

引用:袁媛,张睿添,秦秋云,李妞妞,夏猛,李相玲,张萌.肝细胞癌生物标志物识别及靶向中药预测[J].中医导报,2025,31(7):195-203.

## 肝细胞癌生物标志物识别及靶向中药预测\*

袁媛<sup>1</sup>,张睿添<sup>2</sup>,秦秋云<sup>3</sup>,李妞妞<sup>3</sup>,夏猛<sup>3</sup>,李相玲<sup>4</sup>,张萌<sup>3</sup>

(1.广西中医药大学公共卫生与管理学院,广西南宁 530200;

2.广西中医药大学壮医学院,广西南宁 530200;

3.广西中医肝脾藏象重点研究室,广西南宁 530200;

4.广西中医药大学中医药壮瑶医药研究院,广西南宁 530200)

[摘要] 目的:发掘肝细胞癌(HCC)生物标志物,并预测其靶向中药及活性成分,以期为HCC提供新的治疗策略及用药参考。方法:在TCGA(癌症基因组图谱)数据库肝癌数据集中筛选差异基因,通过加权基因共表达网络分析(WGCNA)确定关键模块基因。利用京都基因与基因组百科全书(KEGG)和基因本体(GO)富集分析揭示其生物学功能。使用机器学习算法评估生物标志物的诊断效能、染色体定位及单细胞特征。采用CIBERSORT算法分析免疫细胞浸润情况。利用中药系统药理学数据库和分析平台(TCMSP)预测生物标志物潜在靶向中药及其活性成分。结果:关键模块基因KEGG通路富集在细胞周期、细胞衰老等信号通路。筛选出MDK(中期因子)、RASGRF2(Ras鸟苷酸释放因子2)、CYP2C9(细胞色素P450 2C9酶)、TERT(端粒逆转录酶)、LCAT(卵磷脂胆固醇酰基转移酶)和ADRA2B( $\alpha$ 2B肾上腺素能受体)共6种生物标志物。免疫细胞浸润分析显示,活化记忆CD4<sup>+</sup>T细胞、T滤泡辅助细胞、M0巨噬细胞、静息肥大细胞等在HCC发生机制中与上述标志物密切相关。中药预测结果表明,包括洋金花、延胡索和华山参在内的48种中药及其57种活性成分可能作为潜在治疗手段。结论:MDK、RASGRF2、CYP2C9、TERT、LCAT和ADRA2B通过影响免疫细胞浸润参与HCC发生发展,可作为HCC的生物标志物。中医药及其活性成分能通过调控这些关键基因发挥治疗作用,为HCC提供了新的治疗策略和用药参考。

[关键词] 肝细胞癌;生物信息学;生物标志物;免疫细胞浸润;中药预测

[中图分类号] R273 [文献标识码] A [文章编号] 1672-951X(2025)07-0195-09

DOI:10.13862/j.cn43-1446/r.2025.07.032

### Identification of Hepatocellular Carcinoma Biomarkers and Prediction Analysis of Targeted Traditional Chinese Medicine

YUAN Yuan<sup>1</sup>, ZHANG Ruitian<sup>2</sup>, QIN Qiuyun<sup>3</sup>, LI Niuniu<sup>3</sup>, XIA Meng<sup>3</sup>, LI Xiangling<sup>4</sup>, ZHANG Meng<sup>3</sup>

(1.School of Public Health and Management, Guangxi University of Chinese Medicine, Nanning Guangxi 530200, China; 2.School of Zhuang Medicine, Guangxi University of Chinese Medicine, Nanning Guangxi 530200, China; 3.Key Laboratory of Liver and Spleen Viscera of Guangxi Chinese Medicine, Nanning Guangxi 530200, China; 4.Institute of Traditional Chinese and Zhuang-Yao Medicine, Guangxi University of Chinese Medicine, Nanning Guangxi 530200, China)

[Abstract] Objective: To identify hepatocellular carcinoma (HCC) biomarkers and predict their targeted traditional Chinese medicine (TCM) and active components, aiming to provide new therapeutic strategies and medication references for HCC. Methods: Differential genes were screened from the liver cancer dataset in The Cancer Genome Atlas (TCGA) database, and key module genes were identified through weighted gene co-expression network analysis (WGCNA). Kyoto Encyclopedia of Genes and Genomes (KEGG) and Gene Ontology (GO) enrichment analyses were used to reveal their biological functions. Machine learning algorithms were applied to evaluate the diagnostic efficacy, chromosomal localization, and single-cell characteristics of the biomarkers. The CIBERSORT algorithm was used to analyze immune cell infiltration. The Traditional Chinese Medicine Systems Pharmacology Database and Analysis Platform (TCMSP) was utilized to predict potential targeted TCM and their active components for the biomarkers. Results: The key module gene KEGG pathway is enriched in signaling pathways such as the cell cycle and cell senescence. Six biomarkers were screened out, including MDK (midkine), RASGRF2 (Ras guanine nucleotide release factor 2), CYP2C9 (cytochrome P450 2C9), TERT (telomerase reverse transcriptase),

\*基金项目:广西高校中青年教师科研基础能力提升项目(2024KY0304);广西中医药大学大学生科研训练课题(2024DXS06);国家中医药管理局高水平中医药重点学科(zyzdxk-2023165);国家自然科学基金项目(82460921)

通信作者:张萌,男,讲师,研究方向为中医肝脾藏象

LCAT (lecithin-cholesterol acyltransferase), and ADRA2B (alpha-2B adrenergic receptor). Immune cell infiltration analysis showed that activated memory CD4+ T cells, T follicular helper cells, M0 macrophages, resting mast cells, etc., were closely related to the above biomarkers in the pathogenesis of HCC. TCM prediction results indicated that 48 kinds of TCM, including Yangjinhua (Daturae Flos), Yanhusuo (Corydalis Rhizoma), and Huashanshen (Physochlainae Radix), along with their 57 active components, might serve as potential therapeutic agents. Conclusion: MDK, RASGRF2, CYP2C9, TERT, LCAT, and ADRA2B are involved in the occurrence and development of HCC by affecting immune cell infiltration and can be used as biomarkers for HCC. TCM and their active components can exert therapeutic effects by regulating these key genes, providing new therapeutic strategies and medication references for HCC.

[Keywords] hepatocellular carcinoma; bioinformatics; biomarkers; immune cell infiltration; traditional Chinese medicine prediction

肝细胞癌(hepatocellular carcinoma, HCC)是一种以高度侵袭性和转移潜能为特征的复杂恶性肿瘤,是全球范围内最常见的原发性肝癌<sup>[1]</sup>。HCC主要由慢性肝病发展而来,特别是病毒性肝炎(如乙型和丙型肝炎)、酒精性肝病及代谢相关性脂肪肝导致的肝硬化<sup>[2]</sup>。我国作为HCC高风险地区,肝癌的发病率和死亡率在恶性肿瘤中分别排在第5位和第2位<sup>[3]</sup>。大量肝癌患者的治疗带来了严重的经济和社会负担<sup>[4]</sup>。因此,深入探究HCC发病机制,寻找新的治疗策略和生物标志物显得尤为重要。随着高通量测序技术的发展,我们能够更深入地理解疾病的分子生物学基础<sup>[5]</sup>。本研究旨在利用加权基因共表达网络分析(weighted correlation network analysis, WGCNA)等分析方法全面挖掘HCC的基因表达数据,以期鉴定出可靠的生物标志物。通过CIBERSORT算法探讨HCC的免疫浸润特征,揭示生物标志物与免疫微环境之间的关联。此外,将生物标志物视为靶点,预测与之相关的中药及活性成分,为HCC的诊断和治疗提供新的视角。

