

基于网络药理学、机器学习和分子对接技术探讨丹参治疗急性脑梗死的作用机制

戴京晶¹, 刘通¹, 汪鑫^{2,3*}, 赵建雄³, 朱加应³, 吕宸蓁⁴, 丁文烨⁴

1. 贵州中医药大学, 贵州 贵阳 550000
2. 上海交通大学医学院附属上海儿童中心贵州医院, 贵州 贵阳 550000
3. 贵州省人民医院, 贵州 贵阳 550000
4. 安徽医科大学 临床医学院, 安徽 合肥 230012

摘要:目的 通过网络药理学、机器学习和分子对接技术探讨丹参治疗急性脑梗死的作用机制。方法 通过中药系统药理学数据库与分析平台 (TCMSP) 与 Swiss Target Prediction 数据库, 预测并筛选丹参的主要活性成分及其作用靶点; 同时, 通过 OMIM、GeneCards 等相关数据库获取脑梗死相关靶点, 经 R 软件韦恩图分析确定交集靶点; 借助 STRING 构建蛋白相互作用 (PPI) 网络, 通过 Cytoscape 构建“丹参-活性成分-靶点-通路”网络, 初筛核心靶点; 整合 GEO 数据集 GSE58294, 经 limma 包差异分析和 WGCNA 共表达网络分析筛选候选靶点, 通过 R 包进行基因本体 (GO) 和京都基因与基因组百科全书 (KEGG) 富集分析; 运用随机森林、支持向量机、逻辑回归 3 种机器学习模型及 SHAP 分析、接受者操作特征 (ROC) 曲线筛选并验证核心基因; 最后经 AutoDock Tools 进行分子对接, 验证核心成分与靶点的结合稳定性。结果 对丹参的 610 个活性成分相关靶点与脑梗死相关的 11 520 个靶点进行交集分析, 共识别出 610 个交集靶点, 核心成分包括丹参酮 II_A 等。GO 富集分析共得到 747 个显著条目, BP 主要涉及肌肉系统过程、肌肉收缩、平滑肌收缩及其正调控, 以及对肽激素的反应等; CC 显著富集于顶端质膜、神经元投射末端、内质网腔、膜微区及膜筏等结构; MF 主要包括蛋白酪氨酸激酶活性、组蛋白 H2AX^{Y142} 与 H3Y⁴¹ 激酶活性, 以及氧化还原酶活性等。表明目标基因在肌肉功能、细胞信号转导及表观遗传调控中发挥重要作用。KEGG 富集分析鉴定出 24 条相关通路, 涉及色氨酸代谢、破骨细胞分化、核因子- κ B (NF- κ B) 信号通路、造血细胞谱系及急性髓系白血病等。交集靶点分别为 B 淋巴细胞瘤-2 相关蛋白 A1 (BCL2A1)、钙调蛋白依赖性蛋白激酶 1 (CAMK1)、组织蛋白酶 K (CTSK)、鞘氨醇 1 磷酸酯受体 3 (S1PR3)、丝裂原激活蛋白激酶激酶 6 (MAP2K6)、类固醇硫酸酯酶 (STS)、钾电压门控通道 1 (KCNE1)。分子对接证实丹参酮 II_A 与核心靶点结合稳定。结论 丹参可通过丹参酮 II_A 等核心成分靶向调控 CTSK、S1PR3 和 BCL2A1 等关键靶点, 干预 NF- κ B 信号通路、色氨酸代谢等核心通路, 发挥治疗急性脑梗死的作用。

关键词: 丹参; 急性脑梗死; 网络药理学; 机器学习; 分子对接; 丹参酮 II_A; 组织蛋白酶 K; 鞘氨醇 1 磷酸酯受体 3; B 淋巴细胞瘤-2 相关蛋白 A1

中图分类号: R285; R286.1 文献标志码: A 文章编号: 1674-5515(2026)02-0284-11

DOI: 10.7501/j.issn.1674-5515.2026.02.004

Mechanism of *Salvia miltiorrhiza* in treatment of acute ischemic stroke based on network pharmacology, machine learning and molecular docking technology

DAI Jingjing¹, LIU Tong¹, WANG Xin^{2,3}, ZHAO Jianxiong³, ZHU Jiaying³, LÜ Chenqu⁴, DING Wenyue⁴

1. Guizhou University of Traditional Chinese Medicine, Guiyang 550000, China
2. Shanghai Children's Medical Center GuiZhou Hospital, Shanghai Jiaotong University School of Medicine, Guiyang 550000, China
3. Guizhou Provincial People's Hospital, Guiyang 550000, China
4. Clinical Medical College, Anhui Medical University, Hefei 230012, China

收稿日期: 2025-11-19

基金项目: 国家自然科学基金资助项目 (82260259)

