

## 慢性阻塞性肺疾病内质网应激特征基因筛选及免疫浸润表现

张 爽, 罗晨阳, 何志义

(广西医科大学第一附属医院呼吸内科, 南宁 530021)

**摘要** 目的: 基于生物信息学、机器学习算法和实验验证揭示内质网应激(ERS)在慢性阻塞性肺疾病(COPD)中的核心基因。方法: 从GEO数据库下载微阵列数据集GSE5058、GSE8545和GSE19407, 以鉴定COPD吸烟者和非吸烟者气道上皮细胞之间的差异表达基因(DEGs), 随后与ERS相关基因重叠后得到共有差异表达基因(ERs-DEGs)并进行富集分析。通过LASSO、SVM-RFE和RF 3种机器学习算法筛选ERS特征基因, 并在GSE10006中验证和评估其诊断效能, 随后进行免疫浸润分析、构建关键基因的lncRNA-miRNA-mRNA ceRNA网络和小鼠肺气肿模型肺组织mRNA表达量验证。结果: 筛选出153个共有差异基因, 其中74个基因表达上调, 79个基因表达下调。GO和KEGG分析显示, ERs-DEGs主要富集于ERS反应、蛋白折叠及多个炎症信号通路, DO分析主要富集于肺血管闭塞性疾病、COPD等。免疫浸润分析提示, COPD样本与多种免疫细胞浸润高度相关。经机器学习算法最终共鉴定出4个特征基因(包括 *THBS1*、*BCL2*、*USP13* 和 *RNFT2*), 且在训练集和验证集中均显示出良好的诊断效能。同时, 选择共表达的mRNA和miRNA构建mRNA-miRNA相互作用网络。RT-PCR结果显示, 与空气组小鼠比较, 香烟烟雾暴露诱导的肺气肿小鼠肺组织 *THBS1* 和 *RNFT2* 的mRNA表达水平升高, *BCL2* 和 *USP13* 的mRNA表达水平下降(均  $P < 0.05$ )。结论: *THBS1*、*BCL2*、*USP13* 和 *RNFT2* 可能是COPD发病过程中ERS形成的核心基因, 有望成为COPD免疫治疗的靶点。

**关键词** 慢性阻塞性肺疾病; 内质网应激; 免疫浸润; 生物信息学

中图分类号: R563.3 文献标志码: A 文章编号: 1005-930X(2024)07-1042-14

DOI: 10.16190/j.cnki.45-1211/r.2024.07.014

## Screening of endoplasmic reticulum stress characteristic genes and immune infiltration manifestations in chronic obstructive pulmonary disease

ZHANG Shuang, LUO Chenyang, HE Zhiyi. (Department of Respiratory Medicine, the First Affiliated Hospital of Guangxi Medical University, Nanning 530021, China)

**Abstract** **Objective:** To identify the core genes of endoplasmic reticulum stress (ERS) in chronic obstructive pulmonary disease (COPD) using the bioinformatics, various machine learning algorithms and experimental validation. **Methods:** The microarray data GSE5058, GSE8545, and GSE19407 were downloaded from the GEO database to identify differentially expressed genes (DEGs) between airway epithelial cells of COPD smokers and non-smokers, and then the common DEGs were obtained after overlapping with ERS-related genes and enriched for analysis. Three machine learning algorithms, LASSO, SVM-RFE, and RF, were used to screen the characteristic genes, and their diagnostic performance was verified and evaluated in the GSE10006. Subsequently, immunoinfiltration analysis was performed. Finally the lncRNA-miRNA-mRNA ceRNA network of key genes was constructed and the mRNA expression levels of lung tissue in the mouse emphysema model were verified. **Results:** A total of 153 common DEGs were screened, of which 74 genes were up-regulated and 79 genes were down-regulated. GO and KEGG analysis showed that ERs-DEGs were mainly enriched in ERS response, protein folding and multiple inflammatory signaling pathways, and DO analysis was mainly enriched in pulmonary vascular occlusive diseases and COPD and so on. Immunoinfiltration analysis showed that COPD samples were highly corre-

[基金项目] 国家自然科学基金资助项目(No.82260012)

[共同第一作者] 张爽、罗晨阳

[通信作者] 何志义, E-mail: zhiyi-river@163.com

[收稿日期] 2024-05-23

lated with a variety of immune cell infiltrations. A total of 4 characteristic genes (including *THBS1*, *BCL2*, *USP13* and *RNFT2*) were finally identified by machine learning algorithms, and they showed good diagnostic performance in both the training set and the validation set. At the same time, co-expressed mRNA and miRNA were selected to construct the mRNA-miRNA interaction network. Reverse transcription-quantitative polymerase chain reaction (RT-qPCR) results showed that compared with the air exposure group mice, the mRNA expression levels of *THBS1* and *RNFT2* in the lung tissues of emphysema mice induced by cigarette smoke exposure were increased, and the mRNA expression levels of *BCL2* and *USP13* were decreased (all  $P<0.05$ ). **Conclusion:** *THBS1*, *BCL2*, *USP13* and *RNFT2* may be the core genes formed by ERS during the pathogenesis of COPD, and are expected to be targets for COPD immunotherapy.

**Keywords** chronic obstructive pulmonary disease; endoplasmic reticulum stress; immune infiltration; bioinformatics

慢性阻塞性肺疾病(chronic obstructive pulmonary disease, COPD)是世界范围内死亡和残疾的主要原因<sup>[1]</sup>,仅次于缺血性心脏病和脑卒中<sup>[2]</sup>。吸烟是COPD的一个关键环境危险因素,与不吸烟者相比,吸烟者有更高的呼吸系统症状和肺功能异常患病率,更高的年FEV1下降率和更高的COPD死亡率<sup>[3]</sup>。迄今为止,COPD的发病机制尚不清楚。越来越多的证据表明,COPD的发病机制涉及多种生物学功能,包括慢性炎症、细胞增殖、细胞凋亡、自噬、铁下垂和内质网应激(endoplasmic reticulum stress, ERS)<sup>[4-7]</sup>。

内质网作为蛋白质合成工厂,参与钙(Ca)的储存和调节、脂质的合成和储存以及葡萄糖代谢<sup>[8]</sup>,作为一种动态的“营养感知”细胞器,能协调能量波动与代谢重编程反应,甚至对细胞命运起到关键性的作用<sup>[9]</sup>。当机体处于氧化应激、缺血、缺氧或病毒感染等状态时,内质网稳态遭到破坏,内质网中错误折叠蛋白异常积累,从而产生ERS<sup>[9-10]</sup>。而吸烟不仅会增加内质网中错误折叠蛋白的负荷,同时也会损害内质网的折叠能力<sup>[11]</sup>;过度暴露于香烟烟雾所引起的ERS状态下还可能诱导气道炎症<sup>[12]</sup>、肺泡上皮和内皮细胞衰老和死亡<sup>[13]</sup>等,导致肺功能损害和其他间接性损伤。因此,探索慢性香烟烟雾暴露(CS)

下ERS形成的核心基因,有助于对COPD的早期诊断和干预,有效延缓COPD的疾病进展。

机器学习这一数据分析技术,近年来在生物学上的应用无处不在,它旨在从多维数据集中通过数据规则和统计假设,寻找数据间的关联,进而建立预测模型,从新的视角阐述疾病的发病机制<sup>[14]</sup>。目前机器学习已运用于构建特异性COPD相关预测模型<sup>[15]</sup>,然而尚未用于阐述COPD和ERS的关系。因此,本文通过分析来自GEO数据库的3个COPD气道上皮细胞微阵列数据集,运用机器学习方法筛选出COPD-ERS相关枢纽基因,为未来COPD的治疗提供新的诊断标志物和潜在治疗靶标。

1 材料和方法

1.1 数据收集及整理

从GEO数据库(<https://www.ncbi.nlm.nih.gov/geo/>)下载基因表达谱数据集GSE5058、GSE8545、GSE19407和GSE10006,格式为MINiML,见表1。纳入GSE5058、GSE8545和GSE19407的气道上皮样本,筛选疾病的差异表达基因(differential expression genes, DEGs)。同时,利用GSE10006对核心差异基因进行验证,探讨吸烟对ERS差异表达的影响。本研究的流程图见图1。

表1 COPD GEO数据集具体数据

芯片名称	平台	样品	不吸烟者	COPD	样本类型
GSE5058	GPL570	Airway eithelial	12	15	Total RNA
GSE8545	GPL570	Airway eithelial	18	18	Total RNA
GSE19407	GPL570	Airway eithelial	47	22	Total RNA
GSE10006	GPL963	Airway eithelial	22	14	Total RNA