## 1 方法

工作流程见图1。

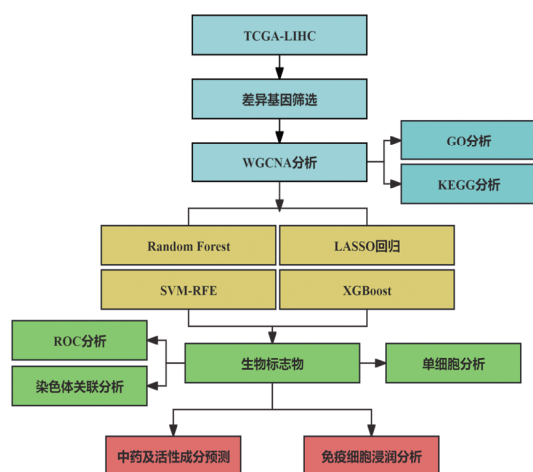


图1 工作流程图

1.1 差异表达基因(DEGs)筛选 从TCGA(癌症基因组图谱)数据库(<https://portal.gdc.cancer.gov/>)中下载肝细胞癌数据集。数据包含352例肿瘤组织和49例正常组织的mRNA测序数据及对应患者的年龄、性别、生存时间、生存状态、分级与分期等临床信息。运用R语言limma包筛选肿瘤组织和正常组织之间的差异表达基因,标准定义为对数倍数变化的绝对值 $|\log_2$

$FC| > 1.5$ 且校正后 $P < 0.05$ 。火山图绘制由R语言tinyarray包实现。

1.2 WGCNA分析 WGCNA用于构建和分析基因共表达网络,旨在揭示基因与特定生物表型之间的关联<sup>[6]</sup>。计算基因间的加权相关性,通过基因之间的相关系数构建层次聚类树处理离群样本。使用pick Soft Threshold函数计算相关系数的平方和基因的平均连接度,以确定最佳软阈值。拓扑重叠矩阵转换,计算相应不相似度,并进行层次聚类分析。通过基因显著性(GS)值和模块成员(MM)值来衡量基因模块与临床信息的关联情况,并识别关键模块<sup>[7]</sup>。

1.3 富集分析 使用clusterProfiler R包对WGCNA识别的关键模块基因集进行基因本体(GO)和京都基因与基因组百科全书(KEGG)富集分析。以 $P < 0.05$ 进行筛选,以确保富集分析结果的可靠性和生物学意义。

1.4 基于机器学习的关键基因筛选与分析 鉴于WGCNA表达数据的复杂性和高维性,采用随机森林(random forest, RF)、支持向量机-递归特征消除(support vector machine recursive feature elimination, SVM-RFE)、LASSO回归(lasso regression)、极端梯度提升(extreme gradient boosting, XGBoost)4种算法进一步筛选特征基因。RF是一种基于决策树构建的监督机器学习算法<sup>[8]</sup>。利用R的randomForest包,设置500棵决策树,筛选最小误差率,基于Gini指数均值获取基因重要性得分,选取得分排名前30位的基因作为候选基因。LASSO模型是一种用于特征选择和回归分析的线性模型,通过引入惩罚函数,缩减冗余变量系数,结合10倍交叉验证选取最佳 $\lambda$ 值防止过度拟合,以筛选特定变量<sup>[9]</sup>。使用R的glmnet包实现。SVM-RFE通过递归移除特征,并使用SVM模型评估特征重要性,以确定最优特征子集<sup>[10]</sup>。通过R的e1071等包分析,结合交叉验证评估SVM-RFE模型,并确定最佳特征数。XGBoost通过增加正则化项,优化后的梯度提升决策树算法<sup>[11]</sup>。借助R的xgboost等包,获取XGBoost模型的特征重要性,筛选重要性大于等于1的特征基因,以柱状图表示。利用韦恩图取4种算法交集,获取HCC相关核心靶点。以受试者工作特征(ROC)曲线验证上述关键基因的诊断效能,一般认为,曲线下面积(area under curve, AUC)大于0.70提示具有临床意义。采用R包circlize展示核心基因染色体关联特征。

1.5 单细胞分析 HCC单细胞转录组数据来源于GEO(<https://www.ncbi.nlm.nih.gov/geo/>)的GSE146115数据集。利用R的Seurat软件包对单细胞RNA测序数据进行标准预处理,质量

控制条件:核糖体基因比例低于20%,每个细胞中检测到的基因数量在200至10 000之间。应用Find Neighbors、Find Clusters和Run UMAP函数对细胞亚群进行无偏聚类分析,使用SingleR软件包对细胞类型进行注释。使用FeaturePlot函数展示目标基因在细胞亚型中的表达分布情况。

1.6 免疫浸润分析 通过R的CIBERSORT法比较两组间免疫细胞的相对比例,并计算各种免疫细胞之间的相关性。借助corplot和ggplot2包可视化结果。

1.7 基于核心基因的中药与活性化合物预测 基于核心基因,利用TCMSP数据库(<https://old.tcm-sp-e.com/>)反向筛选其对应的中药,以口服生物利用度(OB)≥30%且药物相似性(DL)≥0.18为评价标准<sup>[2]</sup>,进一步筛选中药对应的活性化合物。构建网络图反映中药、活性化合物、靶点之间的对应关系。使用雷达图展示中药的性味、归经特征。

## 2 结果

2.1 差异表达基因结果 获取两组间DEGs共2 563个,火山图显示了2 034个上调基因,529个下调基因。(见图2)

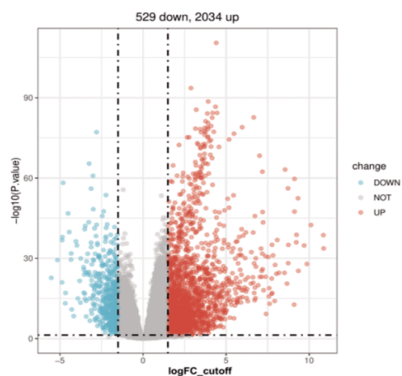
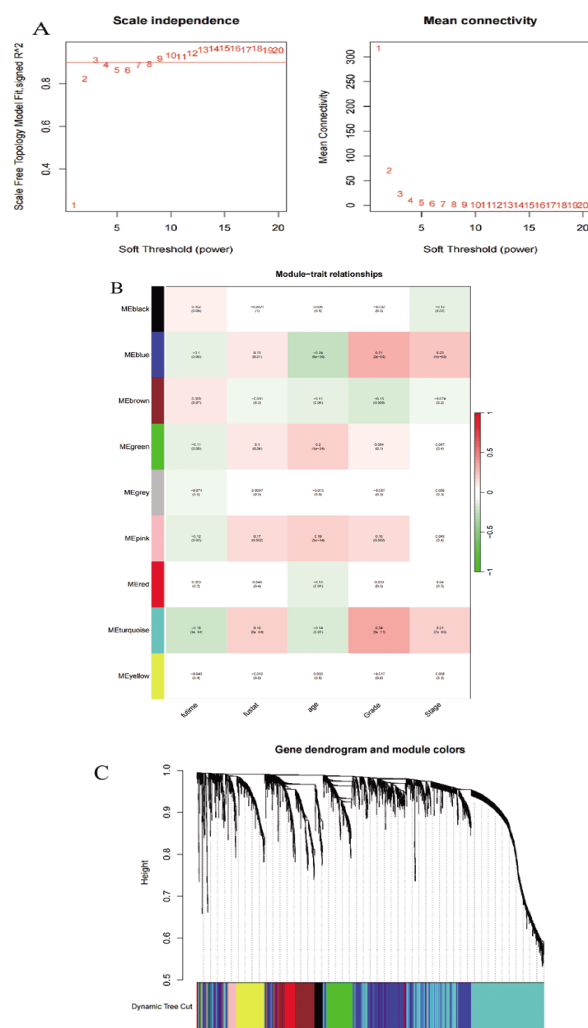


