

慢性呼吸系统疾病关键标志物筛选与辅助诊断模型建立研究

李祖豪, 吴元斌

青海民族大学数学与统计学院, 青海 西宁

收稿日期: 2025年5月13日; 录用日期: 2025年7月17日; 发布日期: 2025年7月28日

摘要

目的: 探讨COVID-19与慢性阻塞性肺疾病及哮喘的生物学共同作用机制并建立基于少数基因的疾病诊断模型。方法: 本研究结合转录组数据分析、差异基因筛选、GO和KEGG功能富集分析、蛋白互作网络构建、统计学分析及机器学习方法, 从多角度解析COVID-19对慢性呼吸系统疾病患者的影响。结果: 筛选出差异共同表达基因, 并通过富集分析揭示其主要参与免疫应答、炎症信号传递及病毒感染相关的分子通路; 蛋白互作网络分析发现了TLR2、MMP9、CXCR4、CCR7、IL1A和CXCL8等关键基因, 可能在炎症调控和疾病进展中起重要作用; 同时, 对于慢性阻塞性肺疾病, 建立了PTPN7等少数基因的诊断模型, 对于哮喘, 建立了CLC等少数基因的诊断模型。结论: COVID-19通过多系统通路加剧慢性呼吸系统疾病的炎症和免疫失调, 为个性化治疗策略提供了新的研究方向。同时, 建立了基于基因的疾病诊断模型, 为未来的治疗和管理提供了重要的启示。

关键词

COVID-19, 慢性呼吸系统疾病, 差异分析, 免疫细胞浸润分析, 机器学习

Research on Key Biomarker Screening and Auxiliary Diagnostic Model Building for Chronic Respiratory Diseases

Zuhao Li, Yuanbin Wu

School of Mathematics and Statistics, Qinghai Nationalities University, Xining Qinghai

Received: May 13th, 2025; accepted: Jul. 17th, 2025; published: Jul. 28th, 2025

文章引用: 李祖豪, 吴元斌. 慢性呼吸系统疾病关键标志物筛选与辅助诊断模型建立研究[J]. 生物医学, 2025, 15(4): 780-797. DOI: 10.12677/hjbm.2025.154084

Abstract

Objective: To explore the common biological mechanisms of interaction between COVID-19 and chronic obstructive pulmonary disease (COPD) as well as asthma, and to establish disease diagnostic models based on a small number of genes. **Methods:** This study combined transcriptome data analysis, differential gene screening, GO and KEGG functional enrichment analysis, protein-protein interaction network construction, statistical analysis, and machine learning methods to comprehensively analyze the impact of COVID-19 on patients with chronic respiratory diseases from multiple perspectives. **Results:** Differential co-expressed genes were identified, and enrichment analysis revealed that they mainly participate in immune responses, inflammatory signaling pathways, and virus infection-related molecular pathways. Protein-protein interaction network analysis identified key genes such as TLR2, MMP9, CXCR4, CCR7, IL1A, and CXCL8, which may play important roles in inflammation regulation and disease progression. Meanwhile, for COPD, a diagnostic model based on a small number of genes including PTPN7 was established, and for asthma, a diagnostic model based on a small number of genes including CLC was established. **Conclusion:** COVID-19 exacerbates the inflammation and immune dysregulation in chronic respiratory diseases through multiple systemic pathways, providing new research directions for personalized treatment strategies. Additionally, the establishment of gene-based disease diagnostic models offers important insights for future treatment and management.

Keywords

COVID-19, Chronic Respiratory Diseases, Differential Analysis, Immune Cell Infiltration Analysis, Machine Learning

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

新型冠状病毒肺炎(Coronavirus Disease 2019, COVID-19)是一种全球性的大流行病,自2019年底首次爆发以来,迅速蔓延,对全球公共卫生、社会经济造成了深远影响[1]。COVID-19作为一种新兴的全球性传染病,对慢性呼吸系统疾病患者造成了显著影响。COVID-19通过多种机制加剧COPD和哮喘的病理过程,包括病毒对呼吸道上皮细胞的直接损伤、引发过度的炎症反应(细胞因子风暴)以及对宿主免疫系统的长期影响。这些机制导致慢性呼吸系统疾病患者的病情加重、住院率和死亡率显著升高。

COVID-19的发病机制复杂,其影响已超越单一疾病的范畴,形成了一种“综合症”模式(syndemics),即COVID-19与其他慢性疾病的共存和相互作用显著加重了病情[2]。特别是,对于慢性呼吸系统疾病患者,如慢性阻塞性肺疾病(Chronic Obstructive Pulmonary Disease, COPD)和哮喘(Asthma),COVID-19的感染显著提高了严重并发症、住院率和死亡率的风险。

COPD和Asthma均是以气道炎症和呼吸功能受损为特征的常见慢性呼吸系统疾病,二者之间既有病理学上的差异,又存在一定的相似性和重叠性。COPD主要由不可逆的气流受限和肺部结构破坏引起,通常与吸烟和长期空气污染有关;而Asthma则以可逆性气道高反应性为主要特征,多为过敏性炎症所致。近年来的研究表明,部分患者同时表现出Asthma和COPD的特征,这种重叠性被称为Asthma-COPD

重叠综合症(ACO) [3], 进一步凸显了这两种疾病在机制上的相互关联。在后 COVID 时代, 针对 COPD 和 Asthma 患者的研究显示, 感染后的长期影响(“长新冠”综合症)可能导致气道慢性炎症持续存在, 从而进一步恶化患者的呼吸功能。这种后遗症不仅涉及呼吸系统, 还可能引发代谢紊乱和心理健康问题。

综上所述, COVID-19 对慢性呼吸系统疾病的影响不仅是急性阶段的危害, 更体现为复杂的多系统相互作用以及对长期健康的深远影响。研究 Asthma 与 COPD 的共病特性, 以及它们与 COVID-19 的交互机制, 不仅有助于理解疾病的共同病理过程, 也为开发新的治疗策略提供了重要的启示。

2. 材料与方法

2.1. 数据来源

从基因表达综合(Gene Expression Omnibus, GEO)数据库(www.ncbi.nlm.nih.gov/geo)下载 GSE164805、GSE148004 数据集用于生物信息学分析 COVID-19 对慢性呼吸系统疾病的影响。通过差异基因分析, 我们识别了 COVID-19 与 COPD、哮喘共有基因表达的变化, 这些基因在免疫应答和炎症反应中起关键作用。此外, 从 GEO 数据库下载了 GSE212331、GSE137268 数据集用于研究分析及辅助诊断。具体说明见表 1。

Table 1. Data set description

表 1. 数据集说明表

GSE 序列	疾病	使用样本	注释平台	组织
GSE164805	COVID-19	10 个新冠患者和 5 个健康个体	GPL26963	外周血单核细胞
GSE148004	COPD、Asthma	9 个 Asthma 患者、7 个 COPD 患者和 9 个健康个体	GPL13497	痰液基因表达
GSE212331	COPD	72 个 COPD 患者和 15 个健康个体	GPL10558	痰液基因表达
GSE137268	Asthma	54 个 Asthma 患者和 15 个健康个体	GPL6104	痰液基因表达

2.2. 筛选共同差异表达基因

采用 R 软件的 Limma 软件包筛选 GSE164805、GSE148004 数据集研究 COVID-19 和 COPD、Asthma 的差异基因表达。分析 P 值以校正 GEO 数据集中的假阳性结果。P < 0.05 和 |Fold Change (FC)| ≥ 2 被定义为筛选 GSE164805 数据集 COVID-19 基因差异表达的阈值, P < 0.05 和 |Fold Change (FC)| ≥ 1 被定义为筛选 GSE148004 数据集 COPD、Asthma 基因差异表达的阈值。分别去除没有对应基因符号的探针集或有多个探针集的基因或取平均值。再通过使用在线维恩图工具

(<https://bioinfogp.cnb.csic.es/tools/venny/index.html>)获得它们的重叠 DEGs。

2.3. 共同 DEGs 的富集分析及 PPI 网络分析

将共同 DEGs 输入微生信在线生物信息学分析(<https://www.bioinformatics.com.cn/>)进行 GO、KEGG 富集分析。将共同 DEGs 输入 STRING 数据库(<https://cn.string-db.org/>)得到蛋白质互作网络分析, 构建一个具有复杂调控关系的 PPI 网络, 其中综合得分超过 0.4 的交互作用被认为具有统计学意义。使用 Cytoscape (<http://www.cytoscape.org>)来可视化这个 PPI 网络。最后, 通过 GeneMANIA 数据库绘制关键模块基因, 通过 PPI 网络的结果结合 Cytoscape 的插件方法筛选关键基因。