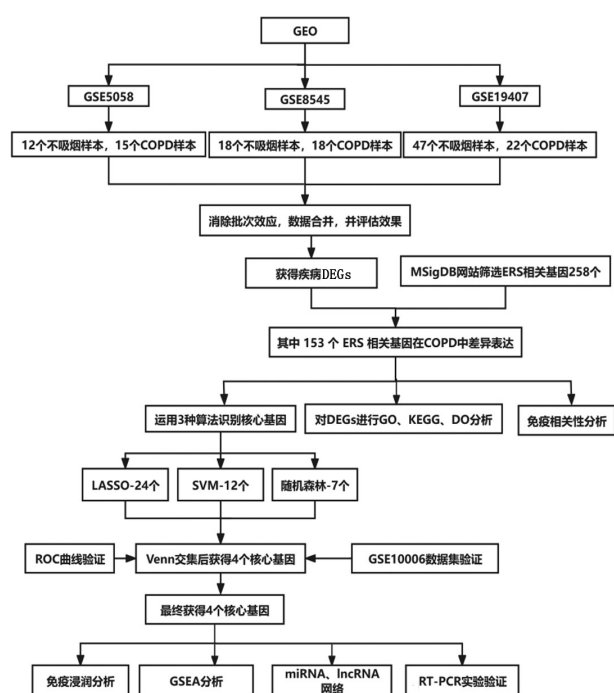


图1 本研究的流程图

利用R软件(V4.2.1)的limma包分别对数据集进行差异基因筛选,判定标准为调整校正后的 $P$ -value  $< 0.05$ ,  $\log_2|\text{Fold-Change (FC)}| \geq 1$ ;  $\log_2|\text{FC}| > 0$  为上调的基因,  $\log_2|\text{FC}| < 0$  为下调的基因。应用limma包与sva包对GSE5058、GSE8545、GSE19407数据集进行合并,并进行批次矫正;ggplot2包进行箱线图及PCA主成分分析,用于评估是否去除了批量效应,最后运用pheatmap包,绘制热图。

从分子特征数据库(Msig DB v2023.1.Hs Mar 2023)中下载ERS相关基因集。利用limma包提取ERS基因,取交集后获得共有差异表达基因(ERS-DEGs);corrplot包分析差异基因之间相关性。

## 1.2 共有差异基因富集分析

使用DOSE包和ClusterProfiler包对共有差异基因进行GO功能注释、KEGG信号通路富集分析和DO富集分析。GO功能注释包括生物学过程(biological process, BP)、分子功能(molecular function, MF)和细胞组成(cellular component, CC)等3个方面。KEGG是一个包含基因组、化学和系统功能信息的综合数据库,其中KEGG PATHWAY数据库是目前生物信息学分析领域应用最广泛的代谢通路数据库。DO富集分析从疾病的角度对基因进行描述。

## 1.3 核心基因的筛选与验证

采用随机森林(random forest, RF)、最小绝对收

缩和选择算子逻辑回归(LASSO)、支持向量机递归特征消除(SVM-REF)3种机器学习算法筛选核心基因,在R语言上应用RF、glmnet和e1071包进行;将筛选的结果取交集。为确定重叠基因的敏感性和特异性,在训练组和验证组数据集中分别运用pROC包绘制受试者工作特征(ROC)曲线,并计算曲线下面积(area under the curve, AUC),确定核心基因在COPD和不吸烟者样本中的诊断效能。

## 1.4 关键基因GSEA分析

从GSEA平台(<http://www.gsea-msigdb.org/>)下载用于GO和KEGG富集分析gmt文件,选择“c2.cp.kegg.symbols.gmt.”作为参考基因集。利用limma包根据COPD患者和正常受试者的分组信息进行方差分析,绘制出按 $t$ 值排序的差异表达基因集上调和下调直方图。

## 1.5 免疫细胞浸润情况的评估

采用CIBERSORT算法评估22个免疫细胞亚群在不吸烟组和COPD疾病组样本中的浸润情况,采用pheatmap包和corrplot包绘制为堆积柱状图,可视化免疫细胞的相对含量,分析其相关性;采用ggplot2包绘制豆荚图以显示不吸烟组和COPD疾病组样本中免疫细胞浸润情况的差异。

## 1.6 lncRNA-miRNA-mRNA ceRNA调控网络的构建

lncRNA通过与miRNA的竞争性结合来调控mRNA的表达。为了解关键基因的靶向ceRNAs,使用了miRanda、Targetscan和miRDB这3个miRNA数据库预测,3个数据库均包含,则认为这个miRNA与这个基因有相关性。根据spongescan数据库(<https://spongescan.rc.ufl.edu>),获得与上述收集的miRNA相互作用的lncRNA。结合miRNA-mRNA和lncRNA-miRNA相互作用,得到lncRNA-miRNA-mRNA ceRNA调控网络,使用Cytoscape 3.9.1软件将结果可视化。

## 1.7 肺气肿小鼠模型的建立

16只雄性C57BL/6J小鼠(6~8周龄,体重20~25 g)购自广西医科大学实验动物中心。将小鼠随机分为2组,空气组( $n=8$ )和CS组( $n=8$ ),分别暴露于室内空气、CS中6个月。CS诱导的肺气肿小鼠模型参照既往研究方法<sup>[16-17]</sup>,即CS组每次暴露于5支香烟(南宁真龙未过滤香烟:10 mg焦油和0.9 mg尼古丁),每次30 min,包括30 min无烟间隔(4次/d, 5 d/周,6个月)。动物实验按照广西医科大学动物



研究伦理委员会批准的方案进行。

将小鼠肺组织用10%福尔马林固定,包埋并切片(4 μm),每只小鼠选择3个不同区域的切片(不包括大气道和血管)。肺组织切片用二甲苯脱蜡,用乙醇水化,然后用苏木精—伊红(HE)染色。使用荧光病理显微镜(Olympus\*/BX53+DP80,日本)选取5个不同的视野拍摄图像,使用Image-Pro Plus软件6.0(Media Cybernetics)测量每组的平均线性截距(MLI),作为肺泡直径的估计。

1.8 RT-PCR 实验验证

依据说明书,用TRIzol试剂分别从CS组和空气组小鼠肺组织中提取总RNA,使用HiScript@III RT SuperMix for qPCR(Vazmye,南京,中国)制备cDNA。Real Time-PCR采用2x Q3 SYBR qPCR Master mix(TOLOBIO,上海,中国),按说明书的指示进行,采用10 μL体系进行RT-PCR扩增。引物序列来自美国国家生物信息中心(NCBI)Primer-BLAST,引物由生工生物工程(上海)有限公司合成引物,见表2。按照制造商的说明,使用 $2^{-\Delta\Delta CT}$ 法计算ERs-DEGs的相对mRNA表达水平。

表2 RT-PCR 引物

基因名称	定位	引物序列(5'~3')
Gapdh	上游	AGAAGGTGGTGAAGCAGGCATC
	下游	CGAAGGTGGAAGAGTGGGAGTTG
Bcl2	上游	CCGTCGTGACTTCGCAGAGATG
	下游	ATCCCTGAAGAGTTCCTCCACCAC
Usp13	上游	TCCCAGAGTTCCAGAGAGCGTATG
	下游	TGTGTGTTGAAGTCCTGCGTTGG
Rnft2	上游	GGGCGGACTCAGGAAAGCATTG
	下游	CTCGGAAGTCAGCTTGGCAGATG
Thbs1	上游	CACTGCCAGAACTCGGTTACCATC
	下游	TCGCATCTGTTGTTGAGGCTGTC

1.9 统计学方法

采用R软件(4.2.1版本)和GraphPad Prism 8.0对数据进行统计分析。计量资料用均数±标准差( $\bar{x} \pm s$ )表示,组间比较采用配对t检验。以 $P < 0.05$ 为差异有统计学意义。

2 结 果

2.1 COPD 差异基因筛选结果

根据筛选标准,GSE5058、GSE8545和GSE19407分别得到3 136个、1 581个和320个差异基因。差异基因火山图见图2A~图2C,前50个差异基因热图见图2D~图2F。对数据集的表达矩阵进行归一化处理后,箱形图的分布趋势基本为直线(图3A)。为了评估组内数据的可重复性,本研究对3个数据集进行了PCA分析,结果显示:合并数据集的不吸烟组与COPD组之间差异显著,归一化后显示数据具有良好的可重复性(图3B)。对差异表达数据进行合并及批次校正后绘制热图(图3C)。