图2 DEGs 火山图

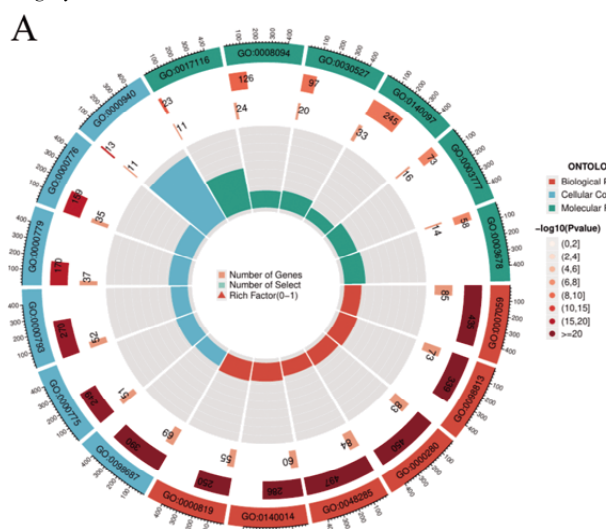
2.2 WGCNA分析 对DEGs进行WGCNA分析。经pick Soft Threshold计算,确定软阈值=3时,平均连通性降至最低。(见图3A)采用动态剪切树法,聚类分析将DEGs分为9个模块,使用不同颜色进行标注,其中turquoise模块基因数量最多,共计994个,grey模块基因数量最少,含11个。各模块基因与临床特

征作相关性分析,发现turquoise模块与futime、fustat的相关性均有统计学意义,故选择turquoise模块进行后续分析。(见图3B、3C)



注:A.软阈值与平均连通度;B.共表达模块相关性;C.聚类树状图。

图3 WGCNA 分析结果



注:A.GO富集分析;B.KEGG富集分析。

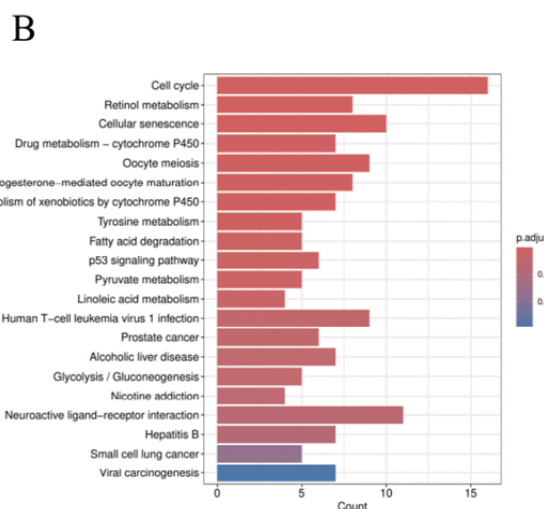


图4 关键模块功能富集分析

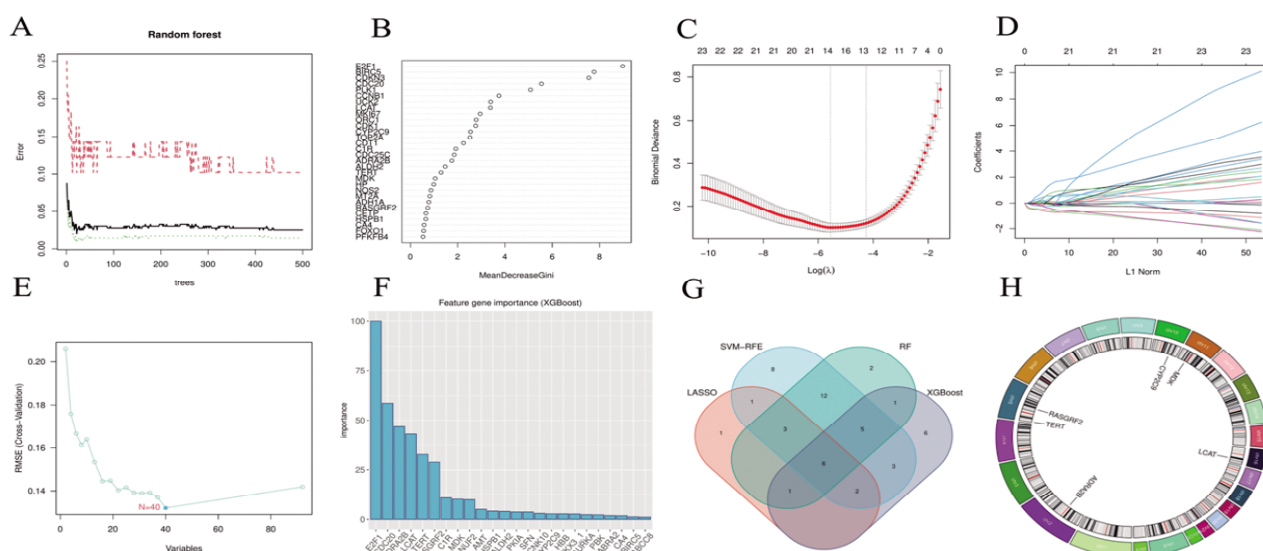
**2.3 富集分析** GO富集分析结果包括生物过程(BP)、分子功能(MF)和细胞组分(CC)共294个条目。其中,BP涉及染色体分离、细胞核分裂、细胞器分裂等生物学过程;MF集中于单链DNA解旋酶活性、ATP活性、染色质的结构成分等功能;CC富集于染色体区域、着丝粒区域、凝聚染色体等组分。图4A显示了BP、MF、CC排名前6的富集结果。如图4B,基因主要富集在细胞周期、视黄醇代谢、细胞衰老、药物代谢-细胞色素P450、细胞色素P450对外源物质的代谢作用、酪氨酸代谢、脂肪酸降解、P53信号通路等信号通路。

**2.4 机器学习模型的构建** 基于turquoise模块,建立4种机器学习模型以进一步识别HCC潜在的生物标志物。RF获取30个预测基因(见图5A、5B)。LASSO回归筛选出14个预测基因(见图5C、5D)。SVM获取排名前30的预测基因(见图5E)。XGBoost筛选出24个预测基因(见图5F)。绘制韦恩图得到6个交集基因,分别为MDK(中期因子)、RASGRF2(Ras鸟苷酸释放因子2)、CYP2C9(细胞色素P450C9酶)、TERT(端粒逆转录酶)、LCAT(卵磷脂胆固醇酰基转移酶)和ADRA2B( $\alpha$ 2B肾上腺素能受体)(见图5G)。上述6个基因分别定位于chr11、chr5、chr10、

chr5、chr16、chr2染色体(见图5H)。ROC分析显示,MDK、RASGRF2、CYP2C9、TERT、LCAT和ADRA2B的AUC值分别为0.923、0.910、0.934、0.947、0.970和0.954,表明6种标志物均具有明确的诊断价值。(见图6)