2.4. 基于 DSigDB 数据库的候选药物识别

为了确定针对 COPD 和 Asthma 的病理机制的候选药物, 我们利用了 Richrweb 平台 (<https://maayanlab.cloud/Enrichr/enrich>) 中的药物签名数据库(DSigDB)。

2.5. 免疫细胞浸润分析

从 GEO 数据库中获得 COPD 和 Asthma 样本的基因表达数据。通过 r 语言中进行免疫细胞浸润分析。输入基因表达矩阵, 输出免疫细胞的比例。比较两种疾病样本中免疫细胞比例差异, 使用箱线图、热图分析展示免疫浸润模式。

2.6. 对关键基因的统计 *t* 检验

使用 python 软件在 GSE212331、GSE137268 数据集上寻找 GSE148004 数据集识别的关键基因, 对两种疾病的关键基因在新的数据集上进行非配对的统计 *t* 检验, 判断关键基因在疾病组和健康组有无统计上的显著差异。 $P < 0.1$, 视为差异有统计学意义。

2.7. 自适应 Lasso 回归筛选关键基因

采用 R 软件的 glmnet、caret 软件包使用自适应的 Lasso 回归进一步从共同 DEGs 筛选关键基因, 通过筛选出的少量关键基因用于诊断 COPD 和 Asthma 两种疾病。

2.8. 机器学习方法建立诊断模型

为了构建诊断模型, 我们采用了逻辑回归建模。我们分别在 GSE212331、GSE137268 上验证我们筛选出的关键识别诊断基因的可靠性。通过原模型与逆概率加权和超参数调优后的模型进行对比说明效果, 研究总流程见图 1。

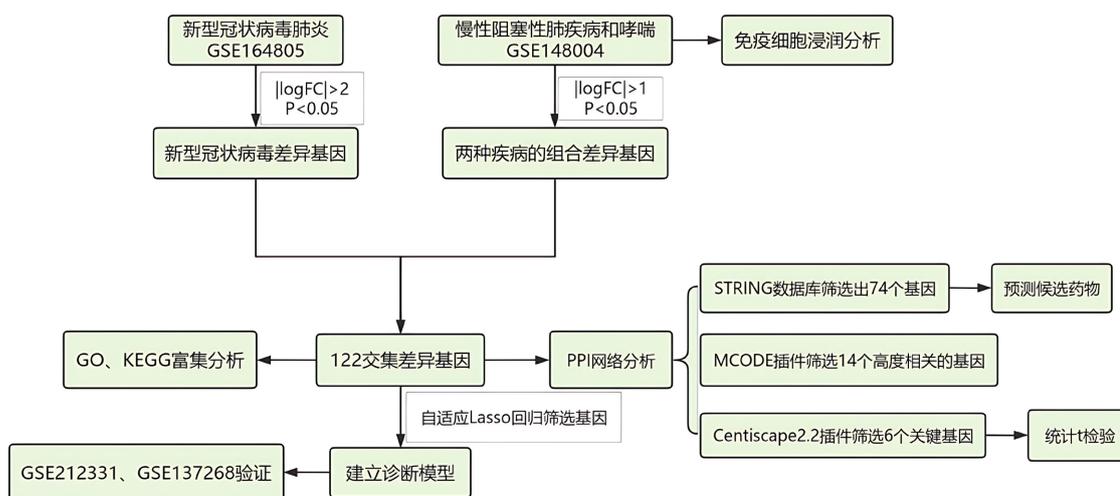


Figure 1. Overall flowchart
图 1. 总流程图

3. 结果

3.1. DEGs 筛选结果

在 GSE164805 数据集中, 经过去除没有对应基因符号的探针集以及对基因名去重后的 DEGs 共 2201

个。在 GSE148004 数据集进行差异分析, 进行同样操作后得到 DEGs 为 892 个。差异分析的结果如图 2(A)、图 2(B)所示。最终得到的交集 DEGs 有 122 个, 如图 3(A)所示。

在 GSE164805 数据集中, 我们筛选出 2201 个差异表达基因(DEGs), 这些基因主要参与免疫应答、炎症信号传递及病毒感染相关的分子通路。通过与 COPD 和哮喘的 DEGs 进行交集分析, 发现 122 个共同 DEGs, 这些基因可能在 COVID-19 与慢性呼吸系统疾病的相互作用中起关键作用。

