

• 数据挖掘与循证医学 •

非特异性间质性肺炎关键基因与免疫细胞浸润分析及中药预测

田仁伟，张少谦，太鹤蓓，张雅婷，陈明竹，罗超，吴文丽，吴宁*

贵州医科大学基础医学院 化学与生物化学实验室，贵州 贵阳 550004

摘要：目的 探寻非特异性间质性肺炎相关的关键基因、发病机制、免疫细胞浸润水平和潜在治疗中药。方法 从 GEO 数据库获取基因芯片 GSE110147，利用 R 语言“limma”等相关拓展包筛选差异表达基因，通过 STRING 和 Cytoscape 软件构建蛋白质互作网络图，并筛选出关键基因。关键基因通过基因芯片 GSE101286 进行验证后，利用 R 语言“clusterProfiler”等拓展包及 Cytoscape 软件中的 GlueGO 插件进基因本体（gene ontology, GO）功能富集分析、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，再基于 CIBERSORT 反卷积算法分析免疫细胞浸润模式。最后将关键基因导入 Coremine Medical 数据库中进行相关中药预测。结果 共筛选出 407 个差异基因和 15 个关键基因。GO 分析显示，非特异性间质性肺炎与剪切体、翻译过程高度相关；KEGG 富集结果显示非特异性间质性肺炎与 RNA 剪接体通路、内质网蛋白质加工通路联系最为密切。免疫细胞浸润分析显示非特异性间质性肺炎患者组织内静息的 CD4⁺记忆 T 细胞、嗜酸性粒细胞等浸润水平较高，而 CD8⁺T 细胞、浆细胞、活化 NK 细胞和 M1 巨噬细胞浸润水平较低。通过关键基因筛选到 18 味中药，其中雷公藤和茯苓有较大治疗潜力。结论 通过生物信息学方法，筛选非特异性间质性肺炎的关键基因，明确了潜在治疗中药，为非特异性间质性肺炎发病机制、治疗新靶点研究及新药研发提供新的方向。

关键词：非特异性间质性肺炎；生物信息学；关键基因；免疫浸润；雷公藤；茯苓

中图分类号：R285 文献标志码：A 文章编号：0253-2670(2023)15-4934-14

DOI: 10.7501/j.issn.0253-2670.2023.15.019

Analysis of key genes and immune cell infiltration of nonspecific interstitial pneumonia and prediction of traditional Chinese medicine

TIAN Ren-wei, ZHANG Shao-qian, TAI He-bei, ZHANG Ya-ting, CHEN Ming-zhu, LUO Chao, WU Wen-li, WU Ning

Laboratory of Chemistry and Biochemistry, Guizhou Medical University, Guiyang 550004, China

Abstract: **Objective** To explore the key genes, pathogenesis, immune cell infiltration level and potential treatment of non-specific interstitial pneumonia. **Methods** The gene chip GSE110147 was obtained from GEO database, and the differentially expressed genes were screened by R language "limma" and other related expansion packages. The protein interaction network map was constructed by STRING and Cytoscape software, and the key genes were screened. After the key genes were verified by gene chip GSE101286, GO function enrichment and KEGG pathway analysis were carried out by using expansion packages such as R language "clusterProfiler" and GlueGO plug-ins in Cytoscape software, and the immune cell infiltration pattern was analyzed based on CIBERSORT deconvolution algorithm. Finally, the key genes were introduced into Coremine Medical Database to predict the related traditional Chinese medicine. **Results** A total of 407 differential genes and 15 key genes such as HSP90AA1, EIF5B and RPS13 were screened. GO analysis showed that nonspecific interstitial pneumonia was highly correlated with shearing body and translation process, while KEGG enrichment showed that nonspecific interstitial pneumonia was most closely related to RNA splice body pathway and endoplasmic reticulum protein processing pathway. Immune cell infiltration analysis showed that the infiltration levels of resting CD4⁺ memory T cells and eosinophils in patients with nonspecific interstitial pneumonia were higher, while the infiltration levels of CD8⁺T cells, plasma cells, activated NK cells and M1 macrophages were lower. Through the screening of key genes, 18 kinds of traditional Chinese medicine were screened, among which Leigongteng (*Tripterygium wilfordii* Hook. f.) and Fuling [*Poria cocos* (Schw.) Wolf]

收稿日期：2023-04-11

基金项目：大学生创新创业训练项目（S202210660105）

作者简介：田仁伟，男，本科在读，主要从事肺部疾病的分子机制研究。E-mail: 1181930691@qq.com

*通信作者：吴宁，女，博士，教授，硕士研究生导师，主要从事贵州特色药用植物防病抗病机制的研究、肺部相关疾病的药物治疗。

E-mail: wuning@gmc.edu.cn

have great therapeutic potential. **Conclusion** This study screened the key genes of non-specific interstitial pneumonia through bioinformatics methods, identified the potential treatment of traditional Chinese medicine, and provided a new direction for the pathogenesis of non-specific interstitial pneumonia, new treatment targets and new drug research and development.

Key words: nonspecific interstitial pneumonia; bioinformatics; key genes; immune infiltration; *Tripterygium wilfordii* Hook. f.; *Poria cocos* (Schw.) Wolf

非特异性间质性肺炎 (nonspecific interstitial pneumonia, NSIP) 属于间质性肺炎的一种, 于 2002 年被单独列出作为一种疾病^[1]。NSIP 临床特征无明显特异性, 多以咳嗽和呼吸困难为主, 少数病例会出现发热、疲劳和体质量下降等症状^[2-3], 故容易造成误诊。目前对于 NSIP 的发病机制并不明确, 但已有研究表明炎症诱导及部分免疫细胞浸润在其发展过程中起到重要作用^[4], 与肺部胶原纤维沉着致肺间质增厚一起作为 NSIP 2 个已知的主要病理改变。故挖掘其关键基因、生物学过程、相关通路和免疫细胞浸润情况有较大意义。目前对于 NSIP 的治疗方法多以激素和免疫抑制剂为主, 但其副作用较大, 治疗效果仍不理想^[5], 故不良反应较少的中药或许会成为其新的治疗思路。本研究主要探讨与 NSIP 主要病理变化的发生机制相关的基因或通路,

分析其免疫浸润模式中突出表现的细胞, 以及针对关键基因预测靶向中药, 以期为 NSIP 的临床诊断和治疗提供理论基础。

1 资料与方法

1.1 数据获取

在 GEO 数据库^[6] (GEO DataSet, <http://www.ncbi.nlm.nih.gov/gds/>) 中以“nonspecific interstitial pneumonia”、“non-specific interstitial pneumonia”、“NSIP”作为检索词, 将种属设置为“human sapiens”进行检索, 选择结果中 2018 年 1 月以后发表的基因芯片 GSE110147 为研究对象, 其中包括 10 个 NSIP 患者肺组织样本和 11 个正常组织样本。GSE101286 为 2017 年发表芯片, 且质量较好, 用于验证关键基因的差异表达, 研究技术路线见图 1。

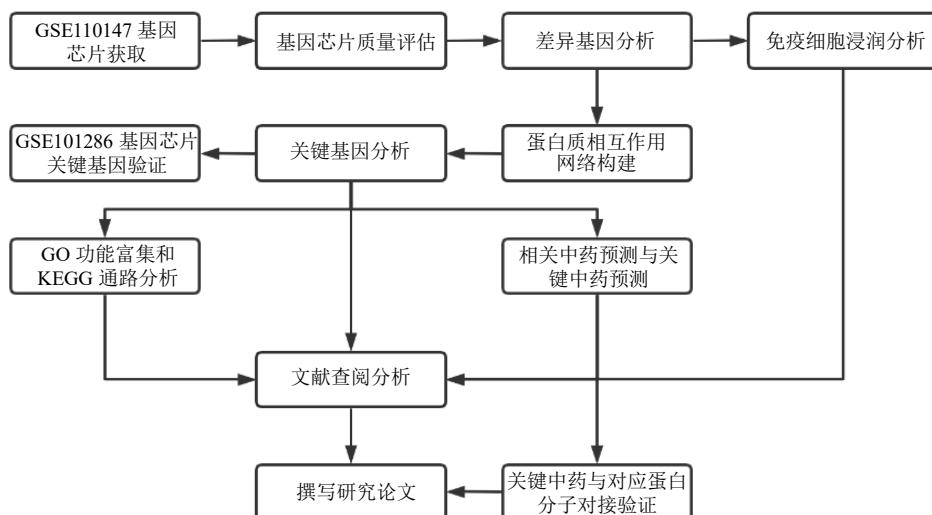


图 1 技术路线流程

Fig. 1 Technical route flow

1.2 数据预处理

运用 R Studio 调用“Bioconductor”“affyPLM”等相关拓展包对 GSE110147 表达谱芯片进行质量评估, 绘制芯片相对对数表达 (relative log expression, RLE) 箱线图、相对标准差 (normalized unscaled standard errors, NUSE) 箱线图。

1.3 数据处理及差异基因筛选

基于“limma”拓展包对 NSIP 和正常组织基因数据中的探针进行归一化处理, 并且对数据进行 τ 检验和贝叶斯检验, 当 1 个基因对应多个探针时取其平均值, 箱线图检验其标准化和分组间聚类情况。对标准化后的 GSE110147 芯片表达谱进行差异分