**2.5 单细胞分析** 基于GSE146115数据集,根据质量控制条件(见图7A)筛选出19 854个细胞。对数据进行UMAP降维和细胞聚类后,共获取17个初始亚群(见图7B)。SingleR共注释出6种细胞亚群,分别是B细胞、单核细胞、T细胞、CD8<sup>+</sup>T细胞、祖细胞和CD4<sup>+</sup>T细胞(见图7C)。在各细胞亚群中对MDK、RASGRF2、CYP2C9、TERT、LCAT、ADRA2B表达分布进行分析发现,MDK在各类细胞中均显著表达;CYP2C9在单核细胞、T细胞、CD8<sup>+</sup>T细胞和祖细胞中有较强表达;RASGRF2、TERT主要在单核细胞中有标注;LCAT在各类细胞中均有表达;ADRA2B在各类细胞中表达较低。(见图7D)

**2.6 免疫细胞浸润分析** 免疫细胞浸润分析发现,组间存在11种免疫细胞的明显差异( $P<0.05$ )。与正常组相比,HCC样本中浆细胞、T调节细胞、静息NK细胞、单核细胞、M0巨噬细胞、M2巨噬细胞、静息树突状细胞、静息肥大细胞、活肥大细胞



注:A、B.RF结果;C、D.Lasso结果;E.SVM结果;F.XGBoost结果;G.四种算法结果韦恩图;H.生物标志物染色体定位。

图5 4种机器学习算法筛选生物标志物

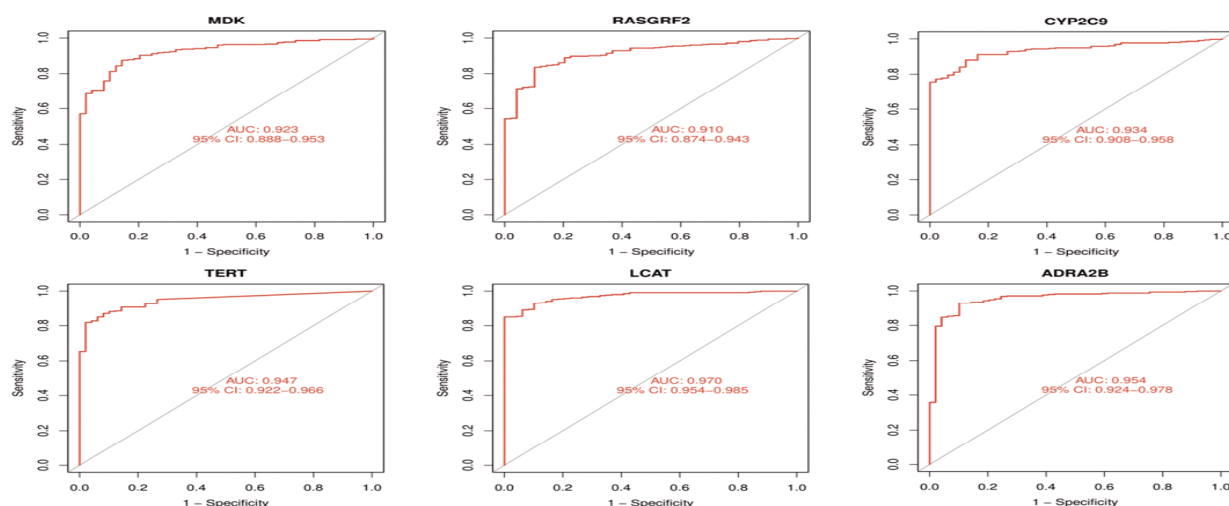
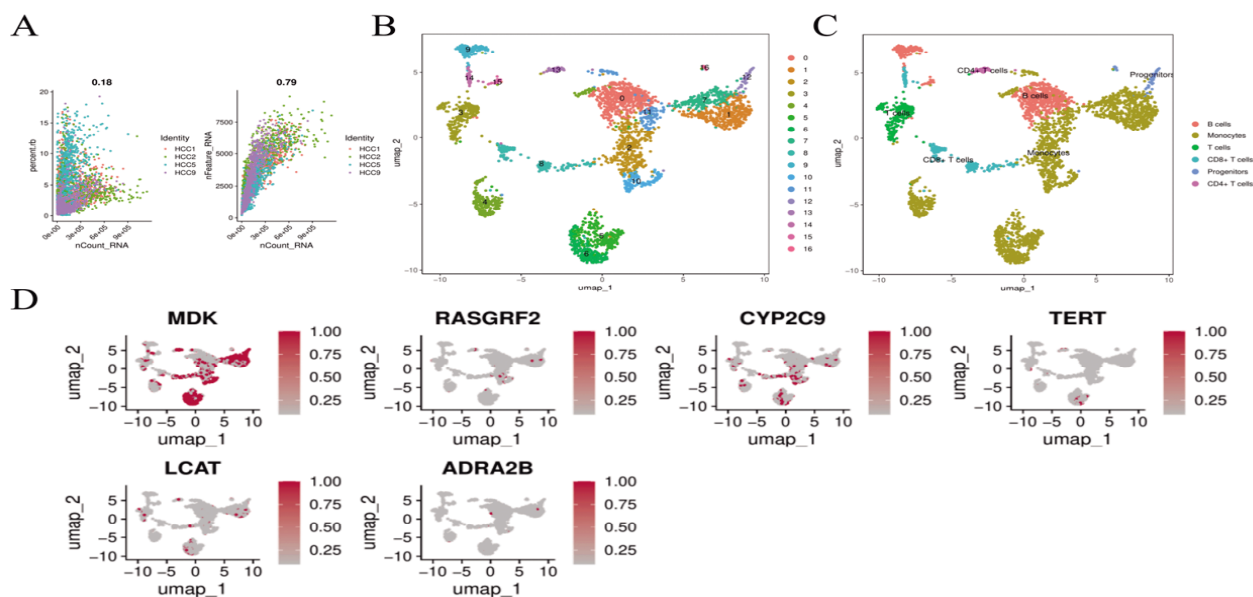