作者简介: 戴京晶, 研究方向为临床疼痛诊疗方法研究。E-mail: 896071358@qq.com

*通信作者: 汪鑫, 博士, 硕士生导师, 主任医师, 研究方向为麻醉学。E-mail: 110667637@qq.com

Abstract: Objective To explore the mechanism of *Salvia miltiorrhiza* in treatment of acute ischemic stroke based on network pharmacology, machine learning and molecular docking technology. **Methods** The main active components and their targets of *S. miltiorrhiza* were predicted and screened by TCMSP and Swiss Target Prediction database. At the same time, the relevant targets of cerebral infarction were obtained through OMIM, Gene Cards and other related databases, and the intersection targets were determined by Wayne diagram analysis of R software. The protein-protein interaction (PPI) network was constructed by STRING, and the “*S. miltiorrhiza*-active ingredient-target-pathway” network was constructed by Cytoscape to screen the core targets. The GEO dataset GSE58294 was integrated, and candidate targets were screened by limma package difference analysis and WGCNA co-expression network analysis. Gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis were performed by R package. Three machine learning models of random forest, support vector machine and logistic regression, SHAP analysis and receiver operating characteristic (ROC) curve were used to screen and verify the core genes. Finally, molecular docking was performed by AutoDock Tools to verify the binding stability of the core components and the target. **Results** A total of 610 intersection targets were identified by intersection analysis of 610 active components related targets and 11 520 targets related to ischemic stroke, and the core components included tanshinone II_A. A total of 747 significant items were obtained by GO enrichment analysis. BP mainly involved muscle system processes, muscle contraction, smooth muscle contraction and its positive regulation, and response to peptide hormones. CC was significantly enriched in the apical plasma membrane, neuronal projection end, endoplasmic reticulum cavity, membrane microdomain and membrane raft. MF mainly included protein tyrosine kinase activity, histone H2AXY142 and H3Y41 kinase activity, and oxidoreductase activity. It showed that the target gene plays an important role in muscle function, cell signal transduction and epigenetic regulation. KEGG enrichment analysis identified 24 related pathways, involving tryptophan metabolism, osteoclast differentiation, nuclear factor- κ B (NF- κ B) signaling pathway, hematopoietic cell lineage and acute myeloid leukemia. The intersection targets were B-cell lymphoma-2-associated protein A1 (BCL2A1), calmodulin-dependent protein kinase 1 (CAMK1), cathepsin K (CTSK), sphingosine 1 phosphate receptor 3 (S1PR3), mitogen-activated protein kinase kinase 6 (MAP2K6), steroid sulfatase (STS), potassium voltage valve channel 1 (KCNE1). Molecular docking confirmed that tanshinone II_A binded stably to the core target. **Conclusion** *S. miltiorrhiza* can target and regulate key targets such as CTSK, S1PR3 and BCL2A1 through core components such as tanshinone II_A, interfere with core pathways such as NF- κ B signaling pathway and tryptophan metabolism, and play a role in the treatment of acute ischemic stroke.

Key words: *Salvia miltiorrhiza* Bge.; acute ischemic stroke; network pharmacology; machine learning; molecular docking; tanshinone II_A; CTSK; S1PR3; BCL2A1

脑梗死俗称中风,在全球年发病人数约为 1 500 万,致死 600 余万,致永久性残疾约 500 万,且复发率高,社会负担沉重^[1-3]。脑梗死占有所有卒中的 71%,主要由神经外科手术、颈动脉粥样硬化、血管栓塞、高血压及高血糖引发^[4-6]。该病呈年轻化趋势,发病率随年龄增长而激增,已成为重大公共卫生挑战^[7-8]。现行治疗包括抗血小板、溶栓、血管内介入及神经保护,重组组织型纤溶酶原激活剂虽然为美国食品药品监督管理局唯一批准的药物,但其治疗窗极窄(发病 3~4.5 h),禁忌证多、并发症风险高,仅 3%~5%患者获益,临床应用严重受限^[9-11]。中医认为脑梗死的基本发病机制为虚、火、风、痰、气、血,血瘀是缺血性卒中的主要病理机制^[12]。因此,以“活血化瘀”为基本疗法治疗脑梗死,往往疗效更佳。据《妇人明理论》记载丹参具有“补血生血,调血敛血,活血逐瘀”的功效;《本草纲目》记载丹参具有“补新血,调经脉”的功效。研究表明丹参具有抗炎、抗凋亡和调节细胞自噬的功效,

且其已被证明治疗脑梗死疗效确切并在临床广泛应用^[13]。基于此,本研究运用网络药理学、机器学习和分子对接技术等,系统解析丹参对脑梗死的干预机制,为脑梗死治疗提供新的思路,为临床治疗寻找更有效的药物。

1 网络药理学

1.1 丹参活性成分及靶点筛选

通过中药系统药理数据库与分析平台(TCMSP)获取丹参的化学成分。依据口服生物利用度(OB)不低于 30%及类药性(DL)≥0.18 的标准,筛选出丹参潜在的活性成分。随后,对相关靶点信息进行合并与去冗余处理,并利用 UniProt 数据库统一规范其基因命名。同时,基于 PubChem 数据库提供的成分 SMILES 标识,借助 Swiss Target Prediction 数据库进一步补充预测丹参活性成分的相关作用靶点。

1.2 脑梗死相关靶点筛选

以“ischemic stroke”为关键词,检索 OMIM 数

据库 (<https://omim.org/>)、GeneCards 数据库 (<https://www.genecards.org/>)、DrugBank 数据库 (<https://go.drugbank.com/>) 和 Therapeutic Target Database 数据库 (<https://www.ttd.idrblab.cn>) 获取脑梗死相关靶点。

1.3 丹参 - 脑梗死交集靶点识别

运用 R 软件 (v4.4.1) 的 Venn Diagram 包绘制丹参与脑梗死相关靶点的韦恩图, 识别两者交集靶点。

1.4 蛋白相互作用 (PPI) 网络和“丹参 - 活性成分 - 靶点 - 通路”网络构建

将药物 - 疾病交集靶点导入 STRING 数据库 (<http://cn.string-db.org>), 设定物种为“Homo sapiens”, 置信度阈值 > 0.900, 构建 PPI 网络。将此网络导入 Cytoscape 3.10.2 软件, 运用 Cytoscape 网络中心性分析插件 (CytoNCA) 获取度值 (degree), 最终选取 degree 最高的成分作为丹参治疗脑梗死的核心成分。