析。差异表达基因 (differentially expressed genes, DEGs) 以 $\log_2(\text{fold change}, \text{FC})$ 的绝对值 >2 和 $P < 0.01$ 为筛选条件, 以 $\log_2\text{FC}$ 的正负代表该基因上调或是下调。

1.4 差异基因蛋白质相互作用 (protein-protein interaction, PPI) 网络构建及关键基因筛选

将获得的 DEGs 以 medium confidence >0.4 为筛选条件, 利用 STRING 数据库^[7] (<https://string-db.org/>) 进行 PPI 分析, 再将得到的 PPI 互作网络导入 Cytoscape 软件 (Version 3.9.1), 利用 CytoHubba 插件中的度 (degree) 值对节点进行排名, 筛选获得关键基因 (hub gene)。

1.5 关键基因验证

将基因芯片 GSE101286 按照“1.2”和“1.3”项下方法进行处理得到各基因的表达情况, 再从中将分析获得的关键基因在基因芯片 GSE101286 表达矩阵中进行提取, 并通过 R 语言对所得结果进行可视化分析, 验证关键基因在 2 个基因芯片中表达趋势是否一致。

1.6 基因本体 (gene ontology, GO) 功能富集分析和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

将得到的关键基因以阈值 $P < 0.05$ 通过 Cytoscape 软件的 GlueGO 插件进行 GO 功能富集分析; 再基于 R Studio 和“clusterProfiler”“org.Hs.eg.db”“ggplot2”“pathview”等扩展包, 设定阈值 $P < 0.05$, 进行 KEGG 通路富集分析, 并且将排名前 2 的通路进行基因定位展示。

1.7 免疫细胞浸润分析

将所获得的 DEGs 基于 CIBERSORT 算法, 以阈值 $P < 0.05$ 进行 22 种免疫细胞浸润分析, 其中包

括 CD8⁺T 淋巴细胞、M1 巨噬细胞、活化的自然杀伤细胞等免疫细胞。

1.8 相关中药预测

将关键基因输入到 Coremine Medical 数据库^[8] (<http://www.coremine.com/>) 中, 下载该基因的相关中药信息, 以 $P < 0.05$ 为条件得到潜在治疗中药。将结果导入 Cytoscape 软件中进行可视化分析, 并筛选出对应 2 个关键基因及以上的关键中药, 再通过《中药大辞典》《中国药典》《中华本草》等分析其性味归经和功能主治, 最后通过查阅文献找出关键中药。

1.9 关键中药与对应基因分子对接验证

对得到的关键中药进行分析并查找其主要成分, 在 PubChem (<http://www.ncbi.nlm.nih.gov/>) 平台下载主要成分的 3D 结构, 从 PDB (<http://www.rcsb.org/>) 数据库下载药物候选靶蛋白的 3D 结构, 通过 PyMOL2.4.0 进行处理, 再通过 AutoDock vina 进行分子对接。对接结果用自由结合能判定优劣, 当自由结合能 $\leq -5.0 \text{ kJ/mol}$, 表示化合物与靶蛋白之间有良好的相互作用关系, 结合能越低, 对接模块越稳定, 二者相互作用的可能性越大, 选择自由结合能最低的对接结果进行可视化展示。

2 结果

2.1 芯片质量分析

对 GSE110147 基因芯片进行质量分析, 结果显示所得 RLE 箱线图各个样本中位数在同一条水平线上, 整体表达水平基本趋于一致。NUSE 相较于 RLE 则更为敏感, 可以进一步剔除降解严重的不可靠样本, NUSE 箱线图结果显示, 整体表达水平同样接近一致, 则所有数据可纳入研究, 见图 2。

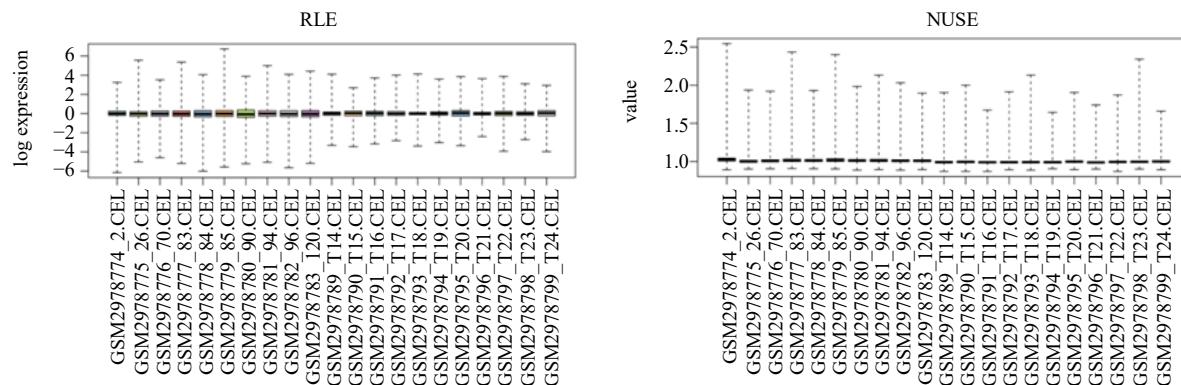


图 2 GSE110147 基因芯片质量分析箱线图

Fig. 2 Box diagram for quality analysis of GSE110147 gene chip

2.2 DEGs 获得

通过 R 包对 GSE110147 基因芯片分析结果显示, 标准化后样本表达基线一致, 共筛选得到

407 个 DEGs (上调基因 273 个, 下调基因 134 个), 并以火山图及热图展示排名前 100 的 DEGs, 见图 3。

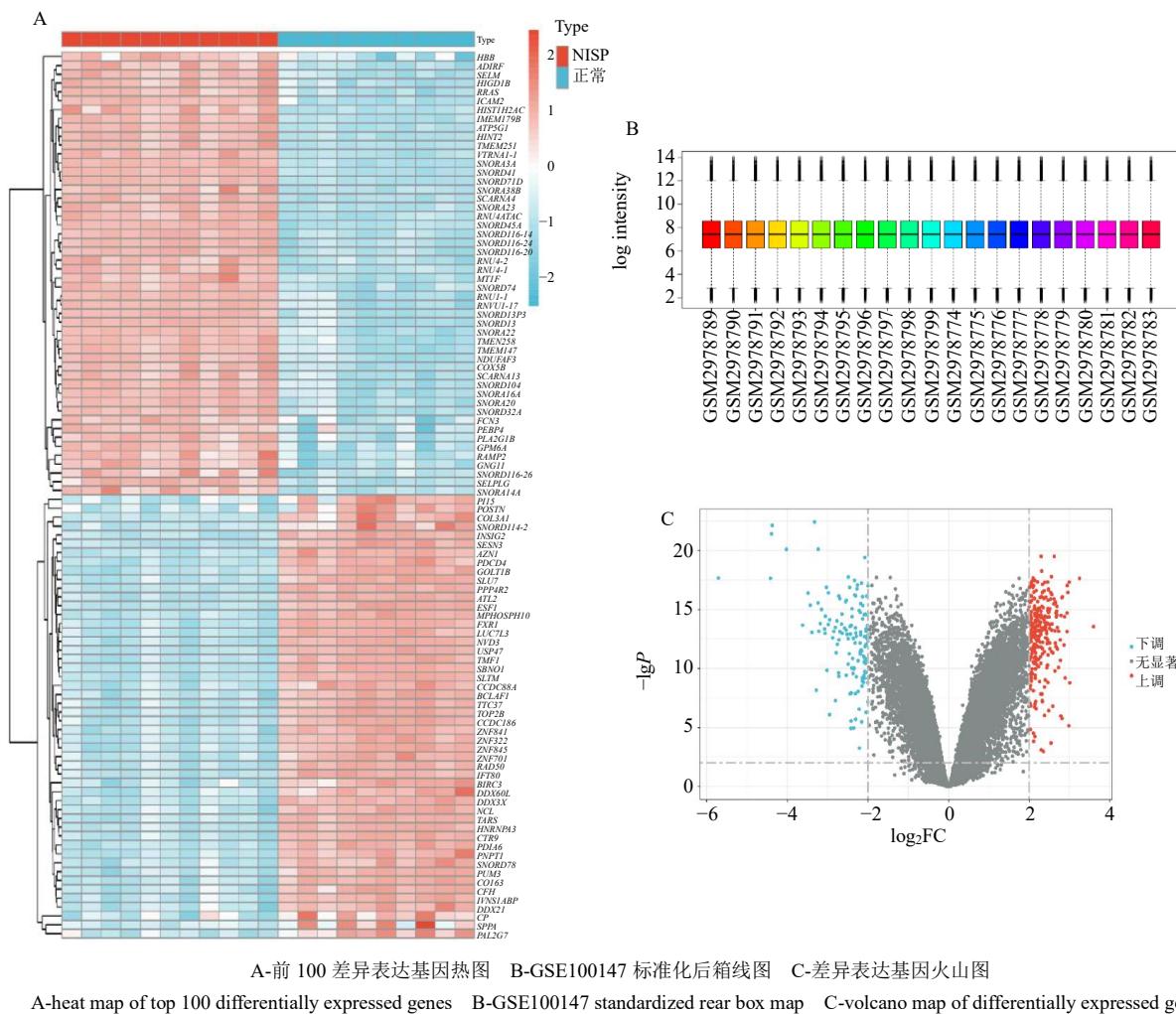


图 3 差异基因分析

Fig. 3 Analysis of differential gene

2.3 PPI 网络构建及关键基因获得

对所得 DEGs 构建 PPI 网络, 其中共有 297 个节点, 节点间共 1160 个相互作用关系。根据度值评分排序, 选出前 15 个基因, 分别为含 WD 重复域蛋白 36 (WD repeat-containing protein 36, *WDR36*)、ATP 结合盒转运蛋白 E1 (ATP-binding cassette sub-family E member 1, *ABCE1*)、*SKIV2L2* (superkiller viralicidic activity 2-like 2)、核糖体蛋白 S13 (Ribosomal protein S13, *RPS13*)、90 kDa 热休克蛋白 α 1 (heat shock protein HSP 90-alpha1, *HSP90AA1*)、真核翻译起始因子 3a (eukaryotic translation initiation factor 3 subunit A, *EIF3A*)、