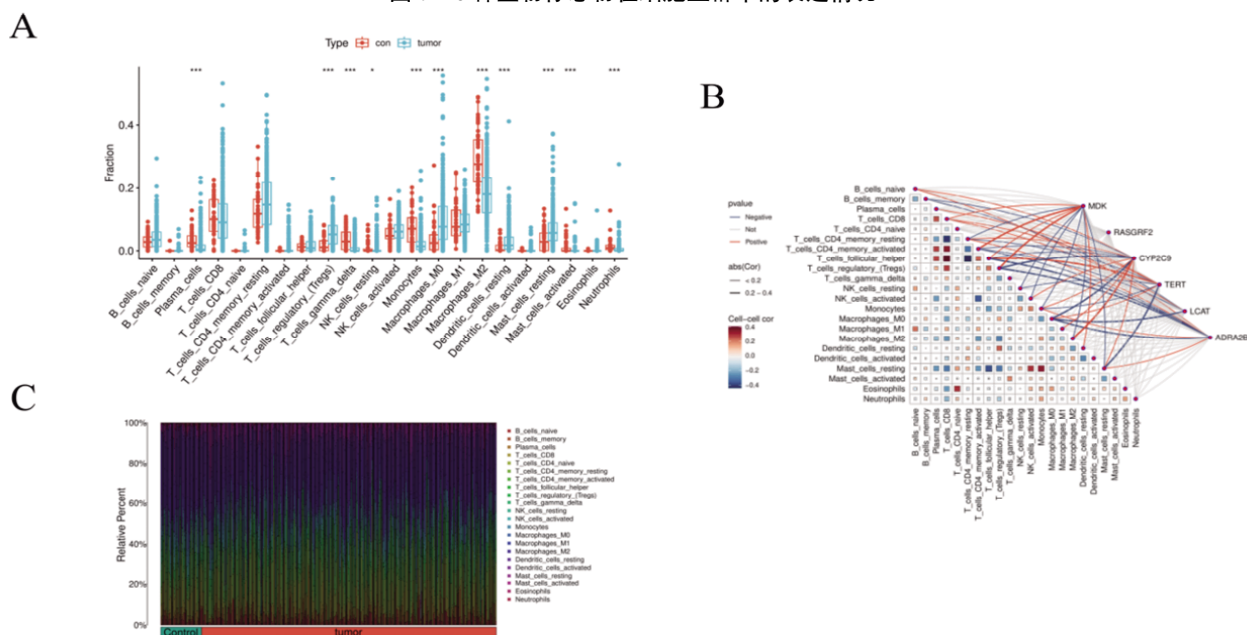
图6 诊断效能的ROC曲线分析





注:A.数据质量控制;B.单细胞聚类;C.细胞注释;D.生物标志物表达水平。

图7 6种生物标志物在细胞亚群中的表达情况



注:A.免疫细胞丰度箱线图;B.免疫细胞相关性热图;C.免疫细胞丰度比例堆积图。

图8 免疫细胞浸润分析

胞、中性粒细胞呈现出显著高表达,而 $\gamma\delta$ T细胞表现出显著低表达。(见图8A和图8C)相关分析发现,MDK与活化记忆CD4<sup>+</sup>T细胞、T调节细胞、T滤泡辅助细胞、CD8<sup>+</sup>T细胞、静息树突状细胞呈正相关,与静息记忆CD4<sup>+</sup>T细胞、静息NK细胞、活化NK细胞、单核细胞、M2巨噬细胞、静息肥大细胞呈负相关。CYP2C9与M2巨噬细胞、幼稚B细胞、静息记忆CD4<sup>+</sup>T细胞、活化NK细胞、单核细胞、M1巨噬细胞、静息肥大细胞呈正相关,与T调节细胞、M0巨噬细胞、记忆B细胞、活化记忆CD4<sup>+</sup>T细胞、T滤泡辅助细胞呈负相关。TERT与T滤泡辅助细胞、浆细胞、CD8<sup>+</sup>T细胞呈正相关,与静息肥大细胞、单核细胞、静息NK细胞、静息记忆CD4<sup>+</sup>T细胞呈负相关。LCAT与CD8<sup>+</sup>T细胞呈正相关,与M0巨噬细胞、记忆B细胞、活化记忆CD4<sup>+</sup>T细胞呈负

相关。ADRA2B与幼稚B细胞、静息记忆CD4<sup>+</sup>T细胞、静息肥大细胞呈正相关,与T滤泡辅助细胞、T调节细胞、M0巨噬细胞、活化记忆CD4<sup>+</sup>T细胞呈负相关。(见图8B)

2.7 中药及其活性成分预测结果 基于上述生物标志物,参考OB $\geq 30\%$ 且DL $\geq 0.18$ 标准,检索TCMSP数据库,共筛选出靶向MDK、RASGRF2、CYP2C9、ADRA2B的中药48味,对应活性成分57个。“生物标志物-活性成分-中药”的网络关系见图9,代表中药的节点颜色越深,表示相关活性物质越多。参考《中华人民共和国药典》《中药大辞典》,将中药性味、归经和功效总结于表1。活性成分的详细信息见表2。以雷达图展示中药性味、归经可以发现,靶向中药的性味以性寒、味苦为主。归经以肝经、肺经为主。(见图10)

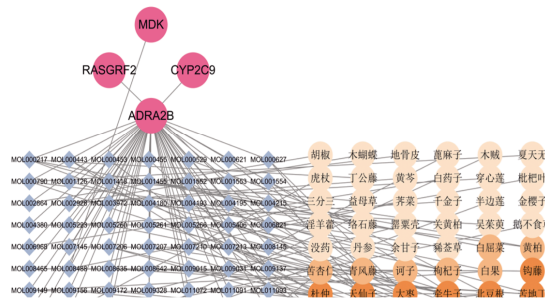
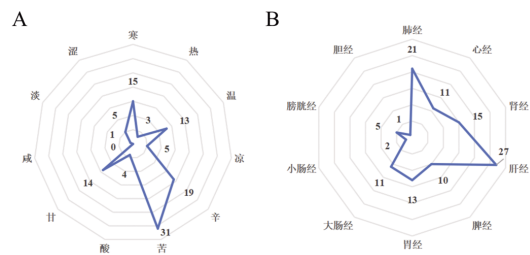


图 9 “生物标志物-成分-中药”网络图



注:A.中药性味雷达图;B.中药归经雷达图。

图 10 靶向中药的性味与归经分析

表 1 中药预测表

中药	性味	归经	功效	中药	性味	归经	功效
洋金花	性温,味辛	归肺经、肝经	平喘止咳、镇痛、解痉	夏天无	性温,味苦、辛	归肝经	活血、通络、行气止痛
延胡索	性温,味辛、苦	归肝经、脾经	活血散瘀、理气止痛	稀莪草	性寒,味辛、苦	归肝经、肾经	祛风湿、利关节、解毒
华山参	性热,味甘、苦	归肺经、心经	平喘止咳、安神镇惊	吴茱萸	性热,味辛、苦	归肝经、脾经、胃经、肾经	散寒止痛、降逆止呕、助阳止泻
莲子心	性寒,味苦	归肾经、心经	清心安神、交通心肾、涩精止血	三分三	性温,味苦、辛	-	解痉镇痛、祛风除湿
荷叶	性平,味苦	归肝经、脾经、胃经	清暑化湿、升发清阳、凉血止血	千金子	性温,味辛	归肝经、肾经、大肠经	逐水消肿、破血消癥
天仙子	性温,味苦、辛	归心经、肝经、胃经	解痉、止痛、安神、定喘	枇杷叶	性寒,味苦	归肺经、胃经	清肺止咳、降逆止呕
牵牛子	性寒,味苦	归肺经、肾经、大肠经	泻水通便、消痰涤饮、杀虫攻积	木贼	性平,味甘、苦	归肺经、肝经	疏风散热、退翳、止血
苦地丁	性寒,味苦	归心经、肝经、大肠经	清热解毒、散结消肿	木蝴蝶	性凉,味苦、甘	归肺经、肝经、胃经	清肺利咽、疏肝和胃
海桐皮	性平,味苦、辛	归肝经、脾经	祛风湿、通经络、止痛	没药	性平,味辛、苦	归心经、肝经、脾经	散瘀定痛、消肿生肌
钩藤	性凉,味甘	归肝经、心经	清热平肝、息风定惊	络石藤	性寒,味苦	归心经、肝经、肾经	祛风通络、活血消肿
杜仲	性温,味甘	归肝经、肾经	补肝肾、强筋骨、安胎	金樱子	性平,味酸、甘、涩	归肾经、膀胱经、大肠经	固精、缩尿、涩肠、止泻
大枣	性温,味甘	归脾经、胃经、心经	补中益气、养血安神	荠菜	性凉,味甘、淡	归肝经、脾经、膀胱经	凉肝止血、平肝明目、清热利湿
北豆根	性寒,味苦	归肺经、胃经、大肠经	清热解毒、消肿止痛	黄芩	性寒,味苦	归肺经、胆经、脾经、大肠经、小肠经	清热燥湿、泻火解毒、止血、安胎
青风藤	性平,味苦、辛	归肝经、脾经	祛风湿、通经络、利小便	虎杖	性寒,味苦	归肝经、胆经、肺经	祛风利湿、散瘀止痛、止咳化痰
苦杏仁	性温,味苦	归肺经、大肠经	降气、止咳平喘、润肠通便	胡椒	性热,味辛	归大肠经、胃经	温中散寒、下气、消痰
黄柏	性寒,味苦	归肾经、膀胱经	清热燥湿、泻火除蒸、解毒疗疮	关黄柏	性寒,味苦	归肾经、膀胱经	清热燥湿、泻火除蒸、解毒疗疮
诃子	性平,味苦、酸、涩	归肺经、大肠经	涩肠敛肺、降水利咽	鹅不食草	性温,味辛	归肺经、肝经	通鼻窍、止咳
枸杞子	性平,味甘	归肝经、肾经	滋补肝肾、益精明目	丁公藤	性温,味辛	归胃经、脾经、肝经	祛风除湿、消肿止痛
白屈菜	性温,味苦	归肺经、胃经	破瘀止血、解痉止痛、止咳平喘	地骨皮	性寒,味甘	归肺经、肝经、肾经	凉血除蒸、清肺降火
白果	性平,味甘、苦、涩	归肺经、肾经	敛肺定喘、止带浊、缩小便	丹参	性寒,味苦	归心经、肝经	祛瘀止痛、活血通经、清心除烦
余甘子	性凉,味甘、酸、涩	归肺经、胃经	清热凉血、消食健胃、生津止咳	穿心莲	性寒,味苦	归肺经、胃经、肝经	清热解毒、凉血、消肿
罂粟壳	性平,味酸、涩	归肺经、大肠经、肾经	敛肺、涩肠、止痛	鹿麻子	性平,味甘、辛	归大肠经、肺经	消肿拔毒、泻下通滞
淫羊藿	性温,味辛、甘	归肝经、肾经	补肾阳、强筋骨、祛风湿	半边莲	性平,味辛	归心经、小肠经、肺经	利尿消肿、清热解毒
益母草	性寒,味苦、辛	归肝经、心经、膀胱经	活血调经、利尿消肿	白药子	性凉,味苦、辛	归肺经、胃经	清热解毒、祛风止痛、凉血止血