1.5 脑梗死相关差异基因筛选

从 GEO 数据库入手, 以“ischemic stroke”为关键词进行检索, 最终选定 GSE58294 数据集用于后续分析。下载该数据集的原始表达矩阵及样本注释文件, 并依照实验设计方案将所有样本划分为对照组和治疗组。为提升数据分布正态性并减小方差波动, 所有芯片表达量数据均进行了 \log_2 转换。在此基础上, 本研究基于 limma 3.58.1 包开展差异表达分析, 通过比较组间表达水平, 并以 $|\log_2FC| > 1$ 且调整后 $P < 0.05$ 为阈值, 筛选得到具有统计学意义的差异表达基因。

1.6 加权基因共表达网络分析 (WGCNA) 网络构建

通过 R 语言中的 WGCNA 包, 采用 Pearson 相关矩阵和平均关联法分析基因配对关系, 将 ID: GSE58294 代入运算, 获取 WGCNA 共表达网络^[14]。采用分层聚类分析剔除任何异常值样本, 然后利用 WGCNA 包中的 pick Soft Threshold 函数, 根据相关系数 (r) > 0.9 的标准, 选择合适的软阈值 (β)。这个过程生成了一个邻接矩阵, 然后将其转换为拓扑重叠矩阵 (TOM)。通过各种 TOM 测量, 通过平均连锁分层聚类将表现出相似表达模式的基因分配到不同的模块中。并计算这些模块与临床特征之间的相关性, 得到相关性热图。

1.7 差异基因、丹参 - 脑梗死交集靶点和 WGCNA 靶点取交集

运用 R 软件的 Venn Diagram 包绘制差异基因、

丹参 - 脑梗死交集靶点、WGCNA 靶点的韦恩图, 识别三者交集靶点。

1.8 基因本体 (GO) 和京都基因与基因组百科全书 (KEGG) 通路富集分析

借助 R 语言中的 org.Hs.eg.db 包将所有靶点基因标识符转换为标准格式, 应用 ggplot2、stringi 及 colorspace 等程序包, 对符合显著性阈值 ($P < 0.05$) 的条目选择生物过程 (BP)、分子功能 (MF) 和细胞组成 (CC) 进行 GO 富集分析, 同时进行 KEGG 通路富集分析。

1.9 机器学习模型预测靶点取交集

使用 R 软件的 limma 包中对 1.7 项中获取到交集靶点和 1.5 项获得的差异基因获得基因矩阵文件, 分别运用逻辑回归、支持向量机、随机森林 3 种机器学习模型筛选靶点, 最后取交集。

1.10 核心基因 SHAP 分析及评估

通过 SHAP 分析明确丹参治疗急性脑梗死的关键基因及作用机制, 结合随机森林、支持向量机、极端梯度提升、梯度提升机和 K 近邻 5 种机器学习模型的曲线下面积 (AUC) 分析验证模型与特征靶点的可靠性。运用 R 软件 kernelshap 包对“1.9”项中的靶点进行 SHAP 特征重要性分析, 并打分, 再使用“pROC”包对靶点进行进一步评估, 最后得到核心靶点。

1.11 分子对接

从蛋白质结构数据库 (PDB, <http://www.rcsb.org>) 下载核心靶点蛋白的三维结构作为受体。自 TCMSPP 获取丹参主要活性成分的二维结构作为配体。应用 AutoDock Tools 软件 (v1.5.6) 执行分子对接, 计算并评估配体 - 受体间的结合能。

2 结果

2.1 丹参活性成分及靶点筛选

从 TCMSPP 和 Swiss Target Prediction 数据库中分别获取了 125、671 个丹参活性成分相关靶点, 去除重复值并进行汇总, 最终得到了 610 个靶点。

2.2 脑梗死相关靶点筛选

通过系统检索 OMIM、GeneCards、DrugBank 和 Therapeutic Target Database 数据库, 获取脑梗死相关靶点。经筛选去重后, 共鉴定出 11 520 个脑梗死相关靶点。

2.3 丹参 - 脑梗死交集靶点识别

通过 R 语言, 对丹参的 610 个活性成分相关靶点与脑梗死相关的 11 520 个靶点执行交集分析。结

果共识别出 610 个交集靶点，这些靶点被视为丹参治疗脑梗死的潜在作用靶点，见图 1。

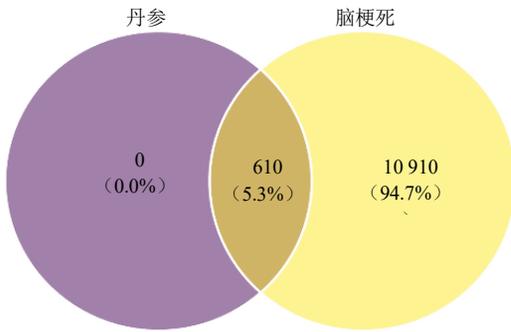


图 1 丹参活性成分相关靶点和脑梗死相关靶点的韦恩图
Fig. 1 Venn diagram of related targets of active ingredients of *Salvia miltiorrhiza* and related targets of ischemic stroke

2.4 PPI 网络和“丹参 - 活性成分 - 靶点 - 通路”网络构建

基于 STRING 数据库构建的 PPI 网络如图 2 所示，图中节点越大，表示该蛋白在网络中越重要。利用 Cytoscape 3.10.2 软件构建的丹参 - 活性成分 - 靶点 - 通路网络如图 3 所示。该网络拓扑结构中箭头节点代表疾病和药物，圆形节点代表基因，方形节点表示活性成分，正方形节点代表活性成分，菱