2.2 COPD 小气道上皮细胞中ERS基因筛选

利用Msig DB数据库下载ERS相关基因,共获得258个,以调整后的 $FDR < 0.05$ 且 $|\log_2 FC| \geq 1$ 为筛选标准,鉴定ERS相关基因在合并后的数据集中差异性表达。最终得到153个ERs-DEGs,其中74个基因表达上调,79个基因表达下调。ERs-DEGs的相对表达情况在热图中展示(图3D),红色越深,代表基因的相对表达量越高,蓝色越深则相对表达量越低。ERs-DEGs之间的相关性在热图中展示(图3E)。

2.3 DEGs的GO、EGG、DO富集分析

对上述得到的153个ERs-DEGs进行富集分析。GO功能注释主要参与ERS、蛋白折叠、泛素样蛋白酶等过程(图4A)。KEGG信号通路富集分析涉及内质网中的蛋白质加工、神经退行性病变、脂质和动脉粥样硬化、帕金森病、P53信号通路、PI3K-Akt信号通路等(图4B)。DO分析主要富集于神经系统肿瘤、肺血管闭塞性疾病、COPD、帕金森病及胰腺癌等(图4C)。

2.4 核心基因的筛选与验证

运用3种机器学习算法对ERs-DEGs进行识别,其中LASSO算法得出24个关键基因(图5A~图5B);SVM-REF算法得出12个关键基因(图5C~图5D);RF算法得出7个关键基因(图5E~图5F)。Venn图最终得到4个重叠的核心基因(图5G),其中THBS1、RNFT2上调( $\log_2 |FC| \geq 1$ ),BCL2、USP13下调( $\log_2 |FC| < -1$ )。



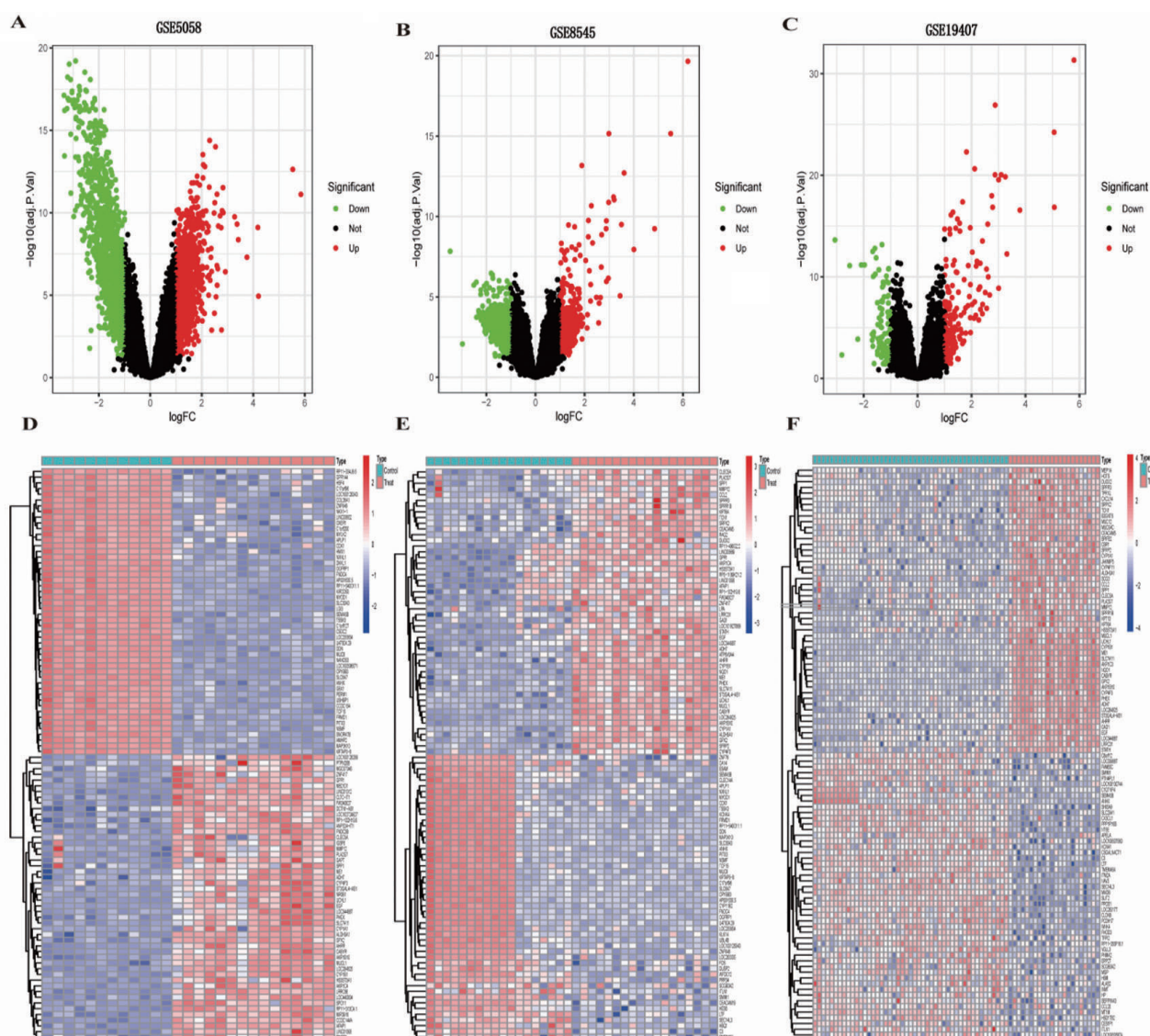


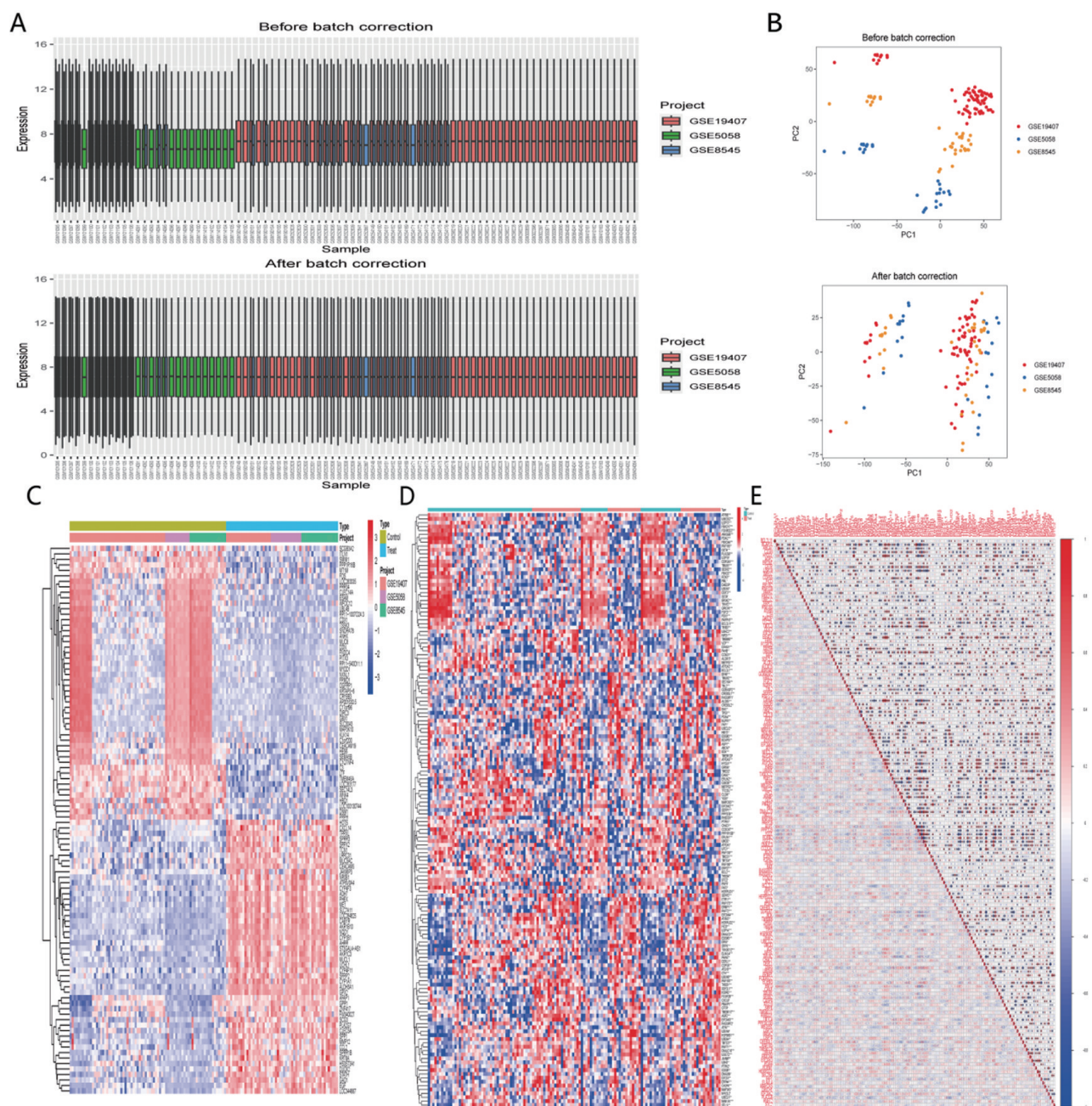
图2 COPD 差异基因火山图(A~C)及热图(D~F)