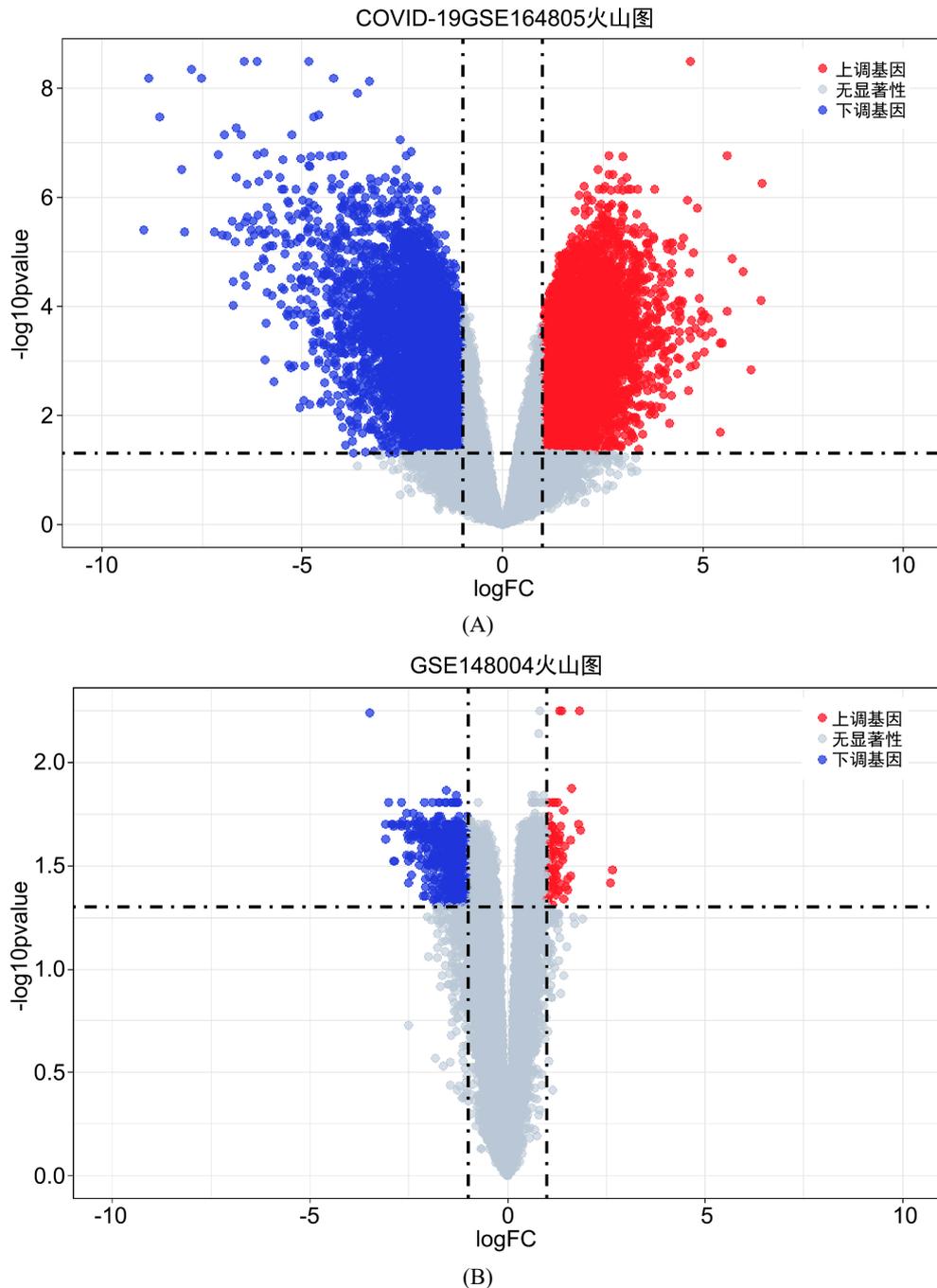


Figure 2. Differential analysis volcano plot
图 2. 差异分析火山图

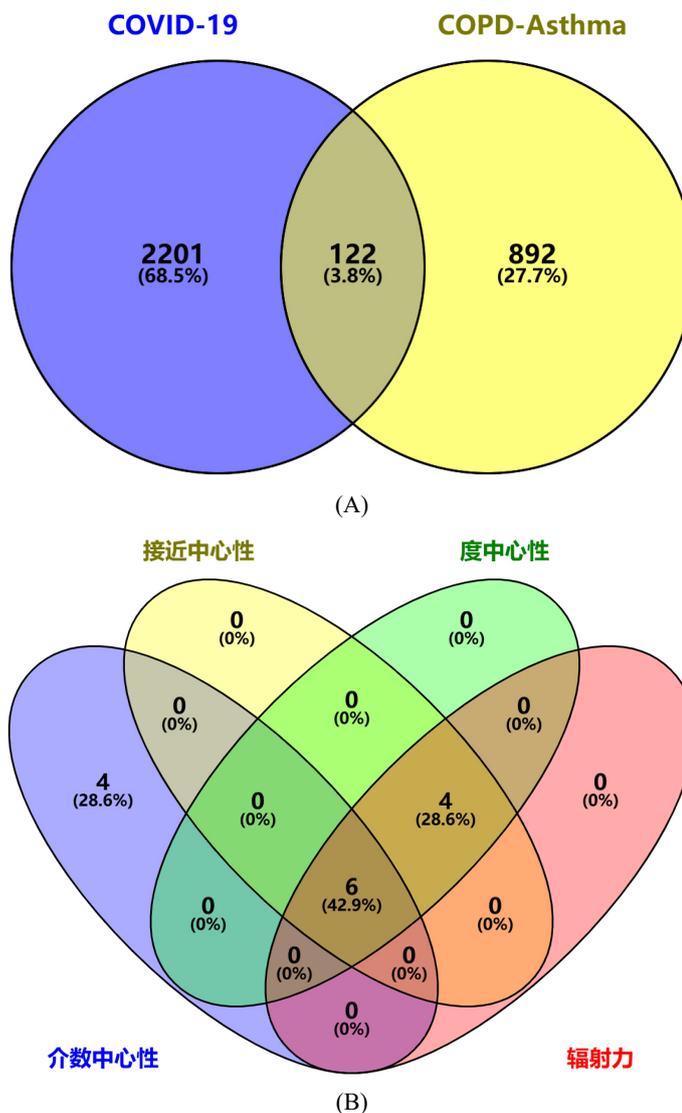
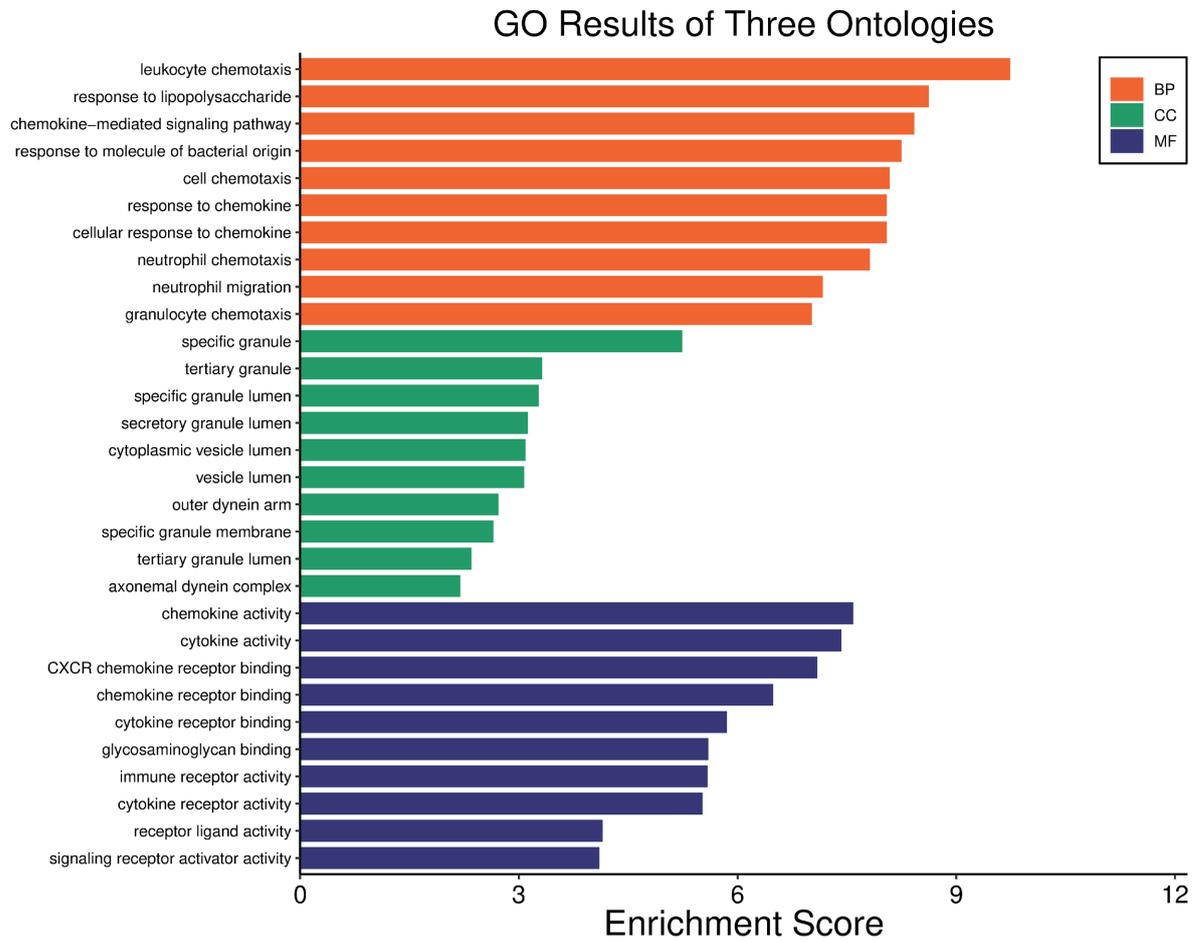


Figure 3. Venn diagram
图 3. 韦恩图

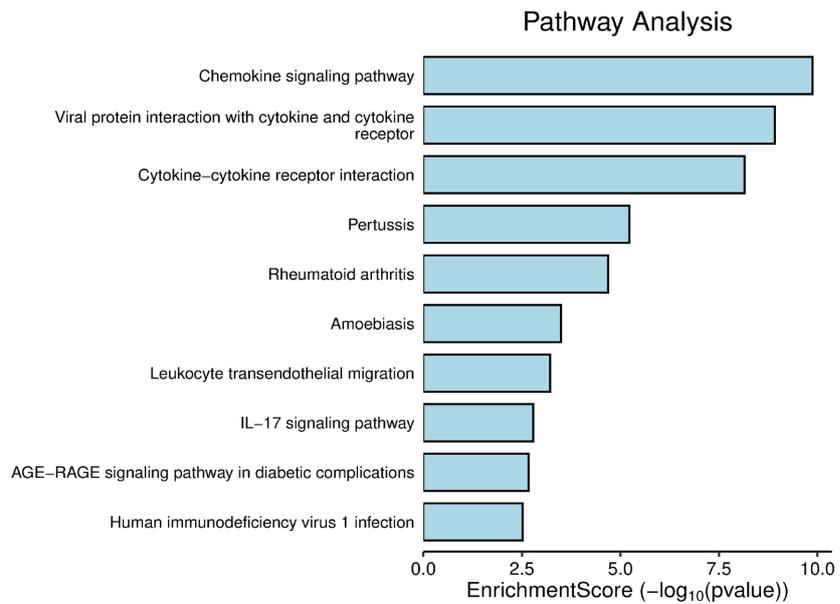
3.2. GO 和 KEGG 富集分析结果

为了分析这些交集差异基因所涉及的生物学功能和通路，我们进行了 GO 和 KEGG 通路富集分析，如图 4 所示。

结果揭示了关键生物学过程，白细胞趋化和对脂多糖的响应，这表明免疫应答增强，与炎症和感染相关。白细胞趋化涉及免疫细胞迁移至炎症或感染部位，而对脂多糖的响应则与细菌感染识别有关。中性粒细胞趋化是炎症早期的关键，涉及其迁移和释放毒性颗粒。分子功能分析显示趋化因子和细胞因子活性，吸引免疫细胞并调节免疫反应。趋化因子通过 CXCR 受体激活免疫细胞。细胞组分分析显示免疫相关的细胞表面受体和信号复合体参与信号传导。KEGG 通路分析显示这些基因主要参与趋化因子信号通路、病毒蛋白与细胞因子和细胞因子受体的相互作用，以及细胞因子-细胞因子受体的相互作用，其中趋化因子信号通路的富集得分最高，表明其在免疫应答中的重要作用。其他两条通路也在免疫应答和病毒感染中显著，显示相关基因的重要性。