POLR2B (DNA-directed RNA polymerase II subunit RPB2)、*SNORD34* (RPL13A 为其编码蛋白, small nucleolar RNA, C/D box 34)、真核翻译起始因子 5B (eukaryotic translation initiation factor 5B, *EIF5B*)、DEAD 框肽 21 (DExD-box helicase 21, *DDX21*)、GTP 结合蛋白 4 (GTP-binding protein 4, *GTPBP4*)、染色体结构维持蛋白 3 (structural maintenance of chromosomes protein 3, *SMC3*)、核仁 pre-rRNA 加工蛋白同系物 (nucleolar pre-rRNA processing protein homolog, *ESF1*)、核仁素 (nucleolin, *NCL*)、谷氨酰脯氨酰 tRNA 合成酶 (glutamate-tRNA ligase, *EPRS*), 其中 *RPS13* 和 *SNORD34* 为下调基因, 其

余为上调基因, 见图 4 和表 1。

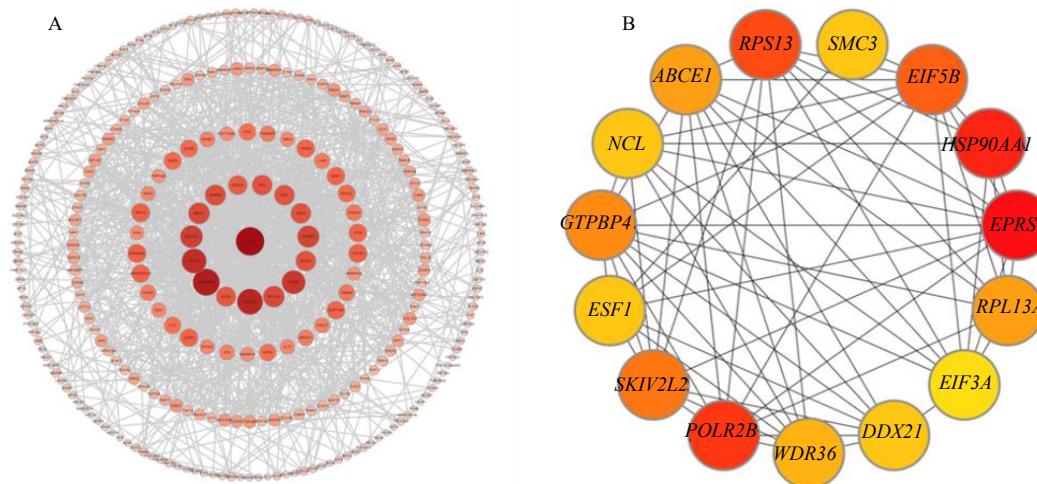
2.4 关键基因验证结果

将关键基因在基因芯片 GSE101286 进行验证后发现, 除基因 *POLR2B*、*SNORD34* 和 *DDX21* 以外 12 个基因所得验证结果与上述结果一致 ($P < 0.05$), 见图 5。

2.5 GO 功能富集和 KEGG 通路分析

关键基因的 GO 分析显示, 主要富集的生物过

程 (biological process, BP) 为蛋白质复性、翻译过程、高尔基组织、细胞骨架依赖型细胞内转运、水解酶活性; 细胞成分 (cellular component, CC) 为剪接复合体、染色体着丝粒区域、核斑点; 分子功能 (molecular function, MF) 为退火活性、微管蛋白结合、RNA 结合。KEGG 通路分析共获得 13 条富集通路, 其中核心靶点与 RNA 剪接体通路、内质网蛋白质加工通路联系最为密切, 除此之外还与黏



A-差异基因 PPI 互作网络 B-度值排名前 15 关键基因 (圆越大、颜色越深代表度值越高, 反之越小)

A-differential gene PPI interaction network map B-degree value ranked among the top 15 key genes (larger circles and darker colors represent higher degree values, and vice versa)

图 4 PPI 互作网络及关键基因

Fig. 4 PPI network and key gene

表 1 差异表达基因中度值排名前 15 的基因

Table 1 Top 15 differentially expressed genes with degree value

基因	度值	log ₂ FC	P 值	校正后的 P 值
<i>EPRS</i>	41	2.493 837	4.21×10^{-15}	4.22×10^{-13}
<i>HSP90AA1</i>	38	2.437 449	8.64×10^{-11}	1.07×10^{-9}
<i>POLR2B</i>	36	2.457 090	1.48×10^{-14}	1.08×10^{-12}
<i>RPS13</i>	34	-2.398 640	9.20×10^{-11}	1.14×10^{-9}
<i>EIF5B</i>	32	2.507 483	3.99×10^{-14}	2.23×10^{-12}
<i>SKIV2L2</i>	30	2.410 001	2.49×10^{-14}	1.57×10^{-12}
<i>GTPBP4</i>	29	2.096 090	1.55×10^{-14}	1.12×10^{-12}
<i>ABCE1</i>	28	2.034 032	1.20×10^{-11}	2.10×10^{-10}
<i>SNORD34</i>	28	-2.012 610	2.24×10^{-14}	1.46×10^{-12}
<i>WDR36</i>	27	2.195 173	4.38×10^{-13}	1.45×10^{-11}
<i>ESF1</i>	26	3.248 177	2.28×10^{-18}	2.54×10^{-15}
<i>DDX21</i>	26	2.725 238	8.17×10^{-11}	1.03×10^{-9}
<i>NCL</i>	26	2.511 015	6.93×10^{-13}	2.11×10^{-11}
<i>SMC3</i>	26	2.097 519	1.09×10^{-13}	4.77×10^{-12}
<i>EIF3A</i>	25	2.127 526	1.22×10^{-11}	2.13×10^{-10}