ROC 曲线用于评估基因对疾病的诊断效能,计算得出AUC值,一般介于0.5~1.0之间,较大的AUC值代表了该基因对该疾病具有较好的诊断性能。在数据集中将数据分为不吸烟组和COPD组,然后对获得的4个疾病特征基因绘制ROC曲线(图5H),4个核心基因AUC值均大于0.8:*BCL2*(AUC=0.868)、*THBS1*(AUC=0.877)、*USP13*(AUC=0.902)、*RNFT2*(AUC=0.904);随后在验证数据集GSE10006中对这4个核心基因进行了验证(图5I),4个基因AUC值均大于0.8,说明*THBS1*、*RNFT2*、*BCL2*、*USP13*作为COPD的特征基因,对疾病的预测具有较高的准

确性。且在验证组COPD样本中*THBS1*、*RNFT2*表达上调( $P<0.001$ )、*BCL2*、*USP13*表达下调( $P<0.001$ )(图5J~图5M),与训练组结果一致。因此,将上述显著差异表达的4个基因作为COPD中ERS相关的关键基因。

## 2.5 关键基因GSEA分析

为进一步研究核心基因的分子生物学功能,进行了GSEA分析,结果显示:这些核心基因与细胞因子受体相互作用、趋化因子信号通路、jak/stat信号通路、自身免疫性甲状腺病、病毒性心肌炎等相关(图6)。



A:去批次前、后合并数据集的箱形图;B:去批次前、后合并数据集的PCA图;C:数据集批次校正后表达热图;D~E:COPD ERs-DEGs热图和共差异基因之间相关性热图。

图3 合并数据集矫正前、后的箱形图、PCA图和表达热图

## 2.6 免疫浸润分析及相关性

免疫环境在 COPD 发生和发展中起着关键作用,采用 CIBERSORT 算法对免疫细胞浸润情况进行分析,22种免疫细胞亚群分布的相对比例见图7A。免疫细胞差异的豆荚图(图7B)显示,COPD 样本 M0 型巨噬细胞、静息的树突状细胞、嗜酸性粒细胞

等呈高表达,而调节性 T 细胞、M2 型巨噬细胞等呈低表达( $P<0.001$ )。22种免疫细胞的相关热图(图7C)显示,活化的肥大细胞和中性粒细胞强正相关关系( $r=0.73$ )、活跃的 CD4 记忆细胞与 M1 型巨噬细胞呈强正相关关系( $r=0.61$ ),而静息的肥大细胞和中性粒细胞呈负相关关系( $r=0.5$ )。



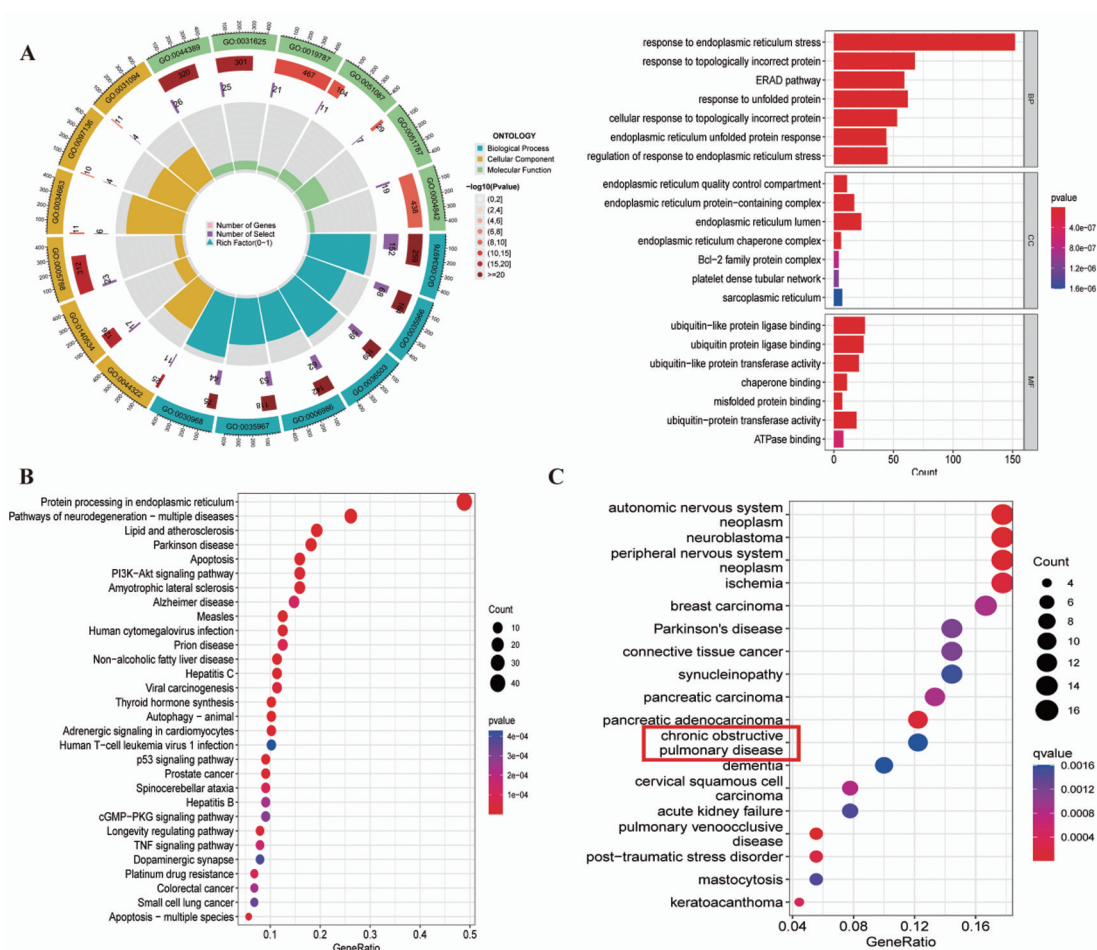


图4 DEGs的GO(A)、KEGG(B)及DO(C)富集分析

对核心基因进行GSEA分析,提示与免疫相关性疾病有关,进一步对关键基因与免疫细胞进行相关性分析(图7D),结果显示:*USP13*的表达水平与幼稚B细胞( $P<0.05$ )、嗜酸性粒细胞( $P<0.05$ )、M2型巨噬细胞( $P<0.05$ )呈正相关关系,与浆细胞呈负相关关系( $P<0.05$ );*THBS1*表达水平与单核细胞( $P<0.05$ )、中性粒细胞( $P<0.01$ )呈正相关关系,与静息的CD4记忆细胞( $P<0.05$ )、滤泡性辅助T细胞( $P<0.05$ )呈负相关关系;*RNFT2*表达水平与静息的树突状细胞( $P<0.001$ )、静息的肥大细胞( $P<0.01$ )、调节性T细胞( $P<0.05$ )呈正相关关系,与M1型巨噬细胞( $P<0.05$ )、活跃的肥大细胞( $P<0.01$ )呈负相关关系;*BCL2*与静息的CD4记忆细胞( $P<0.01$ )呈正相关关系。