形节点代表通路，六边形节点表示药物种类。连接关系清晰展示了活性成分与靶点间的相互作用模式，印证了该丹参通过多组分 - 多靶点实现协同调控的特点。以节点度 (degree) 为拓扑指标进行分析，其中节点 degree 值最高的成分为丹参酮 II_A。

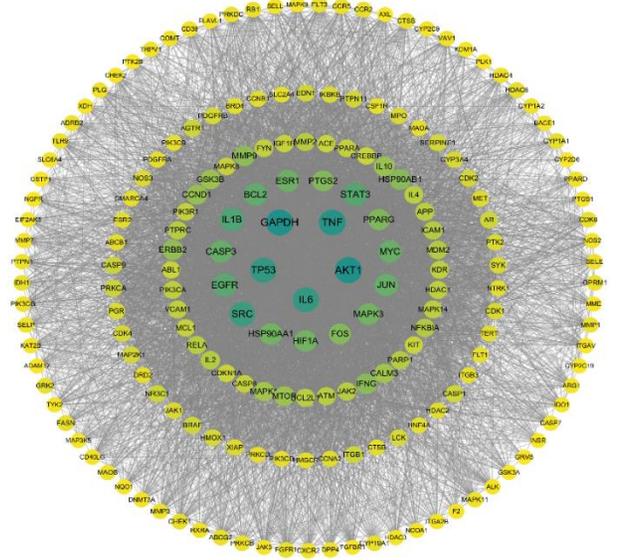


图 2 丹参与脑梗死交集靶点 PPI 网络
Fig. 2 PPI network of intersection targets of *S. miltiorrhiza* and ischemic stroke

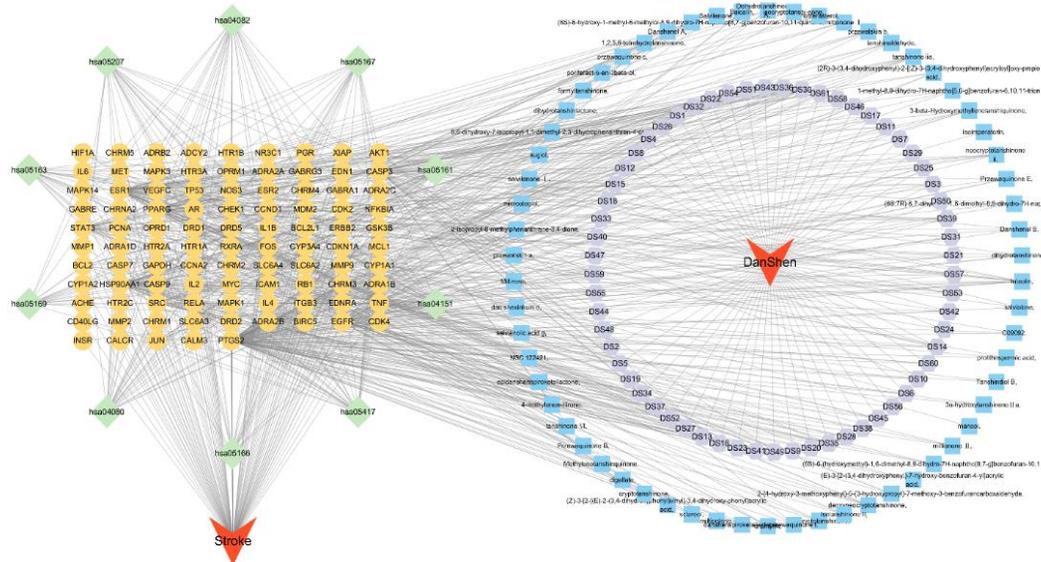


图 3 丹参 - 活性成分 - 靶点 - 通路网络
Fig. 3 *S. miltiorrhiza*-active ingredient-target-pathway network

2.5 脑梗死相关差异基因筛选

脑梗死相关基因芯片数据 GSE58294 通过 batch 整合后，比较了对照组和治疗组的基因表达差异，绘制了热图 (图 4A)、主成分分析图 (图 4B) 和火山

图 (图 4C)，显示了 10 254 个下调基因和 11 402 个上调基因。热图的基因聚类分析进一步证实，不同实验组间基因表达显著差异，反映了多基因网络调控在脑梗死中的作用。