着斑激酶信号通路、癌症中蛋白多糖变化等存在联系, 见图6、7。

2.6 免疫细胞浸润

通过CIBERSORT算法所得结果(图8)显示, 与正常组织相比, NSIP患者组织中CD8⁺T细胞(T

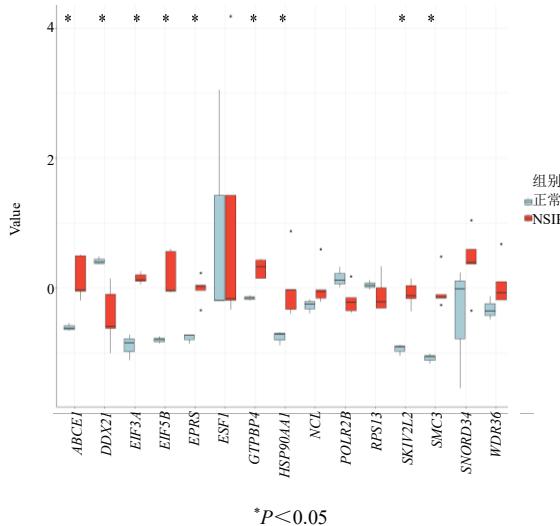


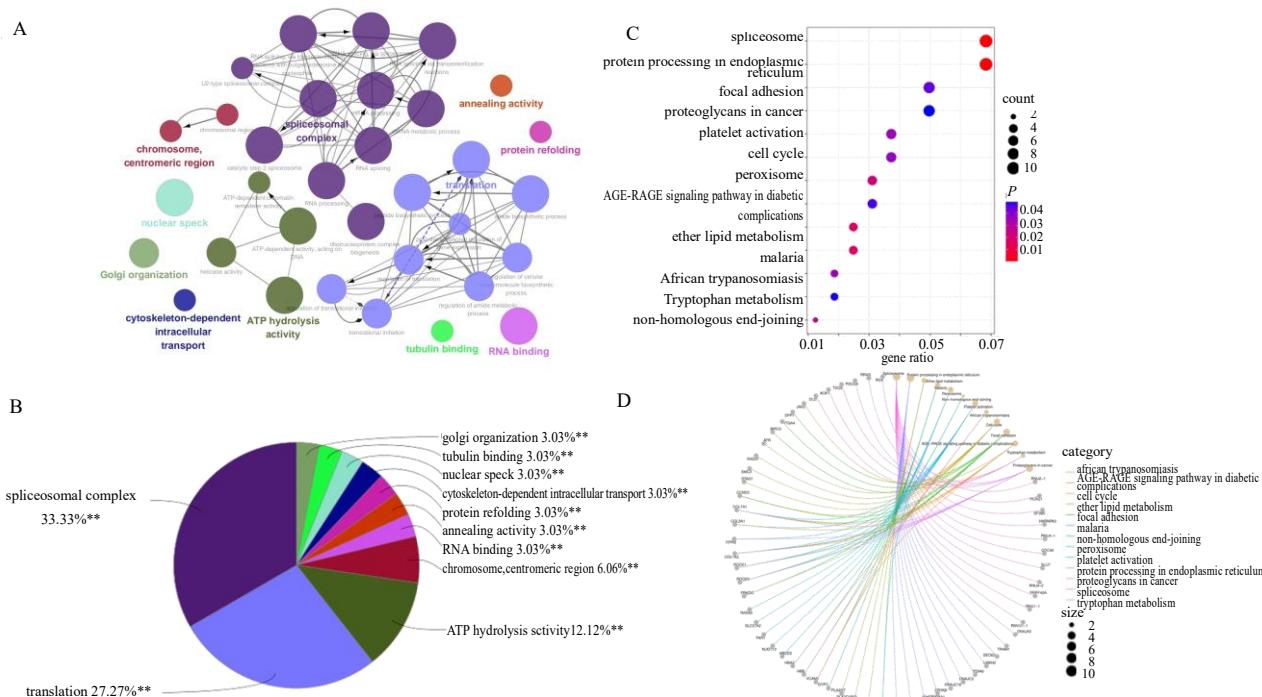
图5 关键基因验证

Fig. 5 Verification of key genes

cells CD8)浸润水平显著降低, 浆细胞(plasma cells)、活化的NK细胞(NK cells activated)和M1巨噬细胞(macrophages M1)表达水平均较低; NSIP患者组织中静息的CD4⁺记忆T细胞(T cells CD4 memory resting)浸润水平显著升高, 幼稚B细胞(B cells naive)、活化的CD4⁺记忆T细胞(T cells CD4 memory activated)、单核细胞(monocytes)、M0巨噬细胞(macrophages M0)、M2巨噬细胞(macrophages M2)、活化的树突状细胞(dendritic cells activated)、肥大细胞(dendritic cells activated)和嗜酸性粒细胞(eosinophils)的表达水平也较高。

而记忆B细胞(B cells memory)、幼稚CD4⁺T细胞(T cells CD4 naive)、滤泡辅助性T细胞(T cells follicular helper)、调节性T细胞(T cells regulatory, Tregs)、静息的NK细胞(NK cells resting)、中性粒细胞(neutrophils)表达水平差异不显著。

同时将差异最为显著的CD8⁺T细胞、静息的CD4⁺记忆T细胞与15个关键基因进行关联性分析, 发现除RPS13和SNORD34外, 其余13个基因与



A-GO功能富集类图 B-GO功能富集饼图 C-KEGG通路富集气泡图 D-KEGG通路富集环状图

A-cluster diagram of GO functional enrichment B-pie chart of GO functional enrichment C-bubble diagram of KEGG pathway enrichment D-ring diagram of KEGG pathway enrichment

图6 GO功能富集和KEGG通路富集分析

Fig. 6 Enrichment analysis of GO function and KEGG pathway

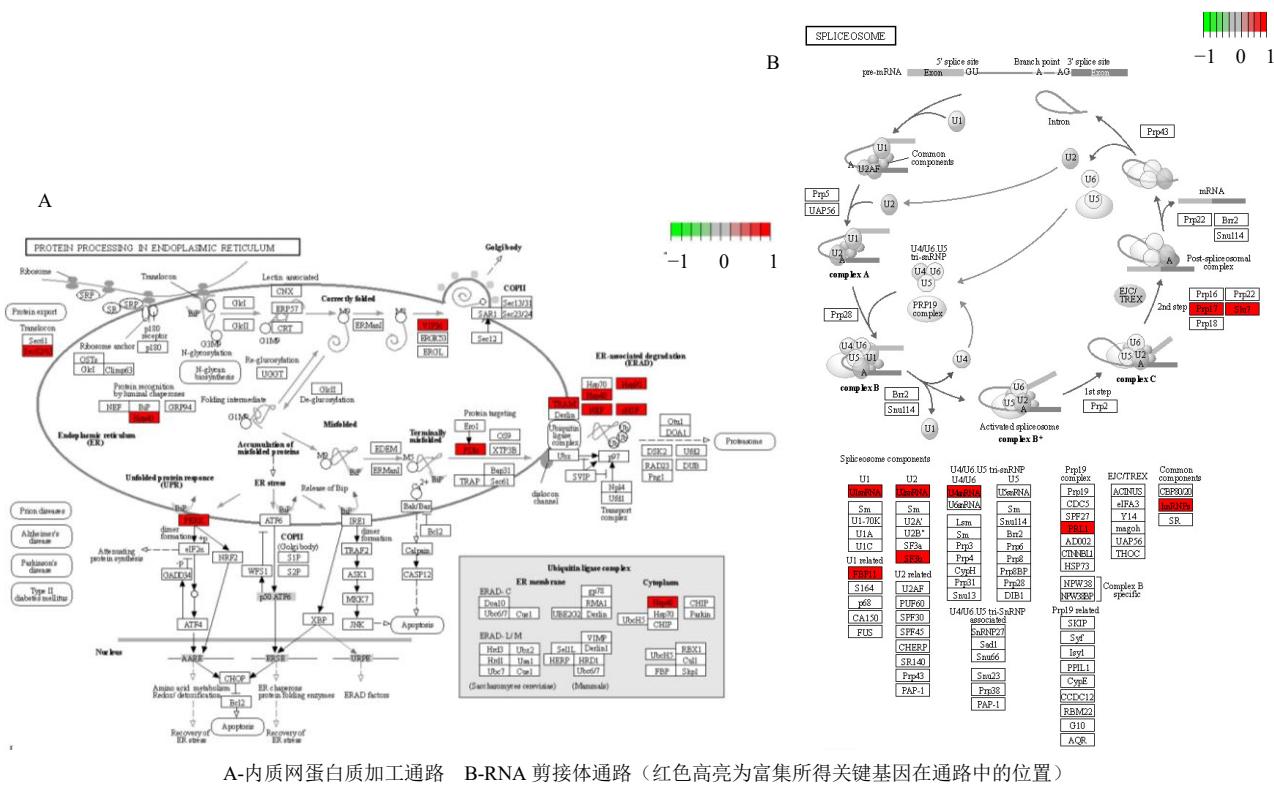


图7 KEGG富集结果排名前2的通路图

Fig. 7 Maps of top 2 pathway in KEGG enrichment results

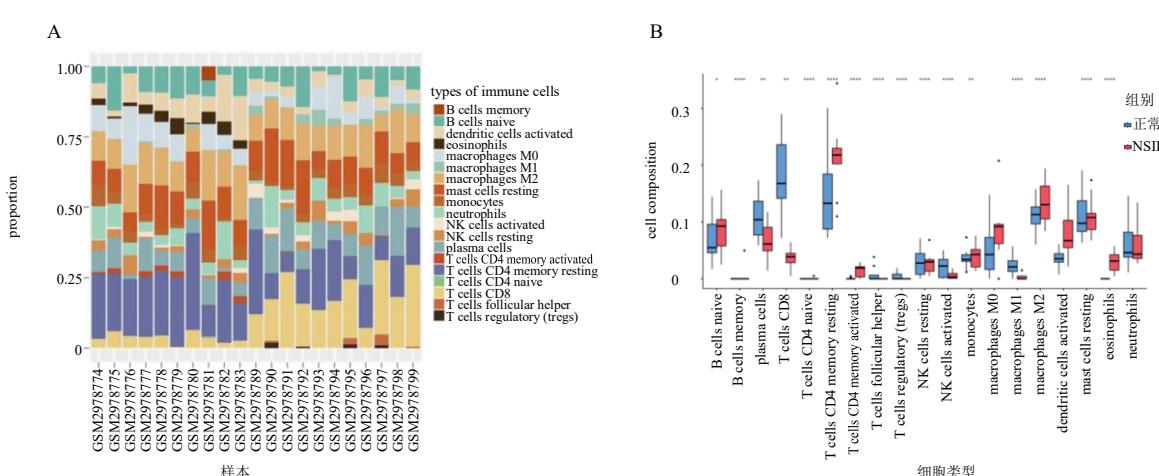


图8 免疫细胞浸润模式分析

Fig. 8 Pattern analysis of immune cell infiltration

CD8⁺T 细胞呈现负相关性，与静息的 CD4⁺记忆 T 细胞呈现正相关性，见图 9 ($P<0.05$ 、 $r>0.3$ 有相关性， $P<0.05$ 、 $r>0.8$ 为显著相关)。

2.7 中药预测分析

通过关键基因共筛选到 18 味中药，分别为冬菇、玉米须、向日葵花、松叶、松香、松节油、松

花粉、石刁柏、芒果核、落葵、雷公藤、葵花盘、浮小麦、茯神、茯苓、北豆根、板栗和浙贝母，其中四气五味主要集中于平、温、甘和苦，归经以肝、脾、心、肾和肺为主，功能主治偏重清热利湿、利尿消肿。通过查阅文献验证后雷公藤与茯苓治疗潜力较大，见图 10 和表 2，性味归经统计分析见图 11。