## 2.7 lncRNA-miRNA-mRNA ceRNA 调控网络的构建

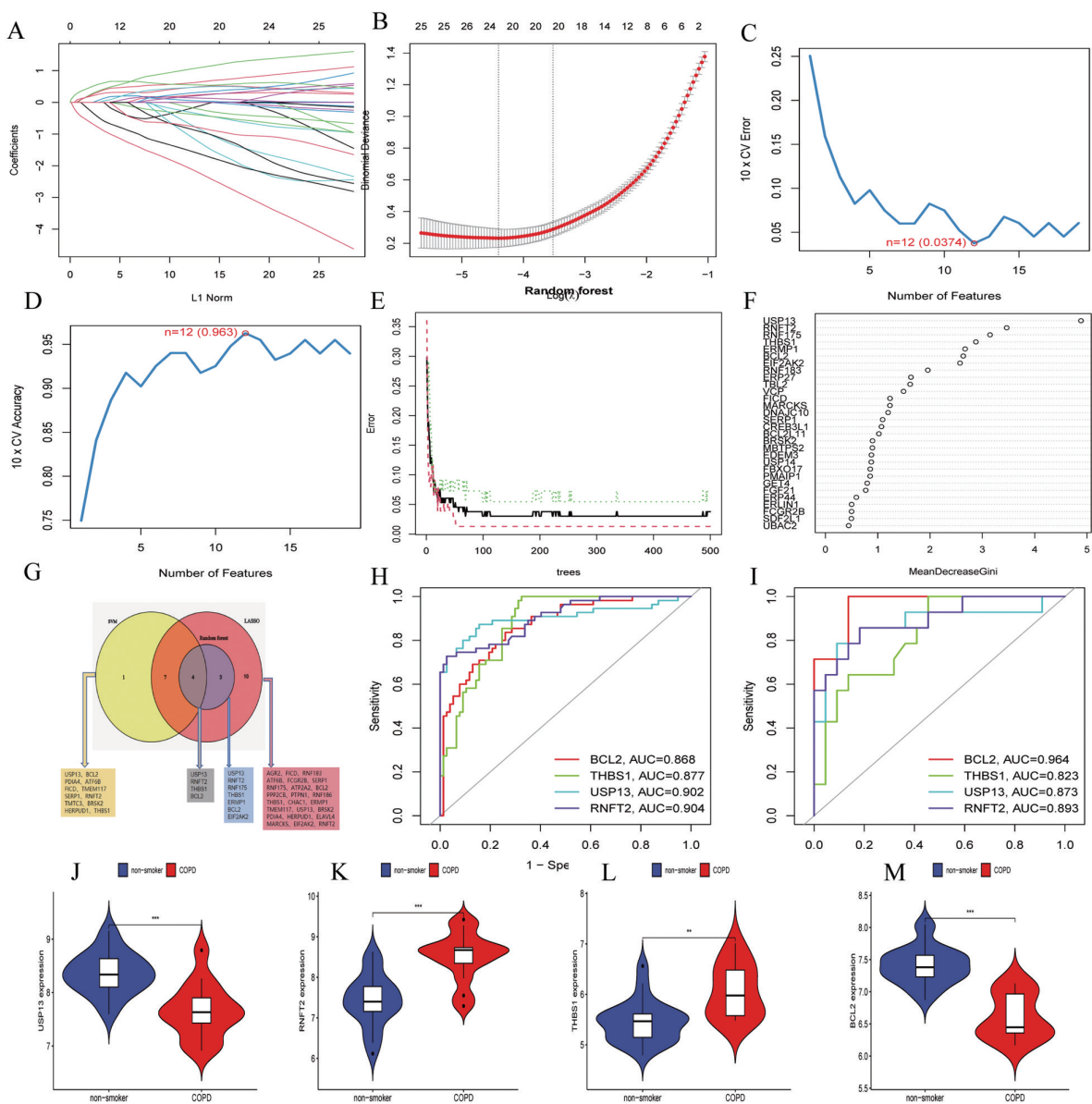
为进一步揭示调控机制,探索miRNA对COPD中关键基因的潜在调控作用,建立了lncRNA-miRNA-mRNA调控网络(图8),该复杂网络包括261个

miRNA、236个lncRNA和4个关键基因。*USP13*、*BCL2*和*THBS1*连接的miRNA较多,分别有83个、77个和76个,而*RNFT2*有25个,其中hsa-miR-3163、hsa-miR-335-3p、hsa-miR-548c-3p、hsa-miR-182-5p和hsa-miR-590-3p与多个关键基因结合。

## 2.8 在肺气肿小鼠模型中的差异验证

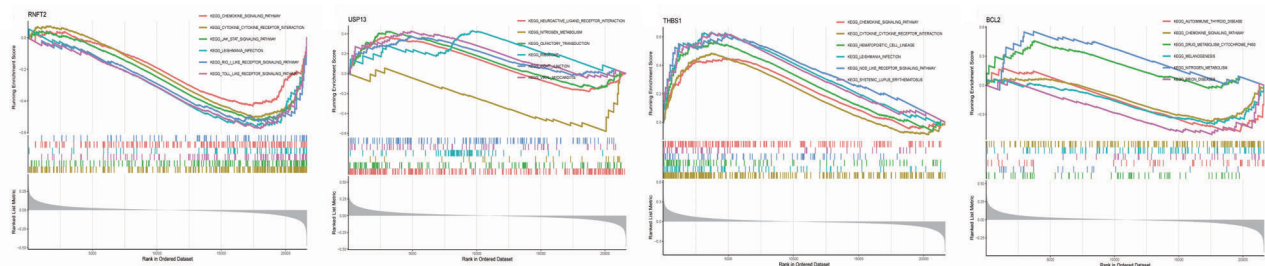
显微镜下观察CS组( $n=8$ )和空气组( $n=8$ )小鼠肺组织HE染色病理切片,空气组小鼠肺泡结构基本完整,炎症细胞浸润较少,肺泡间隔未见明显断裂,肺泡MLI为( $26.5\pm4.178$ )  $\mu\text{m}$ (图9A)。CS组小鼠表现出严重的肺气肿,肺泡间隔断裂融合,肺泡壁变薄,肺泡MLI为( $61.96\pm13.758$ )  $\mu\text{m}$ (图9B)。与空气组小鼠比较,CS组小鼠MLI显著增大( $P<0.01$ )(图9C)。RT-PCR实验验证结果提示,*THBS1*和*RNFT2*的mRNA表达水平在CS组小鼠肺组织中高于空气组,*BCL2*和*USP13*的mRNA表达水平在CS组小鼠肺组织中低于空气组(均 $P<0.05$ ),见图9D~9G。





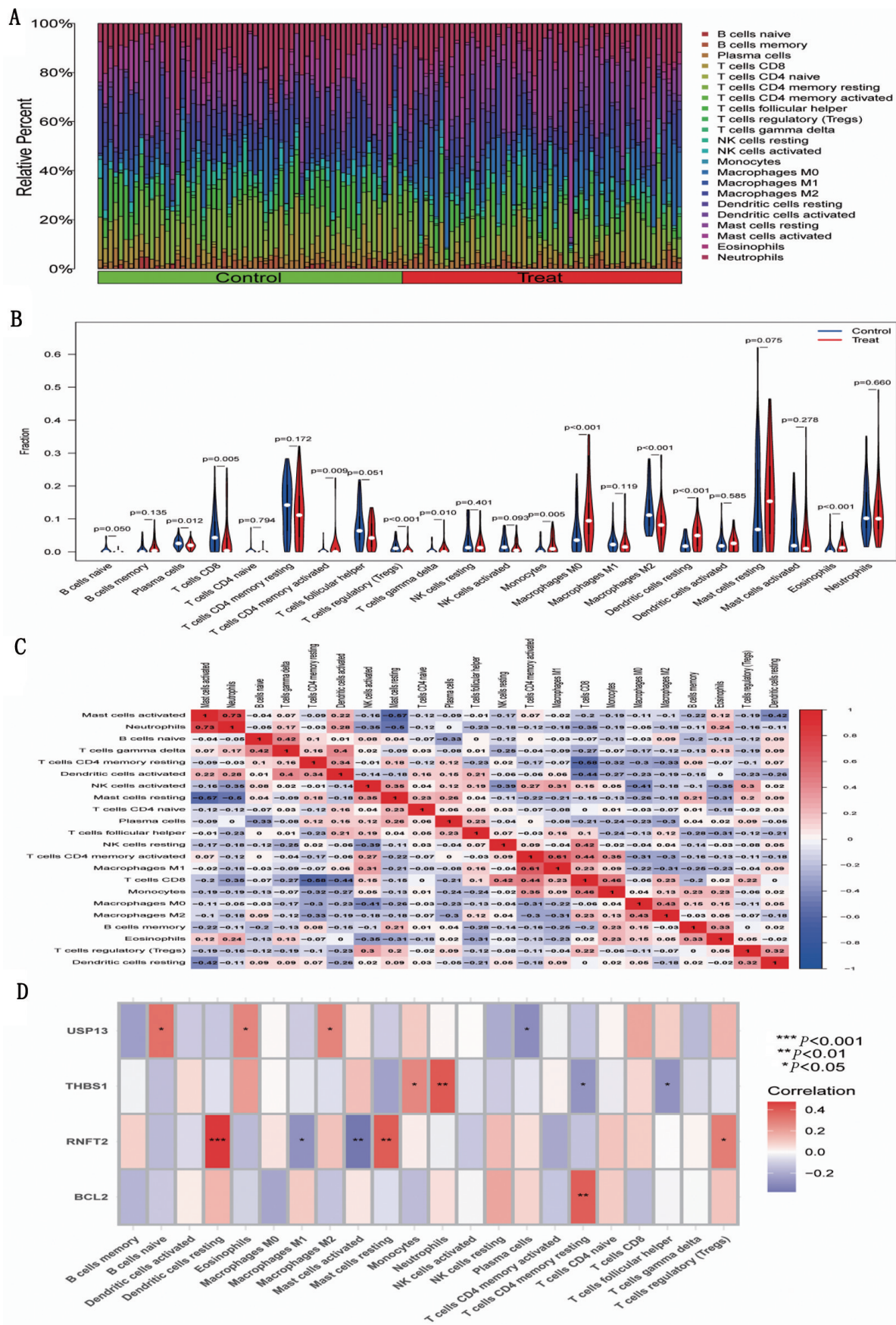
A~B: LASSO 回归分析; C~D: SVM-RFE 分析; E~F: RF 分析; G: Venn 图; H: 训练数据集 (GSE5058、GSE8545、GSE19407) ROC 曲线; I: 验证数据集 (GSE10006) ROC 曲线; J~M: 4 个关键基因在验证数据集中差异表达的小提琴图。