(A)



(B)

Figure 4. Enrichment analysis bar chart
图 4. 富集分析柱状图

使用 STRING 数据库构建了组合得分 > 为 0.4 的重叠 deg 的 PPI 网络, 去除单个基因或者没有和主要基因联系的基因后基于介数中心性可视化的结果如图 5(A)所示。其中包含个 74 节点。通过 Cytoscape 的 MCODE 插件获得了一个高度紧密连接的基因模块, MCODE 算法会遍历网络中的节点和边, 寻找具有高密度的区域。它通过计算每个节点的邻居数量和这些邻居之间的连接情况来确定密集区域。最终发现其中包含 14 个常见的 DEGs 和 83 对相互作用对如图 5(B)所示。

通过插件 centiscape2.2 的四种算法, 我们计算出了前 10 个 hub 基因。具体的算法排序结果如表 2 所示。在图 2(B)韦恩图的交集后, 我们发现了 6 个重叠的 hub 基因, 包括 TLR2、MMP9、CXCR4、CCR7、IL1A 和 CXCL8, 表 3 显示了它们的全名和相关功能。

Table 2. Ranking of key genes

表 2. 关键基因排序表

介数中心性	接近中心性	度中心性	辐射力
TLR2	MMP9	TLR2	MMP9
MMP9	TLR2	MMP9	TLR2
CXCR4	CXCR4	CXCR4	CXCR4
CCR7	IL1A	CXCL8	IL1A
VCAN	CXCL8	CCR7	CXCL8
IL1A	CCR7	IL1A	CCR7
FCER1A	CXCL1	CXCL1	CXCL1
PLIN2	CXCR1	CXCR1	CXCR1
CXCL8	CXCL13	CXCL13	CXCL13
KLRB1	CXCL5	CXCL5	CXCL5

Table 3. Functional table of key genes

表 3. 关键基因功能表

基因符号	基因全名	功能
TLR2	Toll-Like Receptor 2	TLR2 是一种模式识别受体, 它能够识别多种病原体相关分子模式, 参与免疫细胞的激活和细胞因子的产生, 调节炎症反应。
MMP9	Matrix Metalloproteinase 9	MMP9 (基质金属蛋白酶 9)是一种酶, 它参与细胞外基质重塑, 对组织重建、炎症和肿瘤发展等生理和病理过程至关重要。
CXCR4	C-X-C Motif Chemokine Receptor 4	CXCR4 是一种 G 蛋白偶联趋化因子受体, 调节造血、免疫细胞迁移和组织稳态。在神经系统形成和血管生成中至关重要。
CCR7	C-C Motif Chemokine Receptor 7	CCR7 是一种趋化因子受体, 支持适应性免疫应答。在肿瘤中与淋巴转移相关。调节免疫系统中的淋巴细胞归巢和炎症反应。
IL1A	Interleukin 1 Alpha	IL1A (白介素 1 α)是一种促炎细胞因子, 参与炎症反应和免疫应答的调节。激活炎症过程, 并促进其他细胞因子的产生。
CXCL8	C-X-C Motif Chemokine Ligand 8	CXCL8 是 CXC 趋化因子家族的一员, 也是经典的促炎性因子。引导中性粒细胞迁移至感染或炎症部位, 增强活性。有效的血管生成因子, 促进血管生成和组织修复。

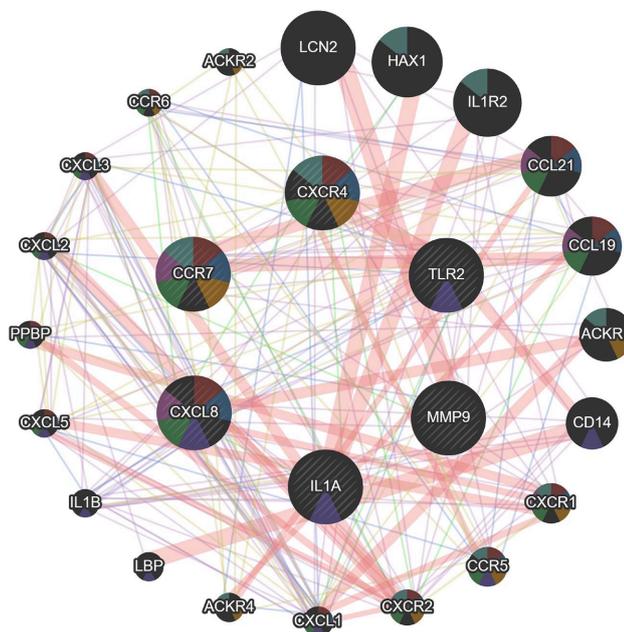


Figure 6. Key genes network diagram

图 6. 关键基因网络图

图 6 为关键基因网络图, 基于 GeneMANIA 数据库, 我们分析了其共表达网络及其相关功能。这些基因表现出复杂的 PPI 网络, 共表达量为 46.44%, 物理相互作用量为 36.72%, 共享蛋白结构域为 11.08%, 协同定位为 4.75%, 遗传相互作用为 1.01%。这些基因与白细胞趋化性、白细胞迁移、细胞趋化性、趋化因子结合、中性粒细胞迁移、细胞对细菌来源的分子的反应、细胞因子结合有关。

3.4. 基于 DSigDB 数据库预测可能的小分子药物

通过 STRING 数据库得到的 74 个基因输入 Enrichr 平台中的药物签名数据库(DSigDB), 利用 richr 上的 DSigDB 药物数据库分析模型基因, 以确定潜在的靶向药物。最终得到了候选药物表 4 如下所示(仅展示前五项)。

Table 4. Genetically targeted drug formulas for COPD and asthma

表 4. 慢性阻塞性肺疾病和哮喘的基因靶向药物表

小分子药物	调整 P 值	结合分数	基因名
benzene	7.81E-12	427.07	CXCL8; IL1R2; PTGER2; IFI6; CXCR4; CXCL1; CSF2RB; MMP9; PAPSS2; IL1A; LGALS2; VCAN; ALDH2; BCL6; HYAL2; HAS1; NAMPT; PELI1; CYP1A1; KCNJ2; TLR2
NICKEL SULFATE	1.45E-11	410.38	CXCL6; CSF3R; CXCL8; IL1R2; CXCR4; CXCL1; CSF2RB; DNAH9; MMP9; CXCL5; IL1A; ZEB1; ADORA2A; NAMPT; CD207; CYP1A1; SNAI1; CCR7; TIMP1; KCNJ2
aspirin	7.70E-11	367.32	CSF3R; CXCL8; KLRB1; IL1R2; PTGER2; PXN; ICAM3; CXCL1; FPR2; FBXO11; MMP9; IL1A; SLPI; ADORA2A; ADORA3; FCER1A; NLRP3; TIMP1; CLC
MS-275	2.12E-8	284.06	GALNT12; CXCL6; CXCL8; IL1R2; PTGER2; CXCR4; CXCL1; IL1RAP; CXCL5; IL1A; VCAN; BCL6; PELI1; CYP1A1; KCNJ2
PMA	8.71E-7	192.51	CXCL6; CXCL8; IL1R2; CYBB; IL1RAP; MMP9; PRK CZ; CXCL5; IL1A; ZEB1; CXCR1; CYP1A1; TIMP1; CCL18

3.5. 免疫细胞浸润分析

COPD 和 Asthma 这种炎症相关疾病, 免疫细胞在疾病的发生和进展中起着关键作用。通过分析基因表达矩阵, 可以推断样本中免疫细胞的浸润水平, 进一步揭示免疫细胞在疾病中的作用机制[4]。图 7 和图 8 分别是两种疾病的热图和箱线图。