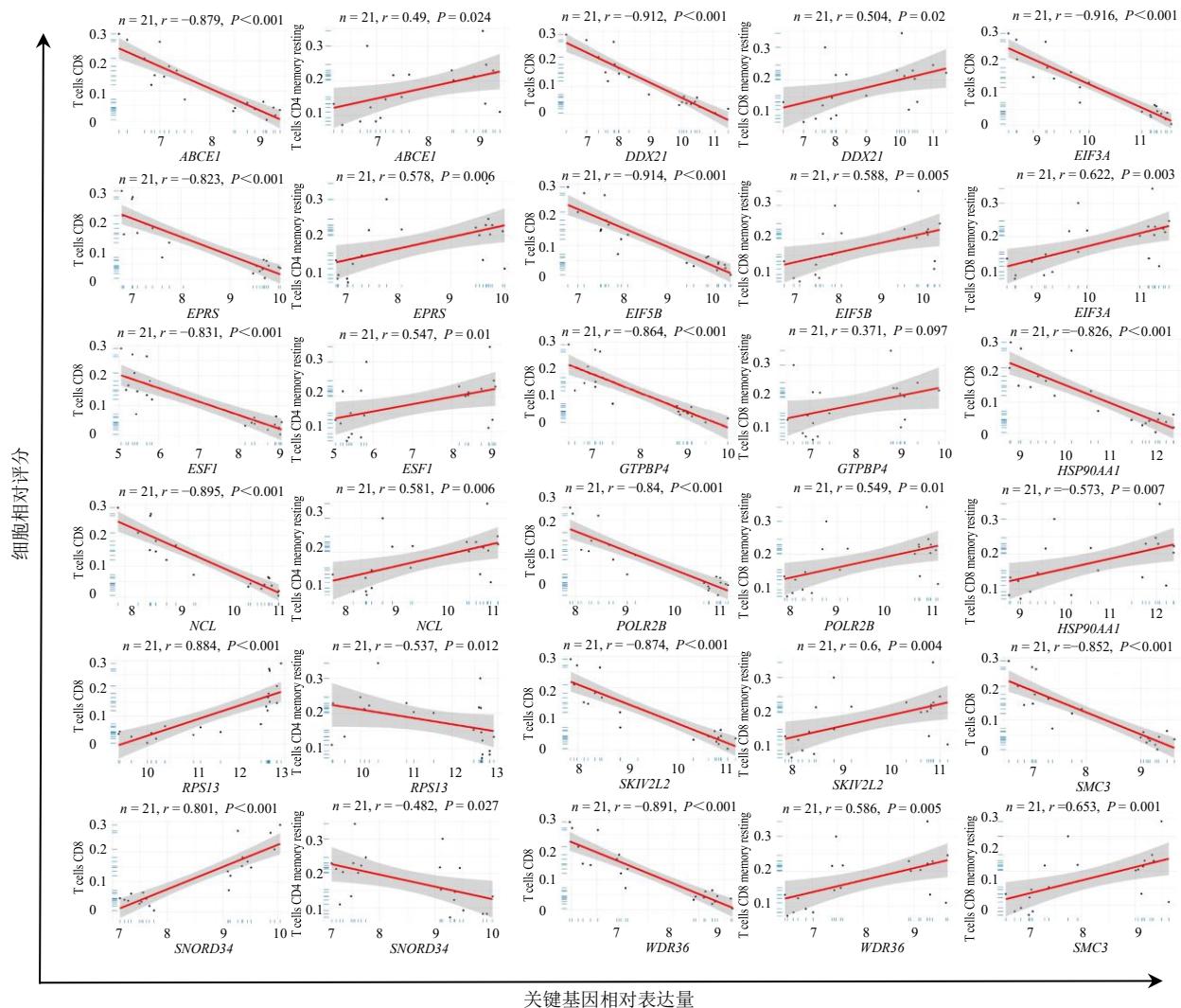
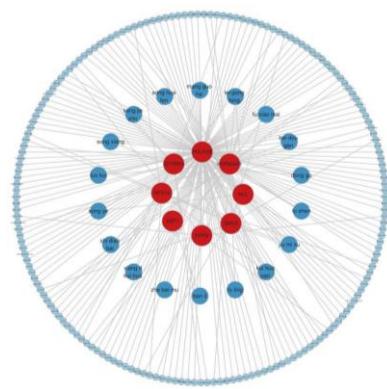


图9 关键基因与CD8⁺ T细胞、静息的CD4⁺记忆T细胞相关性分析

Fig. 9 Analysis of the relationship between key genes and T cells CD8 and T cells CD4 memory resting



内圈红色圆为关键基因，中间较大蓝色圆为18味关键中药，外圈蓝色圆为其余不显著中药
red circle in the inner ring is the key gene, the larger blue circle in the middle is 18 key traditional Chinese medicines, and the blue in the outer ring is the other insignificant traditional Chinese medicine

图10 中药预测分析

Fig. 10 Prediction and analysis of traditional Chinese medicine

2.8 关键中药分子对接验证

通过查阅文献以及口服生物利用度（oral bioavailability, OB）、类药性（drug-likeness, DL），分析关键中药雷公藤和茯苓主要活性物质，并从中选择5种主要活性物质与对应基因转录蛋白进行分子对接，得到其优势构象和自由结合能。结果表明，主要活性物质与其对应的基因转录蛋白的自由结合能均小于-5.0 kJ/mol，并且每个有效成分均可以通过氢键作用于其对应的蛋白氨基酸残基，见表3和图12、13。结果证明雷公藤与茯苓的主要活性物质与其对应蛋白的相互作用可能性较大，有潜在作用价值。

3 讨论

NSIP属于间质性肺炎的一种，是一种慢性肺部疾病，肺泡壁均匀增厚、炎症与免疫细胞浸润和胶

表2 关键基因预测中药分析

Table 2 Prediction analysis of traditional Chinese medicine by key genes

中药	对应基因	性味	归经	功能与主治
冬菇	<i>GTPBP4</i> 、 <i>POLR2B</i>	性寒, 味稍咸, 后微苦	肝、肾经	利肝脏、益肠胃、抗癌
玉米须	<i>RPS13</i> 、 <i>POLR2B</i>	甘、淡, 平	肾、胃、肝、胆经	利尿消肿、清肝利胆
向日葵花	<i>RPS13</i> 、 <i>POLR2B</i>	微甘, 平	肝经	肝肾虚、头晕
松叶	<i>NCL</i> 、 <i>POLR2B</i>	苦, 温	心、脾经	祛风燥湿、杀虫止痒
松香	<i>NCL</i> 、 <i>POLR2B</i>	苦、甘, 温	肝、脾经	燥湿、拔毒、生肌、止痛
松节油	<i>NCL</i> 、 <i>POLR2B</i>	苦、辛, 温	肝、肾经	舒筋活络、消肿止痛
松花粉	<i>NCL</i> 、 <i>POLR2B</i>	苦、甘, 温	肝、脾经	燥湿祛风、生肌止痛
石刁柏	<i>RPS13</i> 、 <i>POLR2B</i>	微甘, 平	肝经	清热利湿、活血散结
芒果核	<i>EPRS</i> 、 <i>POLR2B</i>	酸, 平	胃、小肠经	健胃消食、化痰行气
落葵	<i>HSP90AA1</i> 、 <i>POLR2B</i>	甘、酸, 寒	心、肝、脾、大肠、小肠经	滑肠通便、清热利湿、凉血解毒、活血
雷公藤	<i>NCL</i> 、 <i>HSP90AA1</i>	苦、辛, 寒	心、肝经	祛风除湿、杀虫、解毒
葵花盘	<i>RPS13</i> 、 <i>POLR2B</i>	甘, 寒	肝经	清热平肝、止痛、止血
浮小麦	<i>HSP90AA1</i> 、 <i>POLR2B</i>	甘, 寒	心经	除虚热、止汗
茯神	<i>HSP90AA1</i> 、 <i>POLR2B</i>	甘、淡, 平	心、脾经	宁心、安神、利水
茯苓	<i>HSP90AA1</i> 、 <i>POLR2B</i>	甘、淡, 平	心、脾、肺、肾经	利水渗湿、健脾和胃、宁心安神
北豆根	<i>HSP90AA1</i> 、 <i>POLR3B</i>	苦, 寒	肺、胃、大肠经	清热利咽、祛风除湿、解毒杀虫
板栗	<i>ESF1</i> 、 <i>POLR2B</i>	甘、微咸, 平	脾、肾经	益气健脾、补肾强筋、活血消肿、止血
浙贝母	<i>HSP90AA1</i> 、 <i>POLR2B</i>	苦, 寒	肺、心经	清热化痰、降气止咳、散结消肿

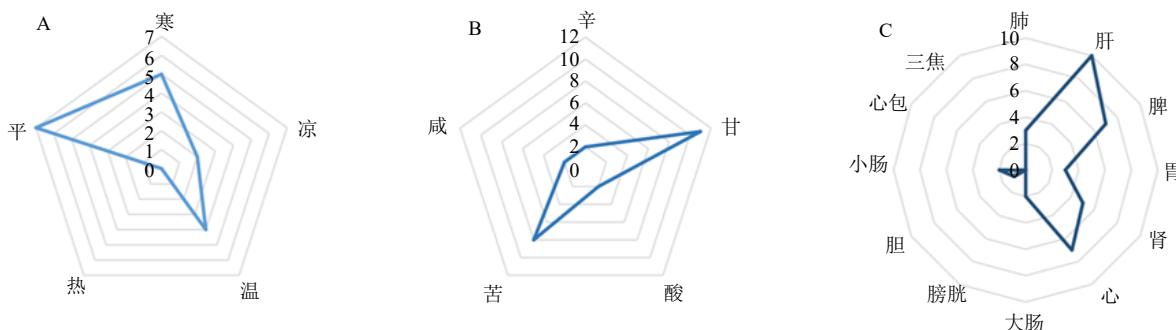


图11 中药四气(A)、五味(B)、归经(C)统计分析

Fig. 11 Statistical Analysis of four *qi* (A), five flavors (B) and meridian tropism (C) of traditional Chinese medicine