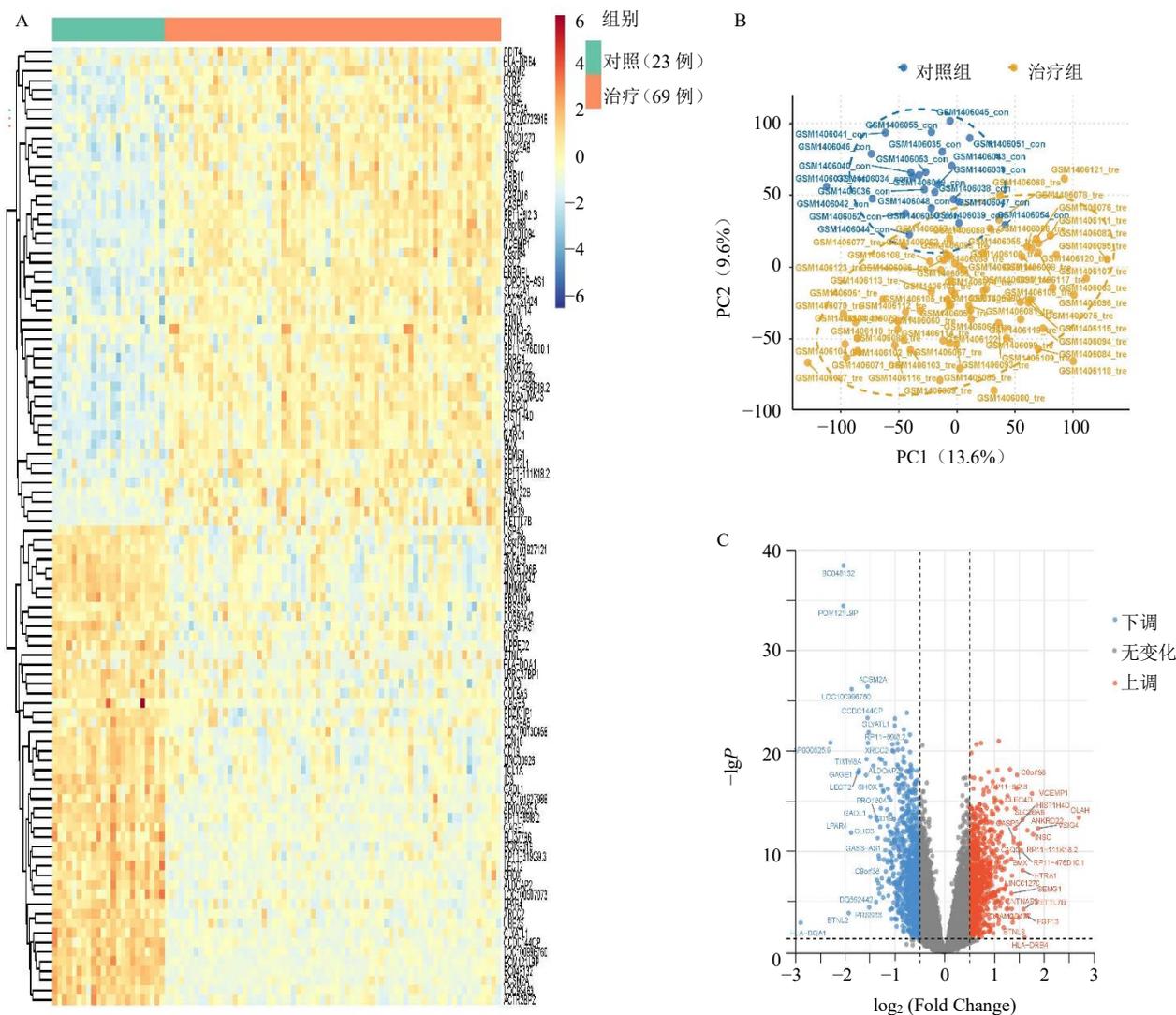


图 4 脑梗死相关差异基因的热图 (A)、主成分分析图 (B) 和火山图 (C)

Fig. 4 Heat map (A), principal component analysis (B) and volcano map (C) of ischemic stroke related differential genes

2.6 WGCNA 网络分析

使用 WGCNA 进行分析, 深入研究了与脑梗死临床特征显著相关的基因模块, 如图 5A 所示, 采用 pick Soft Threshold 函数确定最优软阈值为 8。利用动态剪枝算法进行模块识别, 表达模式相近的基因被划分到同一模块中, 最终得到 10 个模块 (灰色、黄色、红色、绿黄色、蓝色、紫色、黑色、洋红色、绿色和棕色模块) (图 5B)。其中, 蓝色模块的基因表达谱与脑梗死呈显著正相关 ($r=0.76, P=1.2 \times 10^{-18}$), 黑色模块的基因表达谱与脑梗死呈显著负相关 ($r=0.92, P=3.5 \times 10^{-38}$) (图 5C), 提示蓝色和黑色模块可能含有与脑梗死相关的特异性功能基因。由图 5D 可知, 蓝色和黑色核心模块的基

因数量分别为 475、231。

2.7 差异基因、丹参 - 脑梗死交集靶点和 WGCNA 靶点取交集

运用 R 软件的 Venn Diagram 包, 绘制差异基因、药物 - 疾病交集靶点、WGCNA 靶点的韦恩图, 识别三者交集靶点分别为组织蛋白酶 K (CTSK)、鞘氨醇 1 磷酸酯受体 3 (S1PR3)、B 淋巴细胞瘤-2 相关蛋白 A1 (BCL2A1)、酪氨酸激酶 (LCK)、丝裂原激活蛋白激酶激酶 6 (MAP2K6)、精氨酸酶 1 (ARG1)、NAD(P)H 苯醌脱氢酶 2 (NQO2)、钙调蛋白依赖性蛋白激酶 1 (CAMK1)、类固醇硫酸酯酶 (STS)、钾电压门通道 1 (KCNE1)、膜金属内肽酶 (MME) 等 21 个, 见图 6。