图5 核心 ERS 基因的鉴定及诊断效能评估



图中上部曲线图代表每个基因在通路富集中的得分情况,下部分代表每个基因的 RANK 值表达情况,左边代表高表达,右边代表低表达。

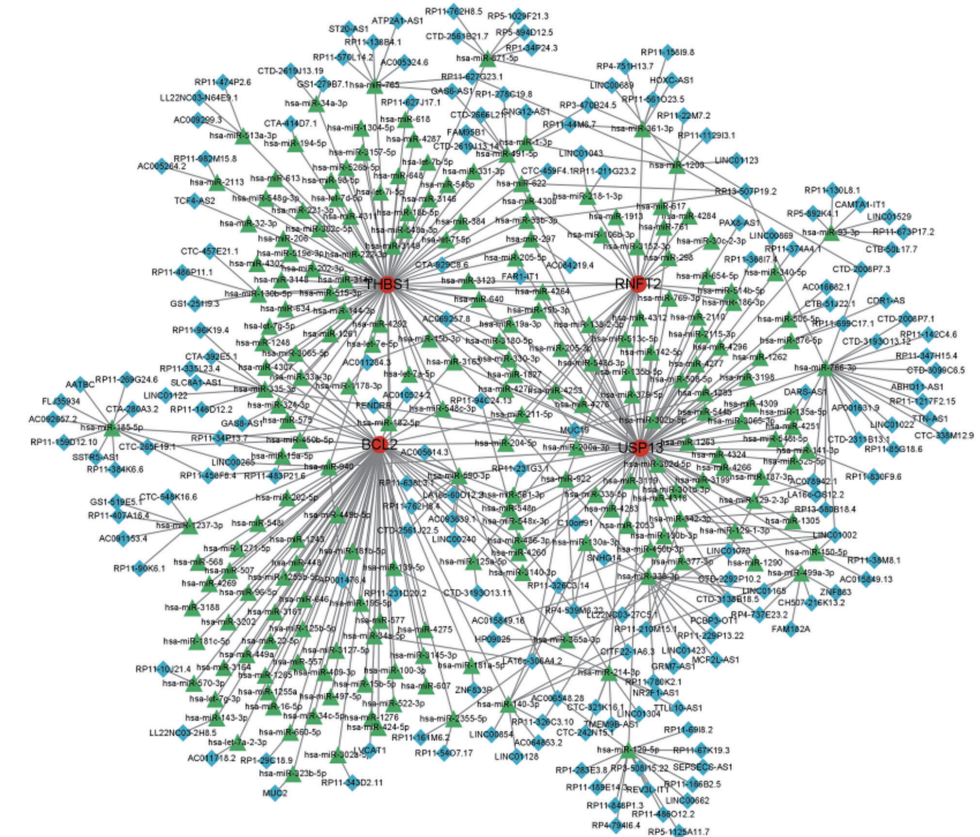
图6 COPD 关键基因 GSEA 分析



A:免疫细胞亚群分布图;B:免疫细胞差异豆荚图;C:免疫细胞相关热图;D:关键基因与免疫细胞的相关性分析。

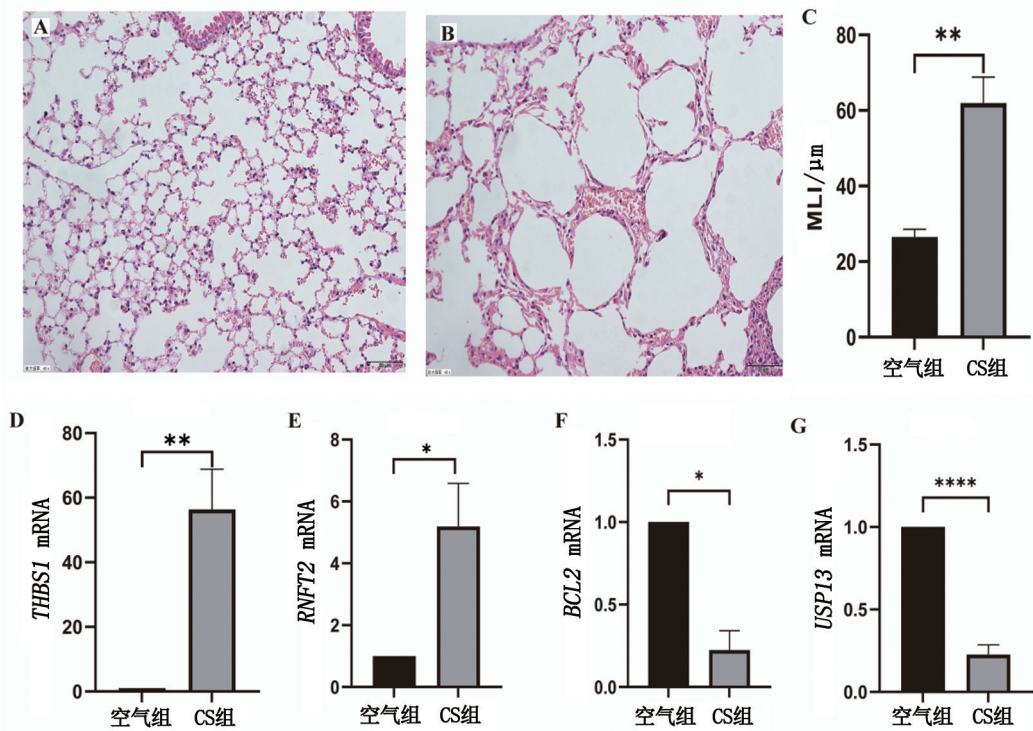
图7 免疫浸润分析





圆形:ERS 核心基因;三角形:miRNA;菱形:lncRNA。

图8 lncRNA-miRNA-mRNA 调控网络



A:空气组小鼠HE染色切片( $\times 200$ );B:CS组小鼠的HE染色切片( $\times 200$ );C:肺泡MLI比较;D~G:关键基因的mRNA相对表达水平;组间比较,\* $P < 0.05$ ,\*\* $P < 0.01$ ,\*\*\*\* $P < 0.0001$ 。

图9 关键基因的表达验证



### 3 讨论

COPD是以持续性、进行性发展且不可逆的气流受限为特征的慢性肺部炎症性疾病,主要病理改变为肺气肿、气道和肺实质慢性炎症以及全身炎症反应。其发病和进展的主要原因是暴露因素,如生物燃料、香烟烟雾和空气污染。尽管吸烟人口比例在1990—2015年期间有所下降(男性下降28%,女性下降29%),但吸烟仍是全球COPD的最主要原因<sup>[18]</sup>。内质网是细胞蛋白质折叠的主要部位,但当不正确折叠的蛋白质过度积累时,内质网稳态破坏会触发ERS<sup>[19]</sup>,如果应激持续太久或太严重,会对细胞造成不可逆转的损害,甚至诱发细胞死亡<sup>[20]</sup>。由香烟烟雾中存在的活性氧(ROS)和活性氮产生的一些活性中间体(如丙烯醛),它们的细胞毒性导致肺部的氧化损伤和蛋白质的错误折叠<sup>[6]</sup>。既往研究也证实了慢性香烟烟雾暴露和香烟烟雾提取物可引发ERS,引发肺上皮细胞凋亡<sup>[21]</sup>。4-PBA是最常用的ERS抑制剂,可下调ERS减少CS诱导的细胞凋亡,从而保护肺泡结构<sup>[22]</sup>。白藜芦醇是一种天然植物多酚,具有抗炎、抗氧化的生物学作用,可减轻ERS所致COPD大鼠肺上皮细胞凋亡,进而减轻肺组织损伤和病理改变<sup>[23]</sup>。麻黄素可通过阻断ERS抑制细胞死亡、炎症反应和氧化应激,从而减轻CS所诱导的COPD<sup>[24]</sup>。以上研究均提示抑制ERS可能成为治疗COPD的有效手段,因此深入挖掘和探索ERS形成和释放的核心基因具有重要意义。