原纤维沉着是其主要的病理变化，常见症状为咳嗽与呼吸困难^[9]。

目前发病机制、诊断依据并不十分清楚，并且治疗方案副作用较多，治疗效果并不理想。生物信息学是多学科交叉融合的一门综合学科，而高通量技术是生物信息学中新一代的测序技术，能以更高通量、更低成本快速完成基因组和转录组的测序^[10-11]。因此，使用生物信息学技术对 NSIP 的关键基因、免疫细胞浸润等情况进行分析，可以对寻找 NSIP 发病机制起到推动作用。同时基于该技术，为 NSIP 寻找潜在治疗靶点，并进一步以此

预测潜在治疗中药，不仅可以推动传统中药发展达到老药新用的目的，以期对临床治疗提供一定研究基础。

3.1 关键基因与 NSIP 内在联系

本研究通过对基因芯片 GSE110147 进行分析，得到 273 个上调基因和 134 个下调基因。并且通过 PPI 网络分析筛选出 15 个与 NSIP 发病机制相关的关键基因。分别为下调基因 *RPS13*、*SNORD34* 和上调基因 *WDR36*、*ABCE1*、*SKIV2L2*、*HSP90AA1*、*EIF3A*、*POLR2B*、*EIF5B*、*DDX21*、*GTPBP4*、*SMC3*、*ESF1*、*NCL* 和 *EPKS1*，与基因芯片 GSE101286 验证

表3 关键中药主要活性物质与对应基因转录蛋白分子对接结果

Table 3 Docking results of key active substances of traditional Chinese medicine with corresponding gene transcriptional proteins

中药	活性成分	结合能/(kJ·mol ⁻¹)			OB/%	DL
		HSP90AA1	NCL	POLR2B		
雷公藤	tripfordin B2	-6.4	-6.0		107.71	0.76
	雷藤氯内酯醇 (tripchlorolide)	-6.8	-5.3		78.72	0.72
	canin	-6.2	-5.5		77.41	0.33
	雷公藤内酯酮 (triptonide)	-5.8	-6.9		68.45	0.68
	川陈皮素 (nobiletin)	-5.7	-6.1		61.67	0.52
茯苓	ergosta-7,22E-dien-3beta-ol	-8.2		-7.4	43.51	0.72
	麦角甾醇过氧化物 (ergosterol peroxide)	-6.7		-8.0	40.36	0.81
	栓菌酸 (trametenolic acid)	-6.9		-8.1	38.71	0.80
	茯苓新酸 C (poricoic acid C)	-6.0		-6.6	38.15	0.75
	啤酒甾醇 (cerevisterol)	-7.1		-7.2	37.96	0.77

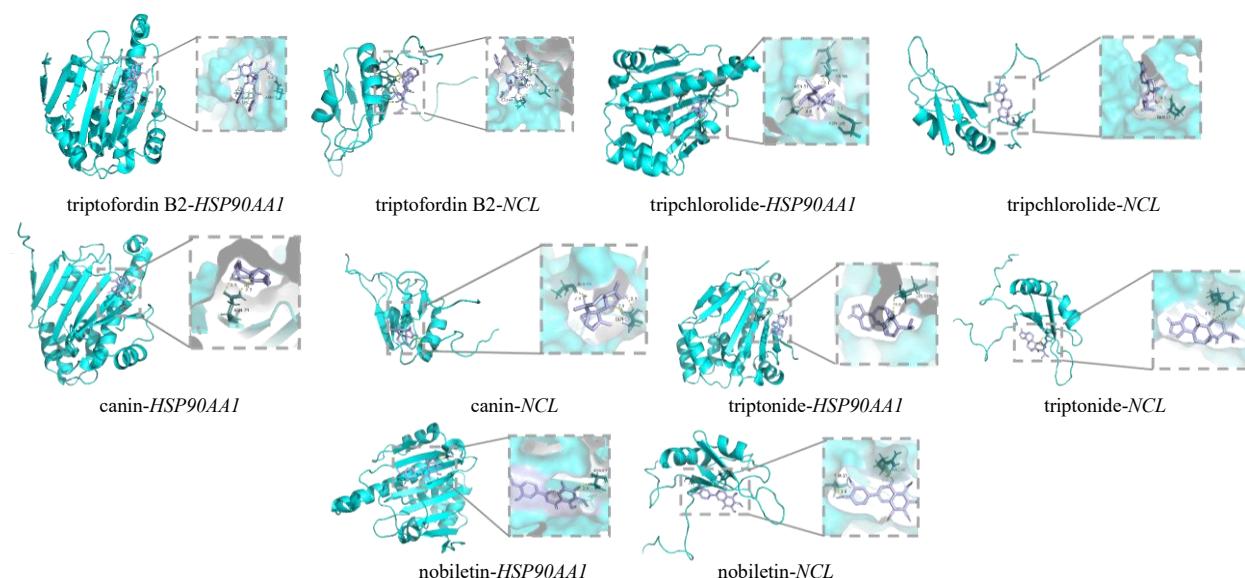


图12 雷公藤有效活性物质与对应基因分子对接验证

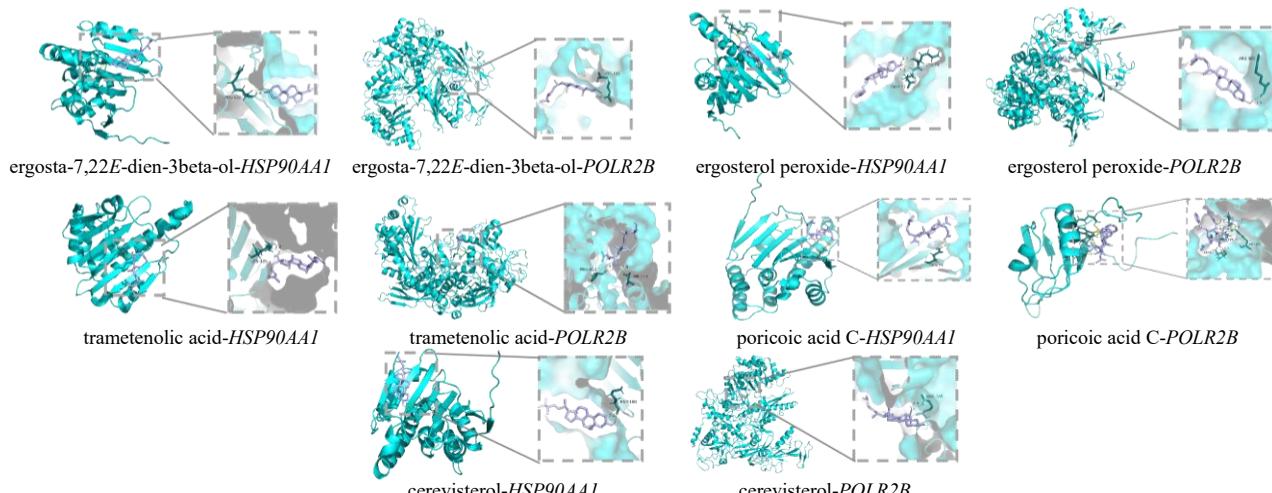
Fig. 12 Verification of docking between effective active substances of *Tripterygium wilfordii* and corresponding gene molecules

图13 茯苓有效活性物质与对应基因分子对接验证

Fig. 13 Verification of docking between effective active substances of *Poria cocos* and corresponding gene molecules

结果基本一致(除基因 *POLR2B*、*SNORD34* 和 *DDX21* 以外)。

15个关键基因中部分基因已有研究表明与肺部疾病有较大关联:*EPRS* 可以作为抗肺纤维化的治疗靶标, 并且通过分泌转化生长因子 β 1 (transforming growth factor β 1, TGF β 1) /信号转导和转录激活因子 (signal transducer and activator of transcription, STAT) 信号通路调节间充质标志物和细胞外基质蛋白的表达, 因此 *EPRS* 可作为治疗特发性肺纤维化的潜在靶标^[12]。*HSP90AA1* 在肺癌患者中高表达还与较差的生存预后相关, 并且可作为辅助诊断指标^[13-14]。*EIF5B* 在肺癌中表达较高且预后不良, 可以通过激活程序性死亡配体 1 (programmed cell death 1 ligand 1, PD-L1) 而抑制 T 细胞活性, 可作为干预肺癌的靶标^[15]。在非小细胞肺癌患者中 *GTPBP4* 的表达水平较高, 在敲低 *GTPBP4* 的小鼠体内发现癌细胞的增殖和侵袭能力降低^[16]。在人的肺腺癌组织中, *ABCE1* 表达水平较高, 与临床的分期也有联系, 在敲低 *ABCE1* 后可以减缓癌细胞的增殖, 并且与降低靶向药物的耐药性有关^[17-18], 而在非小细胞肺癌中, *ABCE1* 可以通过细胞趋化因子 7 (C-C motif chemokine 7, CCL7) 信号通路参与肺癌的进展^[19]。*SMC3* 在人肺癌 A549 细胞中敲低之后会抑制该细胞的增殖和迁移能力^[20]。*EIF3A* 通过参与肺泡上皮细胞间质转化作用而影响肺纤维化的进程, 通过抑制 *EIF3A* 表达可以改善肺纤维化程度, 目前已有针对该靶点进行治疗药物的研究^[21-23]。