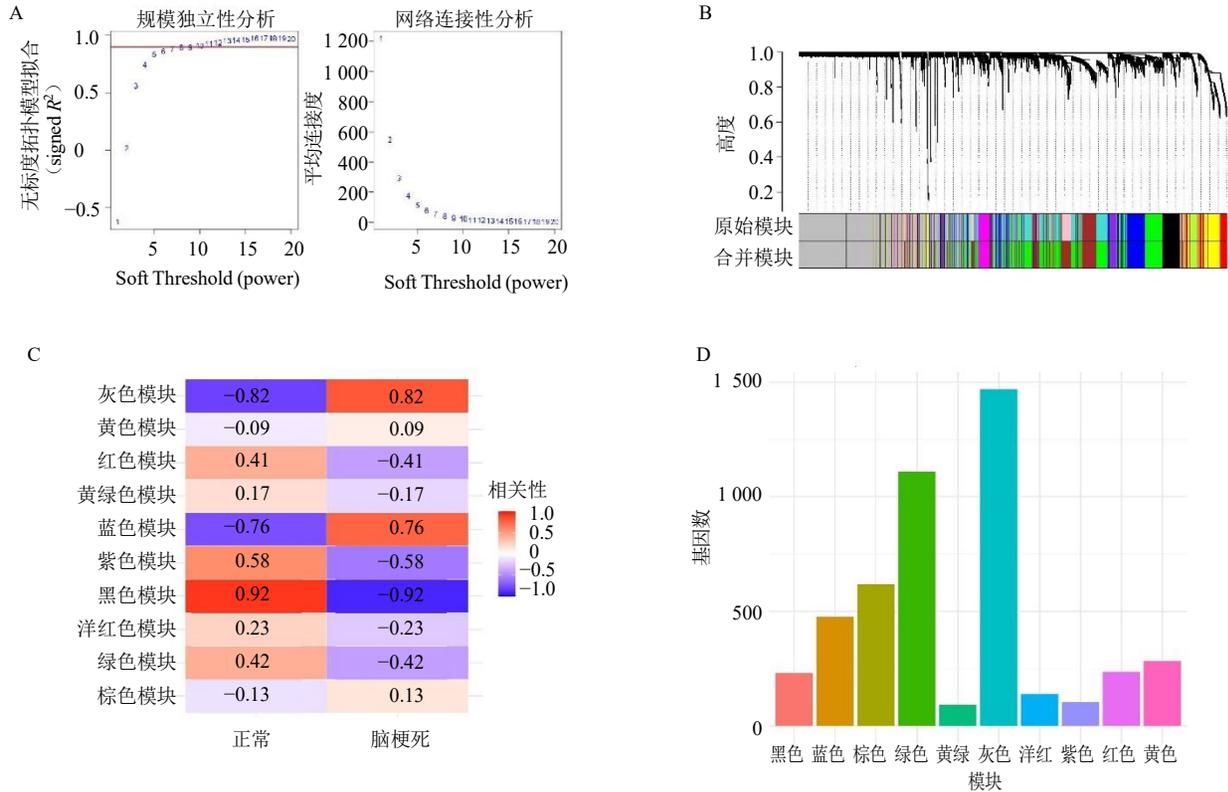


图 5 WGCNA 网络分析中软阈值筛选图 (A)、基因模块分类图 (B)、基因相关热图 (C) 和基因数量柱状图 (D)
 Fig. 5 Soft threshold screening map (A), gene module classification map (B), gene related heat map (C) and gene number histogram (D) in WGCNA network analysis

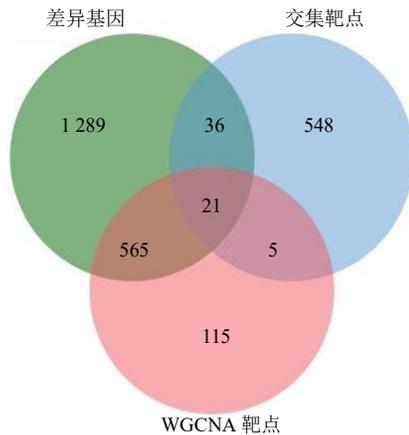


图 6 差异基因、丹参 - 脑梗死交集靶点和 WGCNA 靶点的韦恩图

Fig. 6 Venn diagram of differential genes, *S. miltiorrhiza*-ischemic stroke intersection targets and WGCNA targets

2.8 GO 和 KEGG 通路富集分析

GO 富集分析共得到 747 个显著条目 ($P < 0.05$)。其中, BP 主要涉及肌肉系统过程、肌肉收缩、平滑肌收缩及其正调控, 以及对肽激素的反应等; CC 显著富集于顶端质膜、神经元投射末端、内质网腔、膜微区及膜筏等结构; MF 主要包括蛋白

酪氨酸激酶活性、组蛋白 H2AXY142 与 H3Y41 激酶活性, 以及氧化还原酶活性等。表明目标基因在肌肉功能、细胞信号转导及表观遗传调控中发挥重要作用。KEGG 通路富集分析鉴定出 24 条相关通路, 涉及色氨酸代谢、破骨细胞分化、核因子- κ B (NF- κ B) 信号通路、造血细胞谱系及急性髓系白血病等。提示目标基因可能通过调控免疫炎症反应、细胞分化与代谢过程, 在疾病发生与发展中发挥关键作用, 见图 7。

2.9 机器学习模型预测靶点取交集

运用随机森林、支持向量机和逻辑回归 3 种机器学习模型预测靶点取交集, 得到交集靶点分别为 BCL2A1、CAMK1、CTSK、S1PR3、MAP2K6、STS、KCNE1, 见图 8。

2.10 核心基因 SHAP 分析及评估

特征重要性结果显示, CTSK、S1PR3 和 BCL2A1 为核心特征基因, 其中 CTSK 和 S1PR3 的 SHAP 值分布范围广且与表达量关联显著, BCL2A1 高表达时 SHAP 值正向升高, 见图 9A。特征依赖分析表明, CTSK、S1PR3 表达量与 SHAP 值呈负相