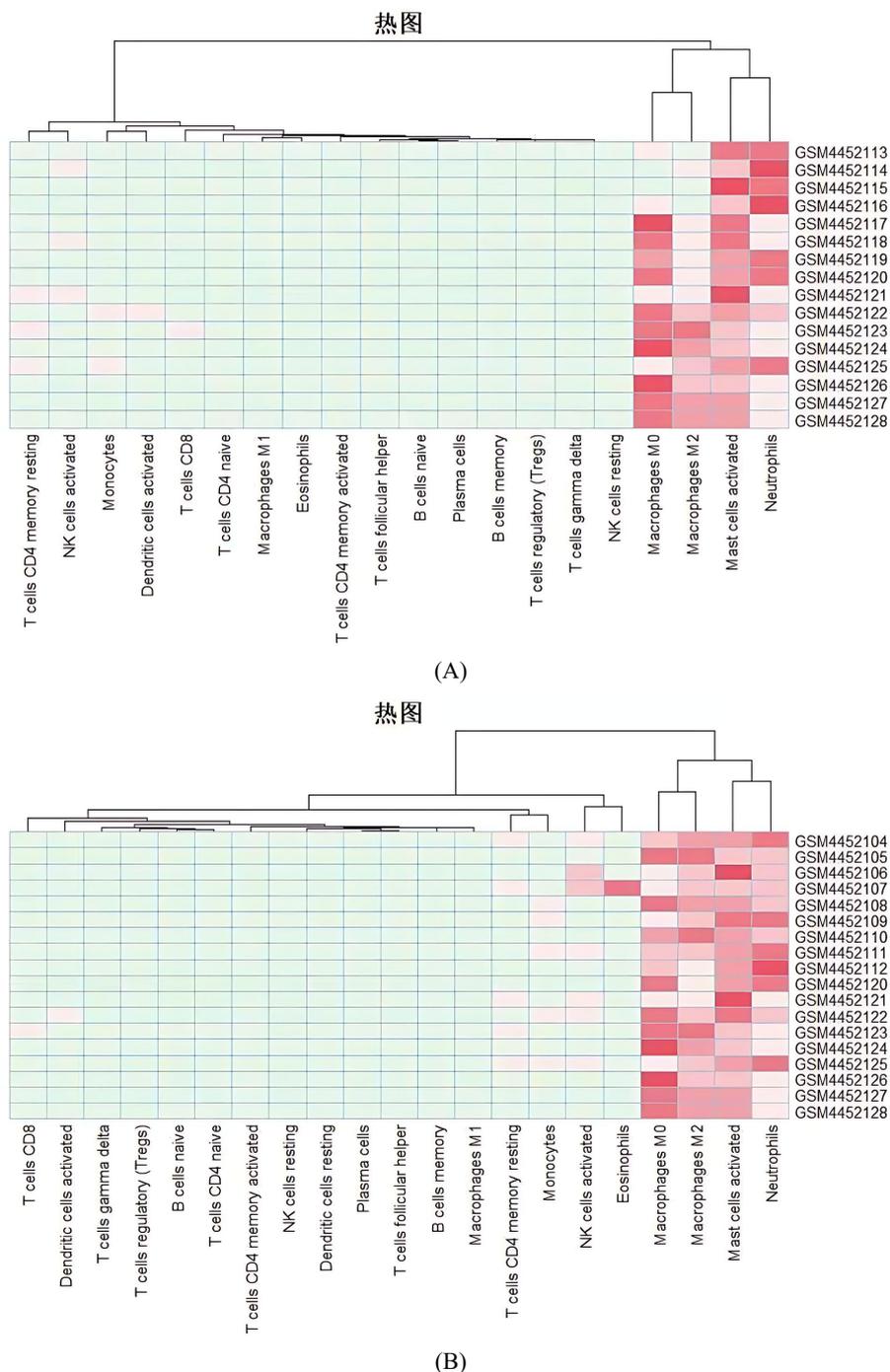


Figure 7. Heatmap of COPD and asthma
图 7. 慢性阻塞性肺疾病和哮喘热图

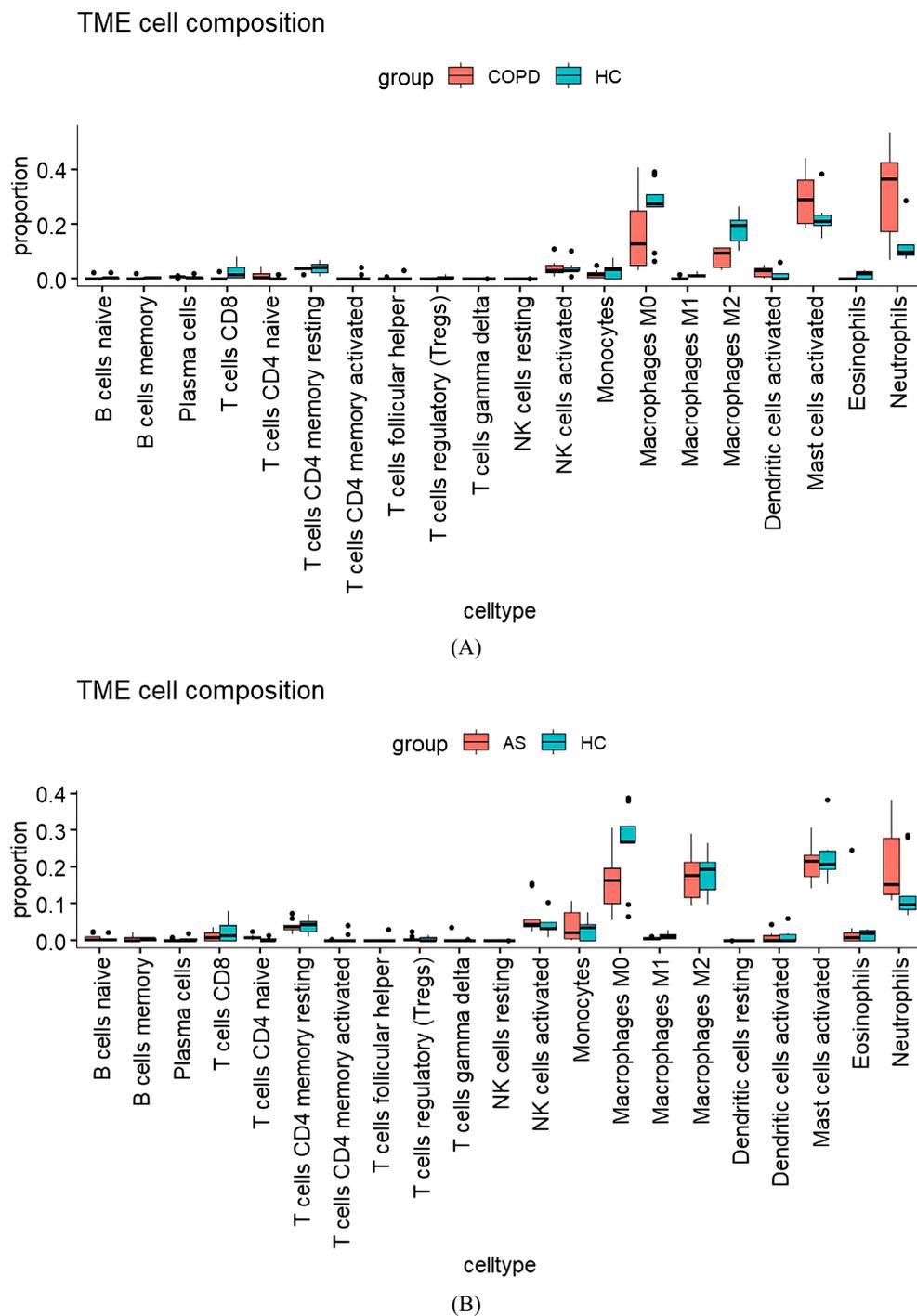


Figure 8. Boxplot of COPD and asthma
图 8. 慢性阻塞性肺疾病和哮喘箱线图

对于 COPD 来说, 差异性细胞类型的中性粒细胞和 M1 型巨噬细胞颜色最深, 说明这些免疫细胞在患病样本中的浸润比例显著较高。调节性 T 细胞和静息树突细胞在部分样本中比例较低, 表明这些细胞可能在疾病中受到抑制。