肺纤维化为 NSIP 重要的病理变化, 且预后不理想。如上所述, 本研究筛选出的关键基因与肺部纤维化有所关联。因此, 在 NSIP 的进展中, 对这些基因的关注可能会给 NSIP 的防治带来新的研究方向。

3.2 富集结果与 NSIP 的内在联系

对差异基因进行 GO 功能富集分析结果显示, NSIP 与剪切体、翻译过程高度相关, 而 KEGG 通路富集分析结果显示 NSIP 与 RNA 剪接体通路、内质网蛋白质加工通路联系最为密切。其中剪切体出现于 2 项富集结果中, 提示其与 NSIP 可能存在紧密联系, 针对剪切体进行检索后, 发现目前研究中, 剪切体与肺癌的发病过程有较大关联性^[24-25], 并且剪切体的相关蛋白 130 在特发性肺纤维化中的纤维化病灶中高表达, 并且与肺部纤维化程度呈正相关, 与其发病过程密切相关^[26], 这提示剪切体可能会在

NSIP 进展过程中与其纤维化程度有所关联。而内质网应激在间质性肺炎纤维化的过程中也起到一定的作用^[27-28], 表明这些生物结构过程和通路为 NSIP 潜在的治疗靶点和研究方向。

结合前文对关键基因的分析, 关键基因与富集结果相互印证彼此的可靠性。而目前在 NSIP 中尚未有对高尔基体、微管蛋白结合、黏着斑激酶信号通路的研究, 也许在 NSIP 的发展过程中有着一定的作用, 可为后续的研究提供方向来揭示 NSIP 尚不明确的发病过程。

3.3 免疫浸润结果分析

对基因芯片进行 22 种免疫细胞浸润分析显示, NSIP 患者组织内静息的 CD4 $^{+}$ 记忆 T 细胞、活化的 CD4 $^{+}$ 记忆 T 细胞、嗜酸性粒细胞、幼稚 B 细胞、单核细胞、肥大细胞等浸润水平较高, 但静息的 CD4 $^{+}$ 记忆 T 细胞浸润水平远远高于活化的 CD4 $^{+}$ 记忆 T 细胞, 而 CD8 $^{+}$ T 细胞、浆细胞、活化 NK 细胞和 M1 巨噬细胞浸润水平较低。提示免疫细胞在 NSIP 的浸润模式有特异性, 多种具有免疫应答或促炎功能的细胞浸润水平显著提升, 故免疫细胞浸润在 NSIP 的发展过程中起到重要的作用。

免疫细胞浸润同样是 NSIP 重要的发病过程, 通过文献分析发现免疫细胞浸润模式与肺纤维化密切相关。而本研究中所得 CD4 $^{+}$ T 细胞和 CD8 $^{+}$ T 细胞在肺纤维化的进程中已被证实存在紧密联系, 外周血中的 T 淋巴细胞可分为 CD4 $^{+}$ T 细胞和 CD8 $^{+}$ T 细胞, CD4 $^{+}$ T 细胞能够通过释放多种细胞因子, 使病原体清除率增加, 在抗纤维化的免疫应答中起到关键作用; 而 CD8 $^{+}$ T 细胞则可以识别被感染的细胞, 使其凋亡, 达到清除病毒的目的同时引起炎症反应, 二者相互作用, 与其他免疫细胞共同维持着机体免疫系统的正常运行。而 CD4 $^{+}$ T 细胞与 CD8 $^{+}$ T 细胞的数量对肺纤维化的发展非常重要, 可通过提升二者比值增强体内的免疫功能, 降低肺纤维化的发生^[29], 同时通过抑制程序性死亡因子-1 增加 CD8 $^{+}$ T 细胞、效应 T 细胞的水平来改善肺纤维化水平也被多次证实有效^[30-32]。本研究结果显示 NSIP 患者体内静息的 CD4 $^{+}$ 记忆 T 细胞水平增加多于活化的 CD4 $^{+}$ 记忆 T 细胞, 转化为效应 T 细胞的能力受到影响。而 CD8 $^{+}$ T 细胞数量则显著减少, 反映 NSIP 患者体内 T 淋巴细胞数量减少, 免疫功能减弱, 抗肺纤维化的能力下降, 更易发生肺纤维化。而其余临床细胞浸润模式的改变, 可能会使得 NSIP 发病

过程中出现有别于其他肺病中肺纤维化的特性。

从关键基因与免疫细胞的相关性分析可知, 13个关键基因与最为显著的静息的 CD4⁺记忆 T 细胞呈现正相关, 与 CD8⁺T 细胞呈现负相关。不仅提示了分析结果的可靠性与一致性, 还反映出针对关键基因与免疫细胞用药来减轻肺纤维化程度值得深入研究。

目前 NSIP 诊断仍存在一些问题, 难以与其他间质性肺炎区别, 而对于免疫浸润模式及其相关基因的深入研究, 可以找出 NSIP 独特的浸润模式, 得到新的依据帮助诊断疾病, 并针对免疫细胞的差异性和肺纤维化的内在联系制定治疗方案, 弥补临幊上诊断困难和治疗效果不理想的现状。

3.4 中药预测结果分析

NSIP 在中医学中并未有详细准确的介绍, 且间质性肺炎在中医理论中也没有统一病名, 但有理论认为该病主要累及肺、脾、肾 3 脏, 致 3 脏皆虚。又根据其症状将其归为肺痿、肺痹, 前者病因归结为先天肾之真阴不足、后天外邪犯肺, 后者病因归结为肺气亏虚。本研究针对关键基因筛选治疗 NSIP 的潜在中药, 得到冬菇、玉米须、向日葵花、松叶、松香、松节油、松花粉、石刁柏、芒果核、落葵、雷公藤、葵花盘、浮小麦、茯神、茯苓、北豆根、板栗和浙贝母共 18 味中药, 主要归为心、肝、脾、肺、肾经, 与 NSIP 中医理论高度符合, 其中除与肺、脾、肾经直接关联以外, 心经和肝经同样有关, 肺气亏虚, 金虚木侮, 加重病情。气为血之帅, 血为气之母, 肺气亏虚, 宗气不足, 则心血运行不畅。故其治疗方案针对不同证候, 以清热化痰、活血化瘀、温阳补气和祛瘀解毒为主。18 味中药中, 浙贝母、北豆根、茯苓、茯神和雷公藤归经功效较为符合。

进一步通过查阅文献, 发现雷公藤和茯苓在肺部疾病上已有不少研究, 与 NSIP 更为符合。茯苓对应基因为 HSP90AA1 和 POLR2B, 雷公藤对应基因为 HSP90AA1 和 NCL, 与内质网蛋白加工通路密切相关。茯苓利水渗湿、健脾宁心, 在慢性阻塞性肺疾病^[33]、放射性肺损伤^[34]等肺部相关疾病中有一定治疗作用, 并且可以提升 CD4⁺ T 细胞的数量与 CD4⁺T 细胞/CD8⁺T 细胞水平^[35]。雷公藤清热解毒、祛风通络, 已有研究表明雷公藤可以通过多途径影响肺纤维化、哮喘、肺腺癌等肺部相关疾病^[36-39], 除此之外还有研究表明雷公藤对内质网应激也有影响, 与其对应内质网加工通路有较强关联^[40], 并且

还可以调节 T 淋巴细胞的表达^[41-42]。

通过分子对接也初步证明雷公藤与茯苓有效成分与其对应基因转录蛋白自由结合能较低, 说明有较大的相互作用可能性, 故雷公藤与茯苓对 NSIP 的潜在治疗价值更具有说服力。

4 结论

综上所述, 本研究以探究 NSIP 的关键基因和免疫浸润模式为主, 通过关键基因富集得到 NSIP 相关的生物学过程和信号通路, 完善其发病机制。通过已有研究推测关键基因在 NSIP 主要病理过程中作用。分析得到 NSIP 主要的免疫浸润模式后, 找到最为显著的免疫细胞与关键基因做相关性研究, 从基因与免疫浸润的角度共同阐述 NSIP 的主要病理过程。最后通过关键基因分析得到潜在治疗中药, 分析其性味归经和功能主治, 通过中医理论和已有文献分析其中最有潜在治疗价值的中药。研究结果有助于进一步阐释 NSIP 的发病机制和发展过程, 为 NSIP 后续实验研究、药物开发和临床应用提供一定的研究基础。但本研究仍然存在不足, NSIP 相关基因芯片数量较少, 后续研究应关注其发展, 及时补充纳入, 除此之外本研究还需进一步实验验证关键基因及相关中药的治疗效果。

利益冲突 所有作者均声明不存在利益冲突

参考文献

- American Thoracic Society, European Respiratory Society. American Thoracic Society/European Respiratory Society international multidisciplinary consensus classification of the idiopathic interstitial pneumonias. This joint statement of the American Thoracic Society (ATS), and the European Respiratory Society (ERS) was adopted by the ATS board of directors, June 2001 and by the ERS Executive Committee, June 2001 [J]. *Am J Respir Crit Care Med*, 2002, 165(2): 277-304.
- Belloli E A, Beckford R, Hadley R, et al. Idiopathic non-specific interstitial pneumonia [J]. *Respirology*, 2016, 21(2): 259-268.
- 王增智, 高杨, 李杰, 等. 非特异性间质性肺炎 6 例临床分析 [J]. 心肺血管病杂志, 2022, 41(10): 1066-1070.
- Hansell D M, Bankier A A, MacMahon H, et al. Fleischner Society: Glossary of terms for thoracic imaging [J]. *Radiology*, 2008, 246(3): 697-722.
- Li X, Chen C, Xu J F, et al. Nonspecific interstitial pneumonia and usual interstitial pneumonia: Comparison of the clinicopathologic features and prognosis [J]. *J Thorac Dis*, 2014, 6(10): 1476-1481.