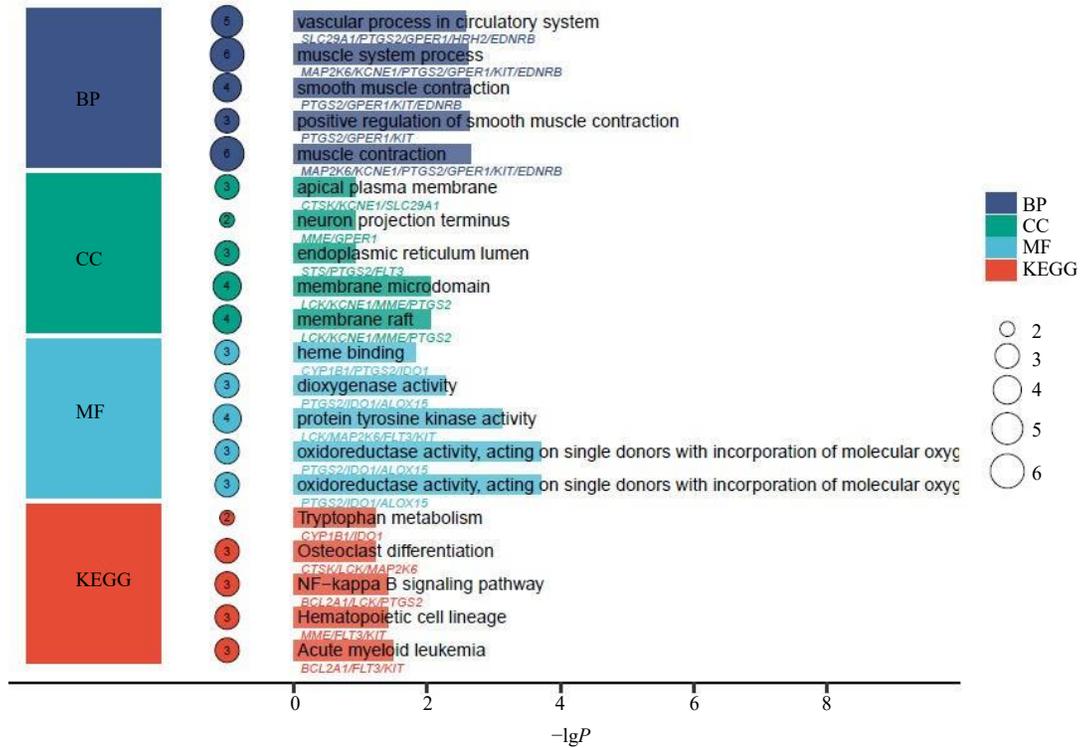


图 7 GO 和 KEGG 富集功能分析

Fig. 7 GO and KEGG enrichment function analysis

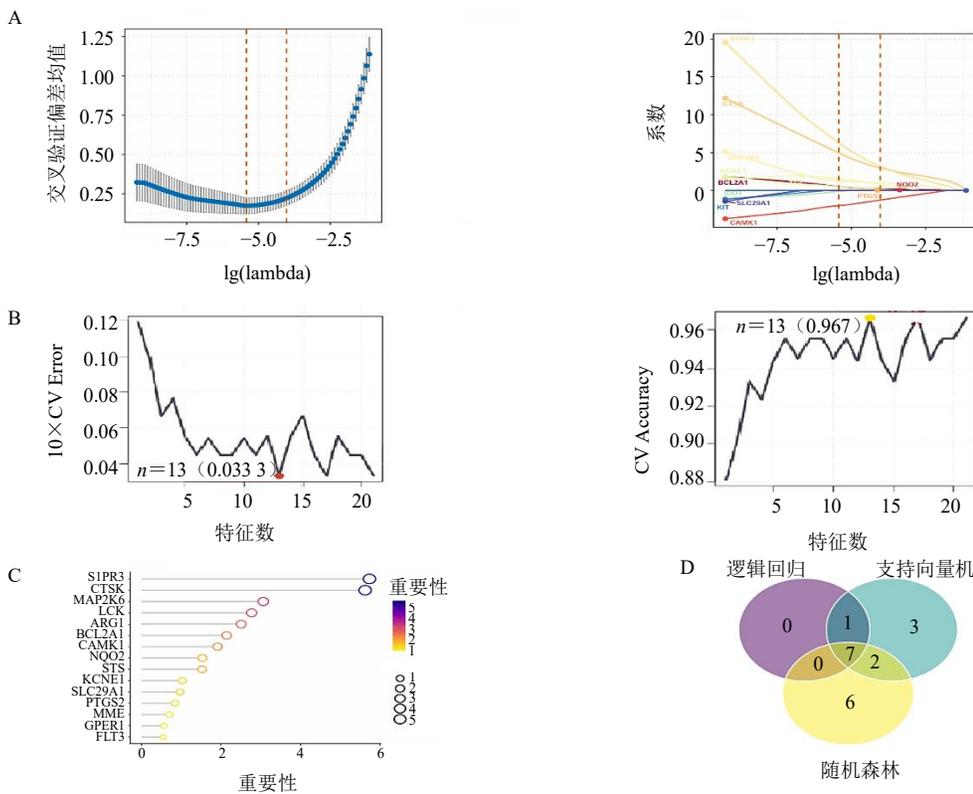


图 8 逻辑回归 (A)、支持向量机 (B)、随机森林 (C) 模型筛选靶点及其交集靶点图 (D)

Fig. 8 Logistic regression (A), support vector machine (B), random forest (C) model screening targets and their intersection target diagram (D)

关, *BCL2A1* 则呈正相关, 提示三者可能通过调控疾病风险评分参与丹参治疗机制, 见图 9B。瀑布图 (图 9C) 与力图 (图 9D) 显示, *BCL2A1* (SHAP 值为 2.52) 为主要正向贡献基因, *KCNE1* (SHAP 值为 -2.61) 和 *S1PR3* (SHAP 值为 -2.03) 为主要负向贡献基因, 且 *CTSK* 与 *S1PR3* 存在表达交互作用, 二者特定表达组合 (*CTSK* 高表达 + *S1PR3* 低表达) 可能增强丹参疗效; 分组分析发现丹参干预可上调 *CTSK*、*BCL2A1* 和下调 *S1PR3*、*KCNE1* 的 SHAP 重要性, 重塑疾病分子网络。AUC 分析显示, 随机森林、支持向量机、极端梯度提升和梯度提升机模型敏感值均达 1.000 (95% CI [1.000, 1.000]), K 近邻模型敏感值为 0.992 (95% CI [0.972, 1.000]), 证实模型诊断效能优异, 核心靶点可作为丹参治疗脑梗死的潜在靶点。