从组间趋势来看, 患病样本整体显示出更多的炎症性免疫细胞。而健康个体样本中可能存在免疫稳

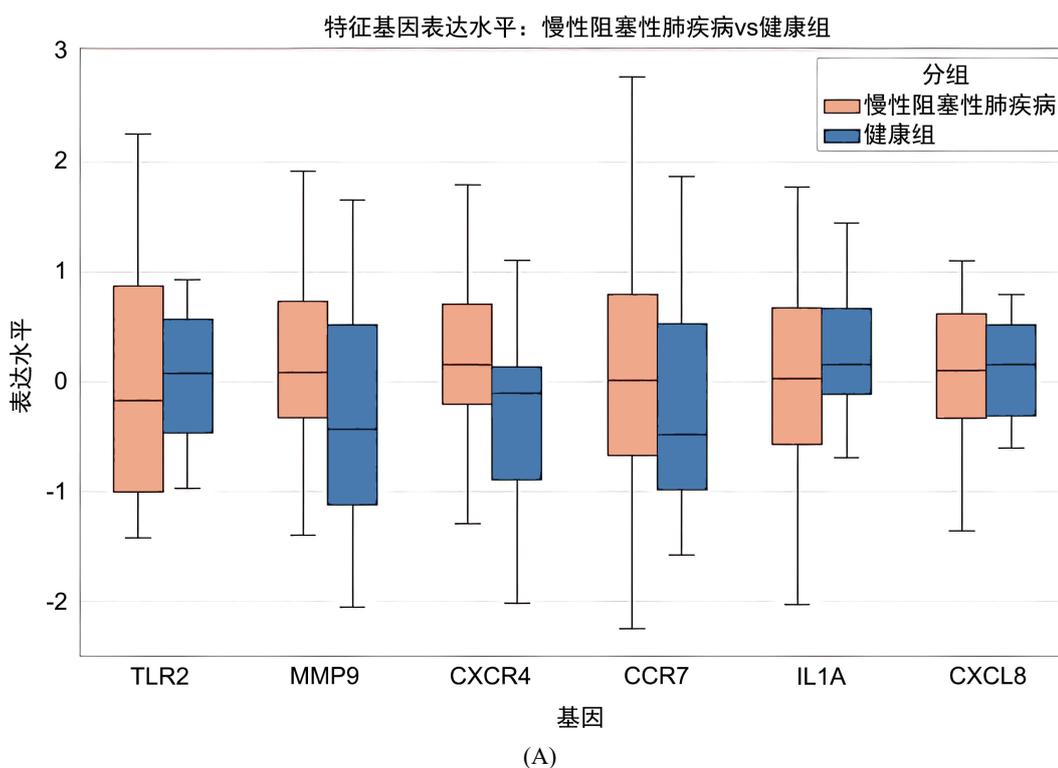
态相关细胞比例的平衡。由此可知, 中性粒细胞和 M1 巨噬细胞主导了疾病的炎症环境。这与 COPD 的慢性炎症病理机制一致。调整型 T 细胞的减少可能导致免疫调节能力下降, 从而加剧炎症反应。M2 巨噬细胞的增加可能反映组织修复, 但如果过度激活可能导致纤维化。中性粒细胞和 M1 巨噬细胞的比例升高可作为疾病的潜在生物标志物。可用于进一步分型 COPD 的炎症亚型。

对于哮喘来说, 差异性细胞类型中的嗜酸性粒细胞和中性粒细胞颜色最为显著, 这暗示着在患病样本里, 这两种免疫细胞的浸润比例有着明显的上升。活化的肥大细胞和 M2 巨噬细胞也在其中扮演了一定角色, 可能参与到组织的修复过程或是引发纤维化现象。调节性 T 细胞呈现低表达状态, 或许意味着机体的免疫调节功能受到了损害。

从患病样本的整体组间趋势来看, 更多地呈现出与炎症相关的免疫细胞特征。相较于健康个体样本中免疫稳态相关细胞比例所维持的平衡状态, 哮喘患病样本的这种情况更为突出。由此可见, 嗜酸性粒细胞和肥大细胞等主导了哮喘的炎症环境, 这与哮喘存在慢性炎症的病理基础相符。调节性 T 细胞数量的减少, 可能导致免疫调节能力变弱, 进而使炎症反应加剧。而 M2 巨噬细胞的增加虽然可能反映出组织在尝试进行修复, 但如果被过度激活, 则有可能引发纤维化问题。嗜酸性粒细胞和肥大细胞的浸润比例升高, 有望作为疾病严重程度评估以及治疗反应预测的潜在生物标志物, 从而为哮喘不同免疫亚型的划分提供重要依据。

3.6. 关键基因 t 检验结果

对于基于介数中心性、接近中心性、度中心性、辐射力四种方法筛选的关键基因的基因表达量进行统计上的 t 检验, 验证患病人群和健康人群之间的基因表达有无显著差异[5]。利用 GSE212331、GSE137268 数据集作为测试集测试关键基因在不同样本下有无显著差异, 图 9 为关键基因表达量箱线图, 表 5 为关键基因的统计 t 检验结果。



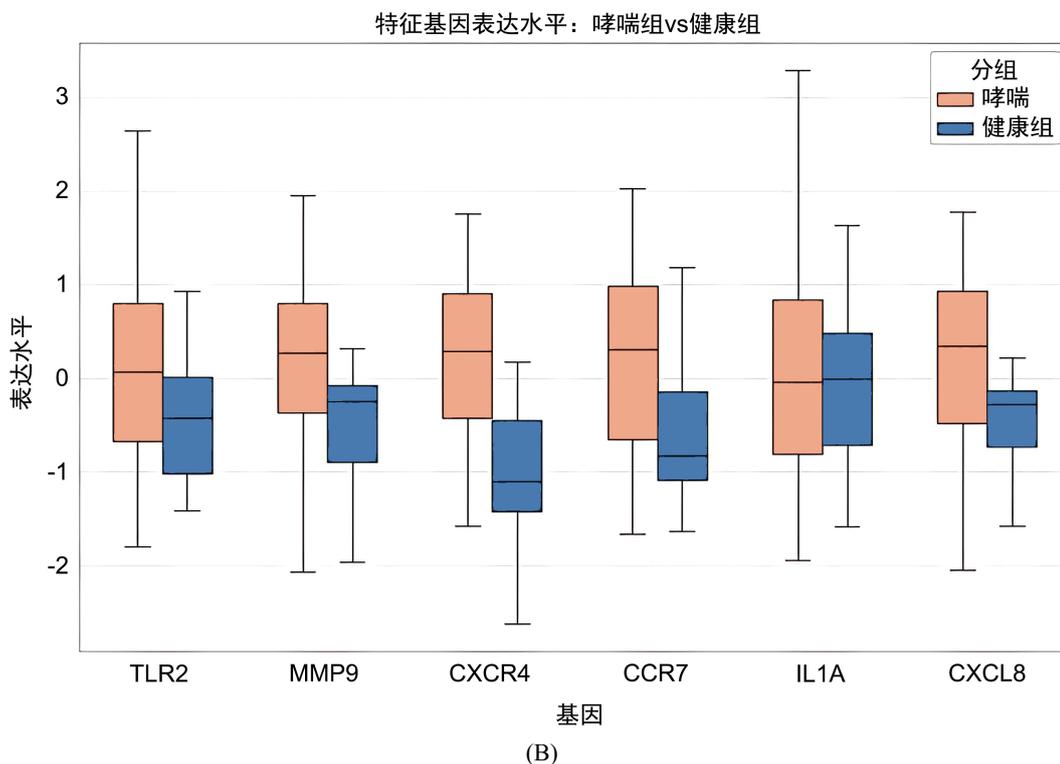


Figure 9. Boxplot of key genes expression
图 9. 关键基因表达量箱线图

Table 5. Statistical *t*-test table
表 5. 统计 *t* 检验表

慢性阻塞性肺疾病基因	t 值	p 值	哮喘基因	t 值	p 值
TLR2	0.25	0.804	TLR2	-2.33	0.026
MMP9	-1.13	0.074	MMP9	-1.53	0.139
CXCR4	-1.91	0.067	CXCR4	-5.79	4.855e-06
CCR7	-0.91	0.072	CCR7	-2.82	0.009
IL1A	1.75	0.088	IL1A	-0.39	0.699
CXCL8	-0.12	0.908	CXCL8	-2.98	0.007

最终结果表明, 对于 COPD 来说, MMP9、CXCR4、CCR7 和 IL1A 可认为有统计上的显著差异。而对于 Asthma 的关键基因来说, TLR2、CXCR4、CCR7 和 CXCL8 可认为有统计上的显著差异。

3.7. 自适应 Lasso 回归筛选基因

在建立医学诊断模型之前, 基因变量之间可能存在高度相关性, 尤其是医学数据中不同生物标志物, 当变量多于样本时, 参数估计会变得极为不稳定, 所以在此种情况下, 使用自适应的 lasso 回归进一步从共同 DEGs 筛选关键基因[6], 通过识别出的少量诊断关键基因用于诊断 COPD 和 Asthma [7]。以下是筛选特征基因的过程见图 10。