表 2 活性成分表

Mol ID	分子名	度值	OB/%	DL
MOL005406	atropine	5	45.97	0.19
MOL000217	(S)-Scoulerine	5	32.28	0.54
MOL000790	Isocorypalmine	4	35.77	0.59
MOL001554	Scopolamine	4	67.97	0.27
MOL008488	rohimbine	4	46.42	0.81
MOL007213	Nuciferin	4	34.43	0.40
MOL002928	Oroxylin A	4	41.37	0.23
MOL001455	(S)-Canadine	4	53.83	0.77
MOL006821	(-)-epigallocatechin-3-gallate	4	55.09	0.77
MOL011093	apohyoscine	3	59.68	0.25
MOL007207	Machiline	3	79.64	0.24
MOL007206	Armepavine	3	69.31	0.29
MOL008635	corynicine	3	38.27	0.81

续表2:

Mol ID	分子名	度值	OB/%	DL
MOL000443	Erythraline	3	49.18	0.55
MOL001552	OIN	3	45.97	0.19
MOL000627	Stepholidine	3	33.11	0.54
MOL009149	Cheilanthifoline	3	46.51	0.72
MOL012922	1-SPD	3	87.35	0.54
MOL012921	stepharine	3	51.55	0.37
MOL009328	5-[[[(1S)-6,7-dimethoxy-2-methyl-3,4-dihydro-1H-isoquinolin-1-yl] methyl]-2-methoxyphenol	2	51.55	0.37
MOL004380	C-Homoerythrinan, 1,6-didehydro-3,15,16-trimethoxy-, (3.β.)-	2	39.14	0.49
MOL001418	galeopsin	2	61.02	0.38
MOL011519	hyoscine	2	49.84	0.27
MOL011491	Datumetine	2	84.74	0.18
MOL004215	leonticine	2	45.79	0.26
MOL004195	CORYDALINE	2	65.84	0.68
MOL004193	Clarkeanidine	2	86.65	0.54
MOL004180	Coronaridine	2	34.97	0.68
MOL003972	1-methyl-2-nonyl-4-quinolone	2	48.42	0.20
MOL001553	atroscine	2	71.93	0.27
MOL000621	16-epi-Isositsirikine	2	49.52	0.59
MOL005266	AGROCLAVIN	2	47.71	0.24
MOL005261	LYSERGOL	2	48.11	0.27
MOL005260	CHANOCLAVINE	2	62.64	0.18
MOL005223	DL-Laudanine	2	69.09	0.37
MOL011949	Allopseudococain	2	34.93	0.23
MOL001126	[(5aS,8aR,9R)-8-oxo-9-(3,4,5-trimethoxyphenyl)-5,5a,6,9-tetrahydroisobenzofurano[6,5-f][1,3]benzodioxol-8a-yl] acetate	2	44.08	0.90
MOL000529	Ibogain	2	50.52	0.63
MOL009172	Pronuciferin	2	32.75	0.37
MOL009156	4'-methyl-N-methylcocaurine	2	55.35	0.26
MOL008642	noryuziphine	2	51.93	0.20
MOL011091	[(1S,3S,5S,7S)-7-hydroxy-8-methyl-8-azabicyclo[3.2.1]octan-3-yl] (2S)-3-hydroxy-2-phenylpropanoate	2	60.94	0.23
MOL013287	Physovenine	2	106.21	0.19
MOL002864	(2E,4Z)-5-(1,3-benzodioxol-5-yl)-1-piperidino-penta-2,4-dien-1-one	2	37.52	0.23
MOL007210	o-Normuciferine	2	33.52	0.36
MOL009137	(R)-(6-methoxy-4-quinolyl)-[(2R,4R,5S)-5-vinylquinuclidin-2-yl] methanol	2	55.88	0.40
MOL000455	Erysotrine	2	113.82	0.44
MOL000453	1,2,6,7-Tetradehydro-3,15-dimethoxyerythrinan-16-ol	2	37.31	0.39
MOL008465	(E)-16,17-Didehydro-17-methoxy-17,18-seco-3-β-yohimban-16-carboxylic acid methyl ester	2	32.75	0.64
MOL011704	1-[(Z)-2-(3,5-dimethoxyphenyl)vinyl]-3,5-dimethoxybenzene	2	41.69	0.21
MOL009031	Cinchonan-9-al, 6'-methoxy-, (9R)-	2	68.22	0.40
MOL009015	(-)-Tabernemontanine	2	58.67	0.61
MOL007145	salviolone	2	31.72	0.24
MOL008145	Costaclavine	2	46.21	0.24
MOL012216	norlobelanine	2	64.08	0.30
MOL006968	(S)-(6-methoxy-4-quinolyl)-[(2S,4S,5S)-5-vinylquinuclidin-2-yl] methanol	2	38.06	0.40
MOL011072	Quinicine	2	75.44	0.33

### 3 讨 论

HCC是一种高度恶性的肿瘤,肝切除和肝移植是HCC早期治疗的主要手段,但适用范围有限。中晚期患者主要采用化疗,并辅以分子靶向和免疫治疗药物,这些治疗虽有一定疗效,但因肿瘤细胞异质性,不同患者的治疗效果差异较大<sup>[13]</sup>。中医药因其良好的疗效和提高生活质量的优点,已被纳入《原发性肝癌诊疗指南(2024年版)》,成为HCC治疗的重要组成部分<sup>[1]</sup>。因此,鉴于HCC治疗手段的局限性,深入研究HCC的发病机制,并探索新的生物标志物至关重要。通过生物标志