本研究通过分析GEO数据库中3个COPD气道上皮细胞微阵列数据集(GSE5058、GSE8545和GSE19407),筛选DEGs作为疾病靶基因,同时在靶基因中进一步筛选ERs-DEGs相关基因,共获得153个ERs-DEGs,并对ERs-DEGs进行了GO、KEGG及DO富集分析。进一步通过3种机器学习算法最后筛选出4个关键ERS基因,分别为*THBS1*、*RNFT2*、*BCL2*和*USP13*,在训练组及验证组数据集中分别建立ROC曲线,它们的AUC值均高于0.8;在肺气肿小鼠肺组织中*RNFT2*和*THBS1*表达上调,*USP13*和*BCL2*表达下调,证实了它们可能是COPD发病机制中ERS形成的核心基因。

吸烟引起的气道炎症也与COPD患者的免疫浸

润有关,对疾病靶基因进行免疫浸润分析提示,与健康小气道上皮细胞样本相比,COPD样本含有较多的M0型巨噬细胞、静息的树突状细胞、嗜酸性粒细胞等,而调节性T细胞、M2型巨噬细胞等则较少。对关键基因进行免疫细胞相关性分析,其中*THBS1*表达水平与中性粒细胞呈强正相关关系,*RNFT2*表达水平与静息的树突状细胞、静息的肥大细胞呈强正相关关系,与活跃的肥大细胞呈强负相关关系;*BCL2*与静息的CD4记忆细胞呈强正相关关系。最后将ceRNA网络构建与基于免疫浸润的免疫细胞分析相结合,有助于识别COPD的诊断标志物,了解调控机制,从而为COPD的临床管理和治疗提供依据及更多靶点。

血小板反应蛋白(THBS)是脊椎动物中一个由分泌基质细胞糖蛋白组成的家族<sup>[25]</sup>,能够结合许多蛋白质并作为细胞外基质(ECM)的相互作用平台,通过编码分泌的基质细胞Ca<sup>2+</sup>结合糖蛋白参与各种生物过程<sup>[26]</sup>。*THBS1*是家族成员之一。*THBS*蛋白被发现与ERS反应有关,在成人组织损伤或活跃地发育重塑期间,或当内质网Ca水平在慢性疾病期间改变时,其与内质网驻地效应物(如BiP和ATF6a)相互作用,导致适应性应激反应、囊泡扩张和蛋白质聚集减少,同时增加蛋白质分泌,对机体起到保护作用<sup>[27]</sup>。对于*THBS1*目前的研究主要集中在抑制肿瘤生长方面,它已被证明是体内新生血管形成和肿瘤发生的有效抑制剂<sup>[28]</sup>,是第一个公认的天然存在的血管生成抑制剂<sup>[29]</sup>。关于*THBS1*在呼吸系统方面的研究,见于家族性肺动脉高压,其在肺血管稳态中发挥重要的作用<sup>[30]</sup>;另有研究显示,*THBS1*可抑制肺癌细胞运动、迁移和侵袭<sup>[31]</sup>。但是*THBS1*和ERS的功能调控关系在COPD中尚未见相关报道,有待于进一步研究。

*BCL2*是Bcl-2家族的抗凋亡成员,定位于线粒体外膜,通过诱导线粒体孔形成,阻止线粒体释放细胞色素C,从而抑制细胞凋亡<sup>[32]</sup>。研究表明,*BCL2*不仅在淋巴样细胞中表达,也在呼吸上皮内化生黏液细胞等上皮组织中表达,并可能通过持续存在黏液细胞和过量黏液分泌,影响慢性气道疾病(如COPD、哮喘)患者肺功能,参与COPD的发生和进展<sup>[33-34]</sup>。DNA甲基转移酶1(DNMT1)引起*BCL2*甲基化状态改变与CS诱导的ROS或氧化应激升高

所致肺细胞凋亡有关<sup>[35]</sup>。本研究通过生物信息学分析和实验验证发现BCL2在气道上皮和慢性香烟烟雾暴露肺气肿小鼠肺组织中表达下降,提示BCL2可能参与COPD的发病机制,但具体致病机制及其与ERS之间的调控关系需要继续深入探讨。

RNFT2是无名指家族的一员,与细胞质和细胞核的蛋白水解有关,在胃肠道、肝脏和肺等人体多个器官中普遍表达<sup>[36]</sup>。目前对RNFT2的研究较少,只有少数报道表明其与胃癌<sup>[36]</sup>预后及抑制乳腺癌<sup>[37]</sup>的转移有关。另外,据报道,RNFT2作为E3连接酶,通过对IL-3炎症细胞因子受体IL3Ra的泛素化和降解来减少肺部炎症和损伤<sup>[38]</sup>。

USP13是半胱氨酸蛋白酶DUB家族的一员,它从泛素结合的蛋白底物中切割泛素分子,在细胞内的蛋白质静止网络中发挥核心作用。研究证明,USP13丢失会损害自噬活性,并通过Beclin 1去泛素化促进肺纤维化<sup>[39]</sup>。USP13的缺乏对细胞外基质生成和炎症反应的影响,可能导致CS暴露条件下小鼠肺泡空间扩大<sup>[40]</sup>。此外,USP13的缺失增加了IL-1R8的多泛素化,降低了其稳定性,从而增强了脂多糖(LPS)诱导的TLR4信号传导和细胞因子的释放,促进了肺损伤后的肺部炎症<sup>[41]</sup>。值得注意的是,上述4个基因在癌症生物学中都有报道,除了BCL2外,另外3个基因之前没有发现与COPD相关,因此需要进一步研究它们在该疾病中的作用。

综上所述,本研究通过数据库筛选COPD-ERS相关的DEGs,对其中存在显著性差异且差异趋势一致的关键基因进行功能富集、免疫浸润分析及构建ceRNA网络,并在肺气肿小鼠肺组织中进行RT-PCR表达验证,发现4个核心基因mRNA的表达与预测趋势一致,其中THBS1、RNFT2表达上调,BCL2、USP13表达下调,证实了THBS1、RNFT2、BCL2和USP13可能通过调控ERS参与COPD的发病过程。但本研究仍存在局限性:在不同的细胞和组织中相同的基因表达存在差异,而本研究的数据集标本来源为气道上皮细胞,在其他不同的组织或细胞中本文所阐述的关键基因是否还发挥关键调控作用,还需对肺组织、外周血单个核细胞等样本进行生物学实验以进一步验证。此外,4个关键基因在吸烟导致的COPD发病机制与ERS之间的关系网络还需要后续研究进一步完善。

## 参考文献:

- [1] CHRISTENSON S A, SMITH B M, BAFADHEL M, et al. Chronic obstructive pulmonary disease [J]. Lancet, 2022, 399(10342): 2227-2242.
- [2] ORGANIZATION W H. Geneva: World Health Organization [Z]. Global Health Estimates, 2022.
- [3] KOHANSAL R, MARTINEZ-CAMBLOR P, AGUSTÍ A, et al. The natural history of chronic airflow obstruction revisited: an analysis of the Framingham offspring cohort [J]. American journal of respiratory and critical care medicine, 2009, 180(1): 3-10.
- [4] SHAPIRO S D, INGENITO E P. The pathogenesis of chronic obstructive pulmonary disease: advances in the past 100 years [J]. American journal of respiratory cell and molecular biology, 2005, 32(5): 367-372.
- [5] YOSHIDA M, MINAGAWA S, ARAYA J, et al. Involvement of cigarette smoke-induced epithelial cell ferroptosis in COPD pathogenesis [J]. Nature communications, 2019, 10(1): 3145.
- [6] YU Y, YANG A, YU G, et al. Endoplasmic reticulum stress in chronic obstructive pulmonary disease: Mechanisms and future perspectives [J]. Biomolecules, 2022, 12(11): 1637.
- [7] ZHENG M, HONG W, GAO M, et al. Long noncoding RNA COPDA1 promotes airway smooth muscle cell proliferation in chronic obstructive pulmonary disease[J]. American journal of respiratory cell and molecular biology, 2019, 61(5): 584-596.
- [8] ALBERTS B. Molecular biology of the cell[M]. 4th edn. Garland sciences, 2022.
- [9] ALMANZA A, CARLESSO A, CHINTHA C, et al. Endoplasmic reticulum stress signalling - from basic mechanisms to clinical applications[J]. The FEBS journal, 2019, 286(2): 241-278.
- [10] ZHAO H, LIU H, YANG Y, et al. Hydrogen sulfide plays an important role by regulating endoplasmic reticulum stress in diabetes-related diseases [J]. International journal of molecular sciences, 2022, 23(13): 7170.
- [11] KELSEN S G. The unfolded protein response in chronic obstructive pulmonary disease [J]. Annals of the American thoracic society, 2016, 13(Suppl 2): 138-145.
- [12] WANG Y, SU N X, PAN S G, et al. Fengbaisan suppresses endoplasmic reticulum stress by up-regulating SIRT1