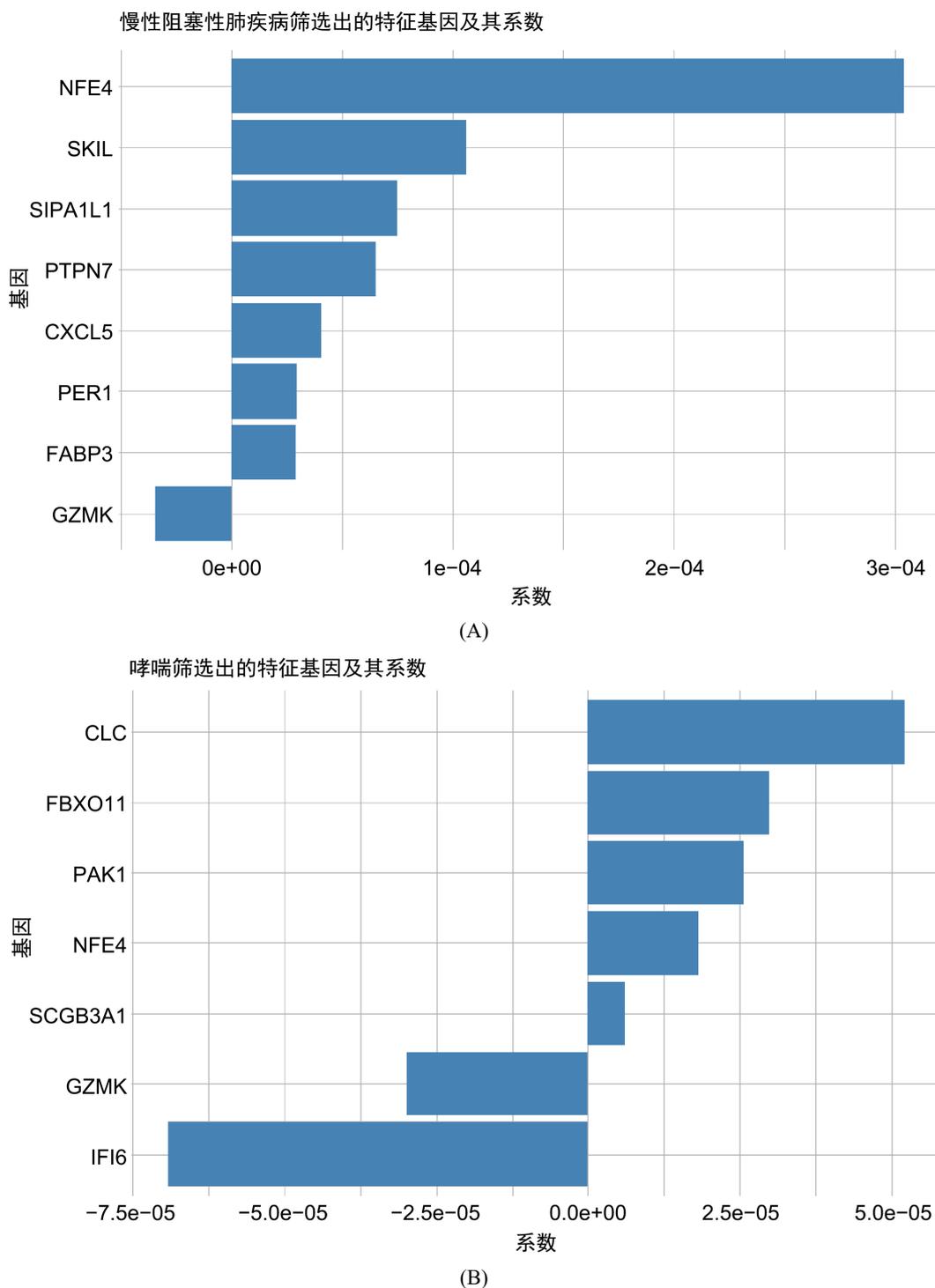


Figure 10. Plot of feature genes and their coefficients
图 10. 特征基因及其系数图

根据自适应 Lasso 回归的结果, 选择非零系数对应的基因, 作为关键特征基因。COPD 筛选的关键特征基因为 PTPN7、NFE4、SKIL、SIPA1L1、CXCL5、PER1、FABP3、GZMK。Asthma 筛选的关键特征基因为 CLC、FBXO11、PAK1、NFE4、SCGB3A1、GZMK、IFI6。

为确定不同疾病基于关键特征基因的诊断模型[6]-[10]。在 GSE148004 数据集上筛选出关键基因后, 使用逻辑回归模型在 GSE212331 数据集上测试 COPD 关键基因诊断效果。在 GSE137268 数据集上测试 Asthma 关键基因的诊断效果。在原始逻辑回归模型的基础上, 本研究采用了逆概率加权的方法, 进行了超参数调优。

以下是原始逻辑回归模型和经过处理后的逻辑回归模型的结果, 见表 6、图 11。

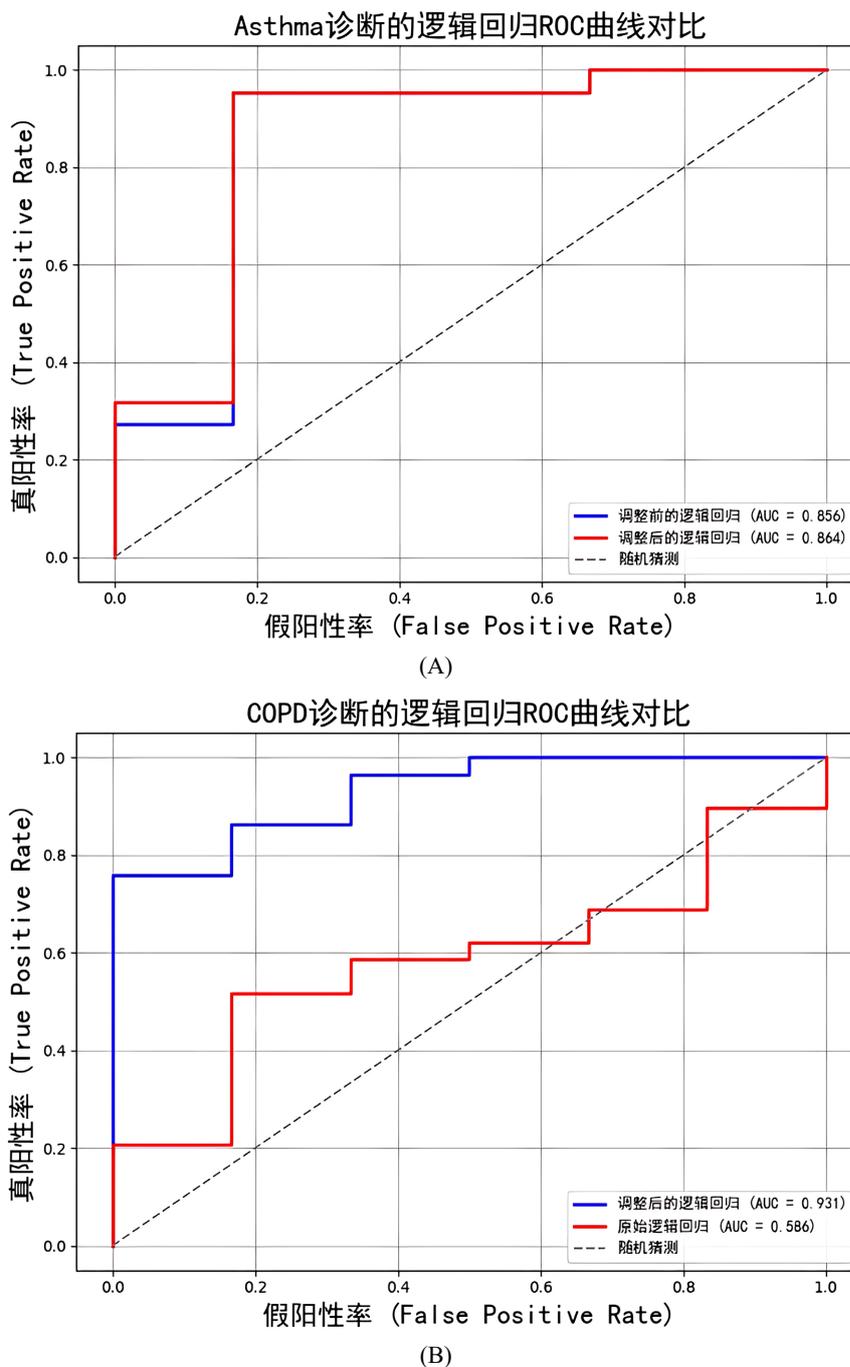


Figure 11. Diagnostic model ROC curve

图 11. 诊断模型 ROC 结果对比图

Table 6. Table of machine learning diagnostic results
表 6. 机器学习诊断结果表

疾病	慢性阻塞性肺疾病			哮喘		
	模型	准确度	灵敏度	特异度	准确度	灵敏度
原始逻辑回归模型	0.771	0.897	0.167	0.857	1.000	0.333
处理后逻辑回归模型	0.886	0.966	0.500	0.929	0.955	0.833