- [6] Barrett T, Wilhite S E, Ledoux P, et al. NCBI GEO: Archive for functional genomics data sets: Update [J]. *Nucleic Acids Res*, 2013, 41: D991-D995.
- [7] Szklarczyk D, Gable A L, Lyon D, et al. STRING v11: Protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets [J]. *Nucleic Acids Res*, 2019, 47(D1): D607-D613.
- [8] Gorji-Bahri G, Moghimi H R, Hashemi A. RAB5A is associated with genes involved in exosome secretion: Integration of bioinformatics analysis and experimental validation [J]. *J Cell Biochem*, 2021, 122(3/4): 425-441.
- [9] Kono M, Nakamura Y, Yoshimura K, et al. Nonspecific interstitial pneumonia preceding diagnosis of collagen vascular disease [J]. *Respir Med*, 2016, 117: 40-47.
- [10] Aldridge S, Teichmann S A. Single cell transcriptomics comes of age [J]. *Nat Commun*, 2020, 11(1): 4307.
- [11] Ouzounis C A, Valencia A. Early bioinformatics: The birth of a discipline: A personal view [J]. *Bioinformatics*, 2003, 19(17): 2176-2190.
- [12] Song D G, Kim D, Jung J W, et al. Glutamyl-prolyl-tRNA synthetase regulates epithelial expression of mesenchymal markers and extracellular matrix proteins: Implications for idiopathic pulmonary fibrosis [J]. *Front Pharmacol*, 2018, 9: 1337.
- [13] 董晓玉, 钟涛, 叶元滋, 等. 血清 HSP90 α 和癌组织基因 HSP90AA1 在肺癌中的高表达及预后价值 [J]. 安徽医科大学学报, 2022, 57(7): 1034-1040.
- [14] Niu M Y, Zhang B, Li L, et al. Targeting HSP90 inhibits proliferation and induces apoptosis through AKT1/ERK pathway in lung cancer [J]. *Front Pharmacol*, 2021, 12: 724192.
- [15] Suresh S, Chen B B, Zhu J F, et al. eIF5B drives integrated stress response-dependent translation of PD-L1 in lung cancer [J]. *Nat Cancer*, 2020, 1(5): 533-545.
- [16] Wu J L, Chen G F, Wang W W, et al. GTPBP4: A new therapeutic target gene promotes tumor progression in non-small cell lung cancer via EMT [J]. *J Oncol*, 2022, 2022: 2164897.
- [17] Ren Y, Li Y H, Tian D L. Role of the ABCE1 gene in human lung adenocarcinoma [J]. *Oncol Rep*, 2012, 27(4): 965-970.
- [18] Wang L H, Lv X Y, Fu X H, et al. miR-153 inhibits the resistance of lung cancer to gefitinib via modulating expression of ABCE1 [J]. *Cancer Biomark*, 2019, 25(4): 361-369.
- [19] Wu Z, Tian Y, Yu Q, et al. The expression and correlation between chemokine CCL7 and ABCE1 in non-small cell lung cancer [J]. *Exp Ther Med*, 2018, 16(4): 3004-3010.
- [20] 余冬梅, 黄颖, 安瑜, 等. SMC3 对肺癌 A549 细胞增殖、迁移和成球能力的影响 [J]. 中国细胞生物学学报, 2017, 39(10): 1320-1325.
- [21] Guo R, Lv Y, Ouyang Y, et al. The role of miR-497/EIF3A axis in TGF β 1-induced epithelial-mesenchymal transition and extracellular matrix in rat alveolar epithelial cells and pulmonary fibroblasts [J]. *J Cell Biochem*, 2017, 118(10): 3401-3408.
- [22] Li X W, Li X H, Du J, et al. Calcitonin gene-related peptide down-regulates bleomycin-induced pulmonary fibrosis [J]. *Can J Physiol Pharmacol*, 2016, 94(12): 1315-1324.
- [23] 李先伟, 左东泽, 沈媛媛, 等. 降钙素基因相关肽对肺纤维化大鼠肺组织 eIF3a、p27 表达的影响 [J]. 中国应用生理学杂志, 2017, 33(1): 16-21.
- [24] Yang Y, Huang T Y, Fan Y H, et al. Significance of spliceosome-related genes in the prediction of prognosis and treatment strategies for lung adenocarcinoma [J]. *Biomed Res Int*, 2022, 2022: 1753563.
- [25] Blijlevens M, Komor M A, Sciarrillo R, et al. Silencing core spliceosome Sm gene expression induces a cytotoxic splicing switch in the proteasome subunit beta 3 mRNA in non-small cell lung cancer cells [J]. *Int J Mol Sci*, 2020, 21(12): 4192.
- [26] Liu K X, Liu D, Feng Y, et al. Spliceosome-associated protein 130: A novel biomarker for idiopathic pulmonary fibrosis [J]. *Ann Transl Med*, 2020, 8(16): 986.
- [27] Dobrinskikh E, Hennessy C E, Kurche J S, et al. Epithelial endoplasmic reticulum stress enhances the risk of Muc5b-associated lung fibrosis [J]. *Am J Respir Cell Mol Biol*, 2023, 68(1): 62-74.
- [28] Deng J L, He Y Q, Sun G C, et al. Tanreqing injection protects against bleomycin-induced pulmonary fibrosis via inhibiting STING-mediated endoplasmic reticulum stress signaling pathway [J]. *J Ethnopharmacol*, 2023, 305: 116071.
- [29] 李瑞琴, 金艳, 李伟, 等. 桂附地黄丸抗大鼠肺纤维化的作用及对外周血 T 淋巴细胞亚群的影响 [J]. 中华中医学学刊, 2016, 34(12): 2938-2940.
- [30] Barber D L, Wherry E J, Masopust D, et al. Restoring function in exhausted CD8 T cells during chronic viral infection [J]. *Nature*, 2006, 439(7077): 682-687.
- [31] Ji M, Liu Y, Li Q, et al. PD-1/PD-L1 pathway in non-small-cell lung cancer and its relation with EGFR mutation [J]. *J Transl Med*, 2015, 13: 5.
- [32] 陈则圣, 吴一鸣, 周婷, 等. PD-1 分子调控肺纤维化的机制进展及其在矽肺病研究中的展望 [J]. 公共卫生与预防医学, 2022, 33(4): 118-123.

- [33] 陈云坤, 刘煌, 张凤, 等. 小柴胡汤合桂枝茯苓丸治疗慢性阻塞性肺疾病急性加重期的临床研究 [J]. 时珍国医国药, 2022, 33(2): 422-424.
- [34] 杨依, 陈凡, 冯瑞兴, 等. 茯苓酸对小鼠急性放射性肺损伤的防护作用机制 [J]. 中国高原医学与生物学杂志, 2022, 43(4): 268-274.
- [35] 张海琴, 刘瑞芬. 桂枝茯苓胶囊对慢性盆腔炎大鼠T细胞亚群和红细胞免疫功能的影响 [J]. 中药药理与临床, 2013, 29(2): 6-8.
- [36] 周禹, 隋雅姿, 杨天明, 等. 雷公藤红素作用于HSP60抗肺纤维化的靶点研究 [J]. 药学学报, 2023, 58(3): 688-694.
- [37] 吕秀红, 郭瑞珍, 刁路明, 等. 雷公藤内酯醇对肺腺癌细胞系A549的体外抑制作用的研究 [J]. 中国药理学通报, 2007, 23(2): 207-210.
- [38] 马秀琴, 黄茂, 郭锡熔, 等. 雷公藤多甙对哮喘大鼠肺组织基质金属蛋白酶9及其抑制剂1表达的影响 [J]. 中国临床康复, 2005, 9(43): 117-119.
- [39] 徐资怡, 石金凤, 鲜静, 等. 雷公藤红素单用和联用抗肿瘤作用机制的研究进展 [J]. 中草药, 2021, 52(14): 4372-4385.
- [40] 徐成波, 郑瑞玑, 廖斌, 等. 雷公藤内酯醇通过CHOP/BCL-2内质网应激途径诱导人多发性骨髓瘤细胞凋亡 [J]. 中国病理生理杂志, 2022, 38(7): 1274-1282.
- [41] Abdin A A, Hasby E A. Modulatory effect of celastrol on Th1/Th2 cytokines profile, TLR2 and CD3⁺ T-lymphocyte expression in a relapsing-remitting model of multiple sclerosis in rats [J]. Eur J Pharmacol, 2014, 742: 102-112.
- [42] Ji W, Li H, Gao F, et al. Effects of *Tripterygium glycosides* on interleukin-17 and CD4⁺ CD25⁺ CD127 low regulatory T-cell expression in the peripheral blood of patients with ankylosing spondylitis [J]. Biomed Rep, 2014, 2(4): 517-520.

[责任编辑 潘明佳]