物的筛选,不仅能提高HCC早期诊断的准确性,还能为个性化治疗提供依据<sup>[14]</sup>。结合生物信息学方法鉴定出的生物标志物,预测其与中药及其有效成分之间的关联,将有助于发现具有潜在治疗价值的中药,从而为HCC的个性化治疗开辟新途径<sup>[15]</sup>。

MDK是一种多功能分泌蛋白,可作为生长因子调节多种信号通路,促进细胞增殖、存活、粘附、迁移和血管生成等,在HCC中显著高表达<sup>[16]</sup>。有研究发现,MDK可以通过Wnt信号传导与β-catenin蛋白相互作用从而诱导上皮间充质转化<sup>[17]</sup>,且

在肿瘤耐药性中扮演了重要角色<sup>[18]</sup>,反映出MDK不仅是HCC早期诊断的标志物,也为肿瘤靶向治疗提供了新视角。RASGRF2属于Ras鸟嘌呤核苷酸交换因子家族成员,是Ras信号通路中的重要调控因子。在结直肠癌中,RASGRF2通过Src/蛋白激酶B(Akt)/核因子 $\kappa$ B(NF- $\kappa$ B)途径调节基质金属蛋白酶-9(MMP-9),促进癌细胞的迁移和侵袭<sup>[19]</sup>。有临床研究<sup>[20]</sup>表明,RASGRF2基因的遗传变异与HCC患者生存率显著相关。目前,关于RASGRF2在肝癌发病机制中的具体作用有待进一步探究。CYP2C9是细胞色素P450的家族成员,细胞色素P450酶系在肝脏中广泛表达,主要参与药物、毒素、激素等物质的代谢过程。CYP2C9在HCC中的表达显著下调,张拓等<sup>[21]</sup>通体外实验证实CYP2C9过表达显著抑制HepG2细胞和MHCC97H细胞的迁移、侵袭能力和EMT过程,是HCC的一种保护性因素。然而,也有研究提出,除了在药物代谢中的作用外,细胞色素P450酶系还参与炎症反应和氧化应激途径。因此,该酶系基因表达的增加可能反映了机体代谢状态的失调,并可能与肿瘤进展有关<sup>[22]</sup>。TERT基因编码端粒酶逆转录酶,是维持端粒长度的关键酶。TERT异常激活,维持端粒长度和基因组稳定性,促进癌细胞的无限增殖,已成为多种癌症预后不良的重要标志物<sup>[23]</sup>。已有研究<sup>[24]</sup>证实,抑制TERT表达可抑制乙肝病毒X蛋白诱导的肝癌发生。LCAT即卵磷脂胆固醇酰基转移酶,主要参与胆固醇逆向转运和高密度脂蛋白(HDL-C)代谢,在HCC肿瘤组织中呈现显著低表达和DNA的高度甲基化<sup>[25]</sup>。近期研究发现,雌激素能够激活LCAT/HDL-C轴,维持胆固醇稳态,进而抑制肝癌发生,表明了靶向胆固醇稳态作为治疗HCC策略的潜在价值<sup>[26]</sup>。ADRA2B是编码 $\alpha$ 2B肾上腺素能受体的基因,有研究<sup>[27]</sup>通过Lasso-Cox回归构建风险预后模型,发现ADRA2B是HCC总生存期预后的重要生物标志物之一,与本研究结果一致。目前,ADRA2B与肿瘤的相关研究开展较少。

结合正常组与肿瘤组免疫细胞浸润与特征基因的相关性分析结果,可以推断MDK、CYP2C9、TERT、LCAT、ADRA2B可能通过调节活化记忆CD4<sup>+</sup>T细胞、T调节细胞、T滤泡辅助细胞、M0巨噬细胞、M2巨噬细胞、静息肥大细胞等特异性免疫细胞参与HCC的发生发展,但其具体调控机制还有待进一步论证。

中医理论考虑分证型、分阶段辨治肝癌<sup>[28]</sup>。早中期肝癌着眼于正气未衰,癌毒内盛,治疗原则重在攻癌,常用疏肝行气、活血消癥、清热利湿解毒等治法,预测中药中多数药物具有上述功效。如:延胡索、夏天无、木蝴蝶均能够疏肝行气,适用于肝气郁结型肝癌;白屈菜、千金子、没药等具有活血消癥的功效,适用于气滞血瘀型肝癌;苦地丁、北豆根、黄柏、豨莶草、黄芩、穿心莲、半边莲等可清热解毒,荷叶、海桐皮、青风藤、三分三、虎杖等可祛湿通络,均适用于湿热毒蕴型肝癌。中晚期肝癌着眼于癌毒炽盛、正气亏虚,治疗原则为攻补兼施,补法以补益肝肾、益气养阴较为常用,如杜仲、大枣、枸杞子、余甘子、淫羊藿、金樱子等。靶向中药中,洋金花、天仙子、牵牛子、白屈菜、罂粟壳、千金子、白药子等在《中华人民共和国药典》中被明确标注具有毒性,符合中医“以毒攻毒”“有故

无殒”的抗肿瘤理念。其中,华山参、洋金花、天仙子、罂粟壳中富含的生物碱,是用于癌痛治疗的潜在药效物质<sup>[29]</sup>。牵牛子、千金子属峻下逐水类中药。现代药理研究<sup>[30-31]</sup>表明,二者均具有抗肝癌活性,其作用机制可能与下调水通道蛋白有关。白屈菜红碱能够通过下调抗凋亡蛋白表达,诱导肝癌HepG2细胞凋亡发生<sup>[32]</sup>,还能够抑制肝癌Hep3B细胞的增殖与侵袭<sup>[33]</sup>。关于白药子抗肝癌的药理研究尚未见明确报道,未来具有较高研究价值。中药资源是现代药物研发的宝库,与上述中药关联的活性成分中,已有许多成分在基础与临床研究中证实了其抗癌效果。表没食子儿茶素没食子酸酯[(-)-epigallocatechin-3-gallate, EGCG]对肝癌、乳腺癌、皮肤癌、肺癌等多种癌症具有肿瘤生长抑制作用,针对EGCG的抗肿瘤潜力,已经进行了一系列临床试验,并取得了一定的成效<sup>[34]</sup>。千层纸素A(oroxylin a)是黄芩的主要黄酮类化合物,在HepG2肝癌细胞中,它可以促进H<sub>2</sub>O<sub>2</sub>生成,从而激活蛋白激酶R样内质网激酶(PERK)-真核翻译起始因子2 $\alpha$ (eIF2 $\alpha$ )-活化转录因子4(ATF4)-C/EBP同源蛋白信号通路(CHOP)信号通路,抑制HepG2的细胞活力,而对正常肝细胞系L02无影响<sup>[35]</sup>。当然,目前结合高通量虚拟筛选与AI技术,我们可以更迅速地从众多天然小分子化合物中识别出具有抗肿瘤活性的候选药物,从而显著降低新药研发的成本和时间。

综上所述,本研究基于WGCNA联合机器学习,挖掘TCGA数据库,发现MDK、RASGRF2、CYP2C9、TERT、LCAT、ADRA2B可作为评估HCC的生物标志物,分析了6种生物标志物与不同类型免疫细胞之间的相关性。在此基础上,预测了可调控上述6种生物标志物的48味中药及其活性成分,为HCC的免疫治疗提供了潜在靶点,同时也为临床中医用药提供了新的策略。