4. 讨论

自 COVID-19 爆发以来, 其对全球人类健康的影响深远。尽管新冠肺炎已得到控制, 但其引发的心血管等并发症依然存在。COVID-19 通过多种机制加剧 COPD 和哮喘的病理过程。首先, COVID-19 引发的细胞因子风暴与 COPD 和哮喘的慢性炎症基础相互作用, 导致更严重的炎症反应和组织损伤。此外, COVID-19 还可能通过调节免疫细胞的功能, 进一步破坏慢性呼吸系统疾病的免疫稳态。前人[11]-[20]对 COVID-19 与多种疾病的关系进行了广泛研究。在慢性呼吸系统疾病方面, 研究了 COPD 恶化的生物标志物, 识别了与不同类型气道炎症相关的临床恶化表型及相关生物标志物, 发现 IL-1 α 在烟雾诱导中性粒细胞炎症中的核心作用及靶向治疗的潜力, 从宿主因子相互作用角度揭示了 COVID-19 与 COPD 的共病机制及地塞米松等的干预作用; 同时分析了哮喘与 COVID-19 的交叉点, 发现关键共宿主因子及潜在治疗药物。这些研究为 COVID-19 相关疾病的诊断、治疗提供了理论依据和新的研究方向。

5. 结论

本研究分析了 COPD 和 Asthma 的基因, 发现 122 个交集基因。在建立蛋白互作网络后发现了 TLR2、MMP9、CXCR4、CCR7、IL1A 和 CXCL8 六个关键基因。研究还建立了基于基因的疾病辅助诊断模型, 对于慢性阻塞性肺疾病来说, 经过逆概率加权和超参数优化后的模型效果远远高于原始逻辑回归模型。对于哮喘来说, AUC 值对于调整前后的模型来说几乎一致, 但调整后的模型在牺牲少许灵敏度的情况下大幅提升了准确度和特异度。本研究揭示了 COVID-19 与 COPD、Asthma 的共同作用机制, 发现了多个关键基因和通路。这些发现不仅为理解 COVID-19 如何加剧慢性呼吸系统疾病提供了重要依据, 还为开发针对 COVID-19 与慢性疾病共病的治疗策略提供了新的方向。

参考文献

- [1] Miro, J., Miroslav, S., Davor, M., *et al.* (2021) Comorbidities and Syndemics in the COVID-19 Age: Challenges and Opportunities for Bringing Separated Branches of Medicine Closer to Each Other. *Psychiatra Danubina*, **33**, 402-413.
- [2] Miro, J. and Ljerka, O. (2013) Comorbidity and Multimorbidity in Medicine Today: Challenges and Opportunities for Bringing Separated Branches of Medicine Closer to Each Other. *Psychiatra Danubina*, **25**, 18-28.
- [3] Barnes, P.J. (2017) Cellular and Molecular Mechanisms of Asthma and COPD. *Clinical Science*, **131**, 1541-1558. <https://doi.org/10.1042/cs20160487>
- [4] Witzleben, V.A., Grages, A., Thomas, J., *et al.* (2024) Immune Checkpoint Expression on Tumor-Infiltrating Lymphocytes (TIL) Is Dependent on HPV Status in Oropharyngeal Carcinoma (OPSCC)—A Single-Cell RNA Sequencing Analysis. *Oral Oncology*, **123**, Article 104251.
- [5] Zou, J., Zhang, W., Chen, X., Su, W. and Yu, D. (2022) Data Mining Reveal the Association between Diabetic Foot Ulcer and Peripheral Artery Disease. *Frontiers in Public Health*, **10**, Article 963426. <https://doi.org/10.3389/fpubh.2022.963426>
- [6] Pei, X., Qi, D., Liu, J., Si, H., Huang, S., Zou, S., *et al.* (2023) Screening Marker Genes of Type 2 Diabetes Mellitus in Mouse Lacrimal Gland by LASSO Regression. *Scientific Reports*, **13**, Article No. 6862. <https://doi.org/10.1038/s41598-023-34072-4>

- [7] Zeng, L. and Chen, Z. (2024) Screening of Genes Characteristic of Pancreatic Cancer by LASSO Regression Combined with Support Vector Machine and Recursive Feature Elimination, and Immune Correlation Analysis. *Journal of International Medical Research*, **52**, 1-10. <https://doi.org/10.1177/03000605241233160>
- [8] Dasgupta, S. (2024) Thinking Beyond Disease Silos: Dysregulated Genes Common in Tuberculosis and Lung Cancer as Identified by Systems Biology and Machine Learning. *OMICS: A Journal of Integrative Biology*, **28**, 347-356. <https://doi.org/10.1089/omi.2024.0116>
- [9] Chen, B., Sun, X., Huang, H., Feng, C., Chen, W. and Wu, D. (2024) An Integrated Machine Learning Framework for Developing and Validating a Diagnostic Model of Major Depressive Disorder Based on Interstitial Cystitis-Related Genes. *Journal of Affective Disorders*, **359**, 22-32. <https://doi.org/10.1016/j.jad.2024.05.061>
- [10] Babichev, S., Yasinska-Damri, L. and Liakh, I. (2023) A Hybrid Model of Cancer Diseases Diagnosis Based on Gene Expression Data with Joint Use of Data Mining Methods and Machine Learning Techniques. *Applied Sciences*, **13**, Article 6022. <https://doi.org/10.3390/app13106022>
- [11] 韩拓, 苑昊, 叶琳樱, 等. COVID-19 与心血管代谢性疾病的临床研究进展[J]. 南方医科大学学报, 2021, 41(4): 628-632.
- [12] Zhang, X., Qin, J., Cheng, X., Shen, L., Zhao, Y., Yuan, Y., *et al.* (2020) In-Hospital Use of Statins Is Associated with a Reduced Risk of Mortality among Individuals with Covid-19. *Cell Metabolism*, **32**, 176-187.E4. <https://doi.org/10.1016/j.cmet.2020.06.015>
- [13] Cheng, X., Liu, Y., Li, H., Zhang, X., Lei, F., Qin, J., *et al.* (2020) Metformin Is Associated with Higher Incidence of Acidosis, but Not Mortality, in Individuals with COVID-19 and Pre-Existing Type 2 Diabetes. *Cell Metabolism*, **32**, 537-547.E3. <https://doi.org/10.1016/j.cmet.2020.08.013>
- [14] Qin, J., Cheng, X., Zhou, F., Lei, F., Akolkar, G., Cai, J., *et al.* (2020) Redefining Cardiac Biomarkers in Predicting Mortality of Inpatients with Covid-19. *Hypertension*, **76**, 1104-1112. <https://doi.org/10.1161/hypertensionaha.120.15528>
- [15] 徐亦鸣, 吕丹丹, 应可净. 2019 冠状病毒病(COVID-19)患者出凝血功能障碍的研究进展[J]. 浙江大学学报(医学版), 2020, 49(3): 340-346.
- [16] 王飞, 刘璐, 薛莹, 等. 新型冠状病毒感染后儿童多系统炎症综合征 4 例[J]. 中国当代儿科杂志, 2023, 25(7): 685-688.
- [17] Bafadhel, M., McKenna, S., Terry, S., Mistry, V., Reid, C., Haldar, P., *et al.* (2011) Acute Exacerbations of Chronic Obstructive Pulmonary Disease. *American Journal of Respiratory and Critical Care Medicine*, **184**, 662-671. <https://doi.org/10.1164/rccm.201104-0597oc>
- [18] Botelho, F.M., Bauer, C.M.T., Finch, D., Nikota, J.K., Zavitz, C.C.J., Kelly, A., *et al.* (2011) IL-1 α /IL-1R1 Expression in Chronic Obstructive Pulmonary Disease and Mechanistic Relevance to Smoke-Induced Neutrophilia in Mice. *PLOS ONE*, **6**, e28457. <https://doi.org/10.1371/journal.pone.0028457>
- [19] Zheng, W., Wang, T., Wu, P., Yan, Q., Liu, C., Wu, H., *et al.* (2021) Host Factor Interaction Networks Identified by Integrative Bioinformatics Analysis Reveals Therapeutic Implications in COPD Patients with Covid-19. *Frontiers in Pharmacology*, **12**, Article 718874. <https://doi.org/10.3389/fphar.2021.718874>
- [20] Jiang, Y., Yan, Q., Liu, C., Peng, C., Zheng, W., Zhuang, H., *et al.* (2022) Insights into Potential Mechanisms of Asthma Patients with COVID-19: A Study Based on the Gene Expression Profiling of Bronchoalveolar Lavage Fluid. *Computers in Biology and Medicine*, **146**, Article 105601. <https://doi.org/10.1016/j.compbiomed.2022.105601>