- expression to protect rats with chronic obstructive pulmonary diseases [J]. *Pharmaceutical biology*, 2020, 58(1): 878-885.
- [13] LIN L, YIN Y, HOU G, et al. Ursolic acid attenuates cigarette smoke-induced emphysema in rats by regulating PERK and Nrf2 pathways [J]. *Pulmonary pharmacology & therapeutics*, 2017, 44: 111-121.
- [14] CAMACHO D M, COLLINS K M, POWERS R K, et al. Next-generation machine learning for biological networks [J]. *Cell*, 2018, 173(7): 1581-1592.
- [15] ZHANG Y, XIA R, LV M, et al. Machine-learning algorithm-based prediction of diagnostic gene biomarkers related to immune infiltration in patients with chronic obstructive pulmonary disease [J]. *Frontiers in immunology*, 2022, 13: 740513.
- [16] ZHANG H, QIU S L, TANG Q Y, et al. Erythromycin suppresses neutrophil extracellular traps in smoking-related chronic pulmonary inflammation [J]. *Cell death & disease*, 2019, 10(9): 678.
- [17] XIAOFEI Y, TINGTING L, XUAN W, et al. Erythromycin attenuates oxidative stress-induced cellular senescence via the PI3K-mTOR signaling pathway in chronic obstructive pulmonary disease [J]. *Frontiers in pharmacology*, 2022, 13: 1043474.
- [18] RABE K F, WATZ H. Chronic obstructive pulmonary disease [J]. *Lancet*, 2017, 389(10082): 1931-1940.
- [19] SCHRÖDER M, KAUFMAN R J. ER stress and the unfolded protein response [J]. *Mutation research*, 2005, 569(1-2): 29-63.
- [20] OZCAN L, TABAS I. Role of endoplasmic reticulum stress in metabolic disease and other disorders [J]. *Annual review of medicine*, 2012, 63: 317-328.
- [21] GAN G, HU R, DAI A, et al. The role of endoplasmic reticulum stress in emphysema results from cigarette smoke exposure [J]. *Cellular physiology and biochemistry*, 2011, 28(4): 725-732.
- [22] WANG Y, WU Z Z, WANG W. Inhibition of endoplasmic reticulum stress alleviates cigarette smoke-induced airway inflammation and emphysema [J]. *Oncotarget*, 2017, 8(44): 77685-77695.
- [23] 李玉屏, 罗百灵, 张立, 等. 白藜芦醇减轻慢性阻塞性肺疾病大鼠内质网应激诱导的肺泡上皮细胞凋亡研究 [J]. *中华结核和呼吸杂志*, 2014, 37(1): 30-35.
- [24] WANG H L, CHEN F Q, WU L J. Ephedrine ameliorates chronic obstructive pulmonary disease (COPD) through restraining endoplasmic reticulum (ER) stress in vitro and in vivo [J]. *International immunopharmacology*, 2022, 103: 107842.
- [25] SCHIPS T G, VANHOUTTE D, VO A, et al. Thrombospondin-3 augments injury-induced cardiomyopathy by intracellular integrin inhibition and sarcolemmal instability [J]. *Nature communications*, 2019, 10(1): 76.
- [26] STENINA O I, TOPOL E J, PLOW E F. Thrombospondins, their polymorphisms, and cardiovascular disease [J]. *Arteriosclerosis, thrombosis, and vascular biology*, 2007, 27(9): 1886-1894.
- [27] LYNCH J M, MAILLET M, VANHOUTTE D, et al. A thrombospondin-dependent pathway for a protective ER stress response [J]. *Cell*, 2012, 149(6): 1257-1268.
- [28] VOLPERT O V, PILI R, SIKDER H A, et al. Id1 regulates angiogenesis through transcriptional repression of thrombospondin-1 [J]. *Cancer cell*, 2002, 2(6): 473-483.
- [29] GOOD D J, POLVERINI P J, RASTINEJAD F, et al. A tumor suppressor-dependent inhibitor of angiogenesis is immunologically and functionally indistinguishable from a fragment of thrombospondin [J]. *Proceedings of the national academy of sciences of the united states of America*, 1990, 87(17): 6624-6628.
- [30] MALONEY J P, STEARMAN R S, BULL T M, et al. Loss-of-function thrombospondin-1 mutations in familial pulmonary hypertension [J]. *American journal of physiology lung cellular and molecular physiology*, 2012, 302(6): L541-554.
- [31] HUANG W T, CHONG I W, CHEN H L, et al. Pigment epithelium-derived factor inhibits lung cancer migration and invasion by upregulating exosomal thrombospondin 1 [J]. *Cancer letters*, 2019, 442: 287-298.
- [32] TAO S C, YUAN T, RUI B Y, et al. Exosomes derived from human platelet-rich plasma prevent apoptosis induced by glucocorticoid-associated endoplasmic reticulum stress in rat osteonecrosis of the femoral head via the Akt/Bad/Bcl-2 signal pathway [J]. *Theranostics*, 2017, 7(3): 733-750.
- [33] SATA M, TAKABATAKE N, INOUE S, et al. Intronic single-nucleotide polymorphisms in Bcl-2 are associated with chronic obstructive pulmonary disease severity [J]. *Respirology*, 2007, 12(1): 34-41.
- [34] HARRIS J F, FISCHER M J, HOTCHKISS J R, et al. Bcl-2 sustains increased mucous and epithelial cell numbers in metaplastic airway epithelium [J]. *American journal of*



- respiratory and critical care medicine, 2005, 171(7): 764-772.
- [35] ZENG H, LI T, HE X, et al. Oxidative stress mediates the apoptosis and epigenetic modification of the Bcl-2 promoter via DNMT1 in a cigarette smoke-induced emphysema model [J]. Respiratory research, 2020, 21(1): 229.
- [36] SASAHARA M, KANDA M, SHIMIZU D, et al. Tissue RNFT2 expression levels are associated with peritoneal recurrence and poor prognosis in gastric cancer [J]. Anti-cancer research, 2021, 41(2): 609-617.
- [37] TAN P, YE Y, HE L, et al. TRIM59 promotes breast cancer motility by suppressing p62-selective autophagic degradation of PDCD10 [J]. PLoS biology, 2018, 16(11): e3000051.
- [38] TONG Y, LEAR T B, EVANKOVICH J, et al. The RNFT2/IL-3R $\alpha$  axis regulates IL-3 signaling and innate immunity [J]. JCI insight, 2020, 5(3): e133652.
- [39] LIU Y, LI Z, XIAO H, et al. USP13 Deficiency impairs autophagy and facilitates age-related lung fibrosis [J]. American journal of respiratory cell and molecular biology, 2023, 68(1): 49-61.
- [40] GREGORY A D, TRAN K C, TAMASKAR A S, et al. USP13 Deficiency aggravates cigarette-smoke-induced alveolar space enlargement [J]. Cell biochemistry and biophysics, 2021, 79(3): 485-491.
- [41] LI L, WEI J, LI S, et al. The deubiquitinase USP13 stabilizes the anti-inflammatory receptor IL-1R8/Sigirr to suppress lung inflammation [J]. EBioMedicine, 2019, 45: 553-562.

本文引用格式:

张 爽,罗晨阳,何志义.慢性阻塞性肺疾病内质网应激特征基因筛选及免疫浸润表现[J]. 广西医科大学学报, 2024, 41(7): 1042-1055. DOI: 10.16190/j.cnki.45-1211/r.2024.07.014

ZHANG S, LUO C Y, HE Z Y. Screening of endoplasmic reticulum stress characteristic genes and immune infiltration manifestations in chronic obstructive pulmonary disease[J]. Journal of Guangxi medical university, 2024, 41(7): 1042-1055. DOI: 10.16190/j.cnki.45-1211/r.2024.07.014