## 参考文献

- [1] 中华人民共和国国家卫生健康委员会医政司.原发性肝癌诊疗指南(2024年版)[J].协和医学杂志,2024,15(3):532-559.
- [2] CHIDAMBARANATHAN-REGHUPATY S, FISHER P B, SARKAR D. Hepatocellular carcinoma (HCC): Epidemiology, etiology and molecular classification[J]. Adv Cancer Res, 2021, 149: 1-61.
- [3] SUNG H, FERLAY J, SIEGEL R L, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries[J]. CA Cancer J Clin, 2021, 71(3): 209-249.
- [4] 杜梅珍,梁善雄,肖晓兰,等.广西某医院2012年-2021年恶性肿瘤住院患者流行病学分析[J].江苏卫生事业管理,2023,34(6):845-849,869.
- [5] 吴亦枫,孙丽蕴,刘若愚,等.银屑病合并重度抑郁障碍差异表达基因及中药预测的生物信息学分析[J].中医导报,2024,30(6):125-132,151.
- [6] LANGFELDER P, HORVATH S. WGCNA: An R package for weighted correlation network analysis [J].



- BMC Bioinformatics, 2008,9:559.
- [7] OU H Y, YE X P, HUANG H S, et al. Constructing a screening model to obtain the functional herbs for the treatment of active ulcerative colitis based on herb-compound-target network and immuno-infiltration analysis[J]. Naunyn Schmiedebergs Arch Pharmacol, 2024, 397(7): 4693-4711.
- [8] Breiman L. Random forests[J]. Mach Learn, 2001, 45: 5-32.
- [9] TIBSHIRANI R. Regression shrinkage and selection via the lasso[J]. J R Stat Soc Ser B Stat Methodol, 1996, 58(1): 267-288.
- [10] 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.
- [11] CHEN T Q, GUESTRIN C. XGBoost: A scalable tree boosting system[C]//Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. San Francisco California USA. ACM, 2016: 785-794.
- [12] 黄依婷, 磨宁芳, 李姐姐, 等. 基于网络药理学和分子对接的逍遥散治疗肝癌并抑郁症的潜在机制探讨[J]. 药物评价研究, 2024, 47(6): 1175-1186.
- [13] 施亮, 顾凯, 吴春晓, 等. 2016年上海市肝癌发病和死亡情况与2002—2016年的变化趋势分析[J]. 肿瘤, 2023, 43(4): 277-286.
- [14] 戴靖宜, 蒋敬庭. 肝细胞肝癌肿瘤标志物诊断的新进展[J]. 诊断学理论与实践, 2023, 22(5): 486-493.
- [15] 李永顺, 谢嘉嘉, 李航, 等. 基于生物信息学及铁死亡分析特发性肺纤维化的差异基因及潜在的治疗中药预测[J]. 药物评价研究, 2023, 46(3): 501-509.
- [16] GOWHARI SHABGAH A, EZZATIFAR F, ARAVINDHAN S, et al. Shedding more light on the role of Midkine in hepatocellular carcinoma: New perspectives on diagnosis and therapy[J]. IUBMB Life, 2021, 73(4): 659-669.
- [17] FILIPPOU P S, KARAGIANNIS G S, CONSTANTINIDOU A. Midkine (MDK) growth factor: A key player in cancer progression and a promising therapeutic target[J]. Oncogene, 2020, 39(10): 2040-2054.
- [18] SAIKIA M, CHEUNG N, SINGH A K, et al. Role of midkine in cancer drug resistance: Regulators of its expression and its molecular targeting[J]. Int J Mol Sci, 2023, 24(10): 8739.
- [19] LU P F, CHEN J J, YAN L H, et al. RasGRF2 promotes migration and invasion of colorectal cancer cells by modulating expression of MMP9 through Src/Akt/NF- $\kappa$ B pathway[J]. Cancer Biol Ther, 2019, 20(4): 435-443.
- [20] LIN Q L, QIU M Q, WEI X Y, et al. Genetic variants of SOS2 MAP2K1 and RASGRF2 in the RAS pathway genes predict survival of HBV-related hepatocellular carcinoma patients[J]. Arch Toxicol, 2023, 97(6): 1599-1611.
- [21] 张拓, 李达军, 王海涛, 等. CYP2C9对肝细胞癌的迁移、侵袭和增殖的影响[J]. 医学新知, 2024, 34(6): 674-681.
- [22] SHEN Y, HUANG J J, JIA L, et al. Bioinformatics and machine learning driven key genes screening for hepatocellular carcinoma[J]. Biochem Biophys Rep, 2024, 37: 101587.
- [23] DRATWA M, WYSOCZAŃSKA B, LACINA P, et al. TERT-regulation and roles in cancer formation[J]. Front Immunol, 2020, 11: 589929.
- [24] RAGHUNANDAN M, GEELEN D, MAJEROVA E, et al. NHP2 downregulation counteracts hTR-mediated activation of the DNA damage response at ALT telomeres [J]. EMBO J, 2021, 40(6): e106336.
- [25] LONG J Y, CHEN P P, LIN J Z, et al. DNA methylation-driven genes for constructing diagnostic, prognostic, and recurrence models for hepatocellular carcinoma[J]. Theranostics, 2019, 9(24): 7251-7267.
- [26] HE W Z, WANG M, ZHANG X C, et al. Estrogen induces LCAT to maintain cholesterol homeostasis and suppress hepatocellular carcinoma development[J]. Cancer Res, 2024, 84(15): 2417-2431.
- [27] ZHANG Z Q, LI J, HE T S, et al. Two predictive precision medicine tools for hepatocellular carcinoma[J]. Cancer Cell Int, 2019, 19: 290.
- [28] 中华中医药学会肝胆病分会. 原发性肝癌中医诊疗指南[J]. 中西医结合肝病杂志, 2024, 34(4): 385-390.
- [29] 陈岩, 唐莹莹, 杨莉, 等. “有毒”中药生物碱类成分的毒性及代谢研究进展[J]. 药学报, 2023, 58(11): 3285-3295.
- [30] 王梦琪, 林海红, 杜钢军. 牵牛子水煎液对小鼠H22肝癌皮下移植瘤的抗肿瘤作用和机制[J]. 河南大学学报(医学版), 2019, 38(2): 77-84.
- [31] 贺丽娜, 刘力, 贺晓旭, 等. 千金子提取物调控AK125512对肝癌细胞增殖、迁移、侵袭的影响研究[J]. 肝胆外科杂志, 2022, 30(3): 228-232, 239.
- [32] 韩聪, 朱国福. 白屈菜红碱诱导人肝癌细胞HepG2凋亡及机制探讨[J]. 中国实验方剂学杂志, 2016, 22(11): 127-130.
- [33] ZHU Y Z, PAN Y Y, ZHANG G B, et al. Chelerythrine inhibits human hepatocellular carcinoma metastasis in vitro[J]. Biol Pharm Bull, 2018, 41(1): 36-46.
- [34] 董雅楠. 表没食子儿茶素没食子酸酯抗肿瘤作用机制的研究进展[J]. 现代药物与临床, 2023, 38(4): 1008-1013.
- [35] 张志强, 王琳. 黄酮类化合物抗肝细胞癌HepG2细胞系的研究进展[J]. 化工时刊, 2022, 36(12): 29-31.

(收稿日期: 2025-01-27 编辑: 刘国华)