin suppresses airway inflammation and modulates expression of caveolin-1 and flotillin-1 in airway smooth muscle cells of asthmatic rats [J]. *Ann Transl Med*, 2019, 7(18): 469. doi:10.21037/atm.2019.08.43.
- [27] Williams EJ, Negewo NA, Baines KJ. Role of the NLRP3 inflammasome in asthma: relationship with neutrophilic inflammation, obesity, and therapeutic options [J]. *J Allergy Clin Immunol*, 2021, 147(6): 2060-2062.
- [28] Kim HY, Lee HJ, Chang YJ, et al. Interleukin-17-producing innate lymphoid cells and the NLRP3 inflammasome facilitate obesity-associated airway hyperreactivity [J]. *Nat Med*, 2014, 20(1): 54-61.
- [29] Ma M, Li GY, Qi MH, et al. Inhibition of the inflammasome activity of NLRP3 attenuates HDM-induced allergic asthma [J]. *Front Immunol*, 2021, 12: 718779. doi:10.3389/fimmu.2021.718779.
- [30] Tan HT, Hagner S, Ruchti F, et al. Tight junction, mucin, and inflammasome-related molecules are differentially expressed in eosinophilic, mixed, and neutrophilic experimental asthma in mice [J]. *Allergy*, 2019, 74(2): 294-307.
- [31] Li LF, Yang L, Cheng SQ, et al. Lung adenocarcinoma-intrinsic GBE1 signaling inhibits anti-tumor immunity [J]. *Mol Cancer*, 2019, 18(1): 108. doi:10.1186/s12943-019-1027-x.
- [32] Liang YC, Lei YY, Liang M, et al. GBE1 is an independent prognostic marker and associated with CD163⁺ tumor-associated macrophage infiltration in lung adenocarcinoma [J]. *Front Oncol*, 2021, 11: 781344. doi:10.3389/fonc.2021.781344.
- [33] Wei WP, Huang JQ, Ma Y, et al. IL-1 signaling pathway molecules as key markers in childhood asthma [J]. *Pediatr Allergy Immunol*, 2021, 32(2): 305-313.
- [34] Chen L, Heikkinen L, Wang CL, et al. Trends in the development of miRNA bioinformatics tools [J]. *Brief Bioinform*, 2019, 20(5): 1836-1852.
- [35] Li JH, Liu S, Zhou H, et al. starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data [J]. *Nucleic Acids Res*, 2014, 42(Database issue): D92-D97.
- [36] Agarwal V, Bell GW, Nam JW, et al. Predicting effective microRNA target sites in mammalian mRNAs [J]. *Elife*, 2015, 4: e05005. doi:10.7554/eLife.05005.
- [37] Chen YH, Wang XW. miRDB: an online database for prediction of functional microRNA targets [J]. *Nucleic Acids Res*, 2020, 48(D1): D127-D131.
- [38] Kang JJ, Tang Q, He J, et al. RNAInter v4.0: RNA interactome repository with redefined confidence scoring system and improved accessibility [J]. *Nucleic Acids Res*, 2022, 50(D1): D326-D332.
- [39] Tang HP, Han XL, Li TT, et al. Protective effect of miR-138-5p inhibition modified human mesenchymal stem cell on ovalbumin-induced allergic rhinitis and asthma syndrome [J]. *J Cell Mol Med*, 2021, 25(11): 5038-5049.
- [40] Liu F, Zhang J, Zhang D, et al. Follistatin-related protein 1 in asthma: miR-200b-3p interactions affect airway remodeling and inflammation phenotype [J]. *Int Immunopharmacol*, 2022, 109: 108793. doi:10.1016/j.intimp.2022.108793.
- [41] Kim BS, Jung JY, Jeon JY, et al. Circulating hsa-miR-30e-5p, hsa-miR-92a-3p, and hsa-miR-223-3p may be novel biomarkers in systemic lupus erythematosus [J]. *HLA*, 2016, 8(4): 187-193.

(编辑:相峰)