为验证核心基因对丹参治疗脑梗死的诊断价

值, 对 7 个关键基因 (*CTSK*、*S1PR3*、*BCL2A1*、*MAP2K6*、*STS*、*CAMK1*、*KCNE1*) 进行接受者操作特征 (ROC) 曲线分析, 如图 10 所示, 评估其区分对照组与丹参干预组的效能。结果显示, 这些基因的 AUC 值均高于 0.85, 表现出良好的诊断潜力: 其中 *CTSK* 的诊断效能最优, AUC 值达 0.955 (95% CI [0.892, 0.999]); *S1PR3* 次之, AUC 值为 0.943 (95% CI [0.890, 0.982]); *BCL2A1* (AUC=0.899, 95% CI [0.818, 0.962])、*MAP2K6* (AUC=0.896, 95% CI [0.810, 0.959])、*STS* (AUC=0.886, 95% CI [0.813, 0.944])、*CAMK1* (AUC=0.876, 95% CI [0.803, 0.938]) 及 *KCNE1* (AUC=0.863, 95% CI [0.779, 0.936]) 均展现出稳定的诊断性能。上述结果证实, *CTSK* 和 *S1PR3* 等核心基因可作为丹参治疗脑梗死的潜在诊断标志物, 为临床疗效评估提供可靠的分子依据。

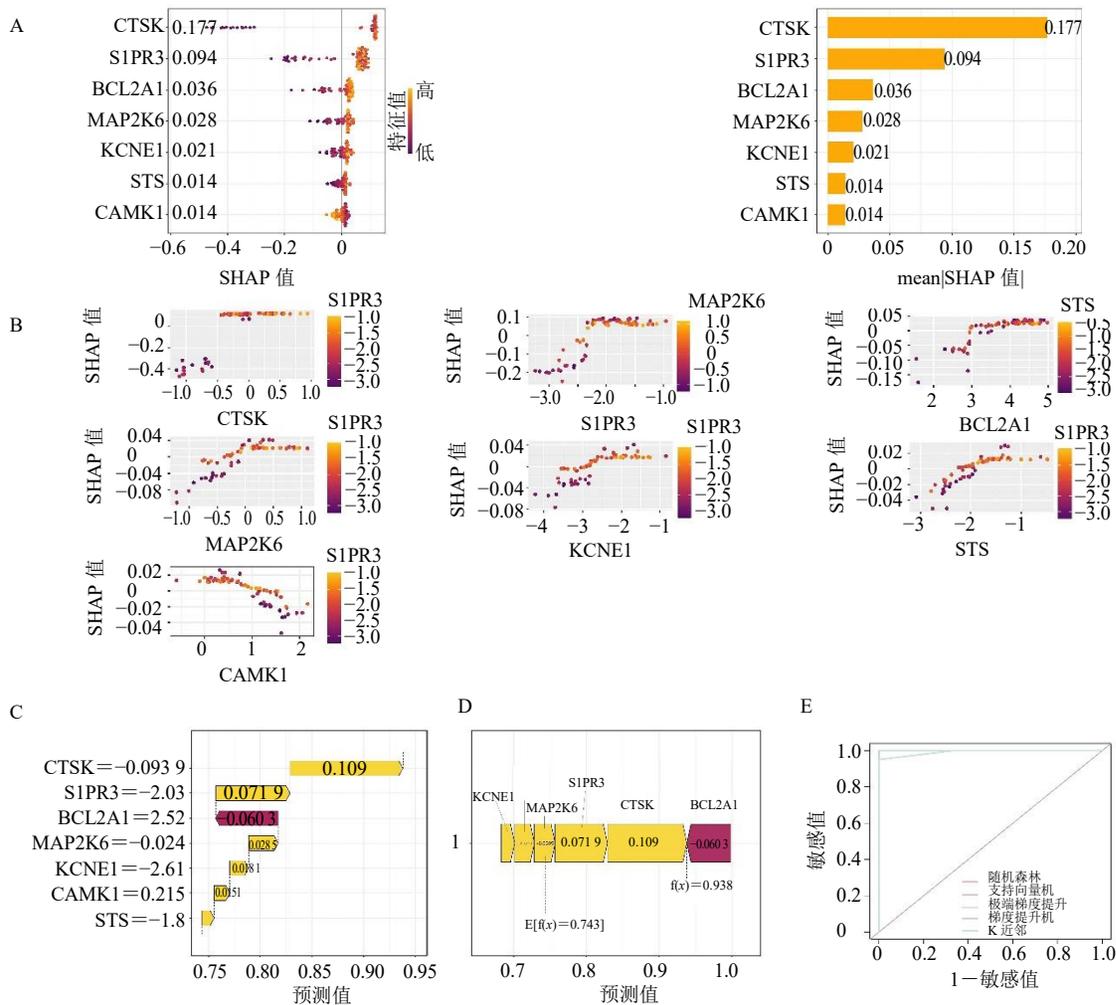


图 9 核心基因的特征重要性分析 (A)、特征依赖分析 (B)、特征重要性瀑布 (C)、特征重要性力图 (D) 和 AUC 分析 (E)

Fig. 9 Feature importance analysis (A), feature dependence analysis (B), feature importance waterfall (C), feature importance diagram (D) and AUC analysis (E) of core genes

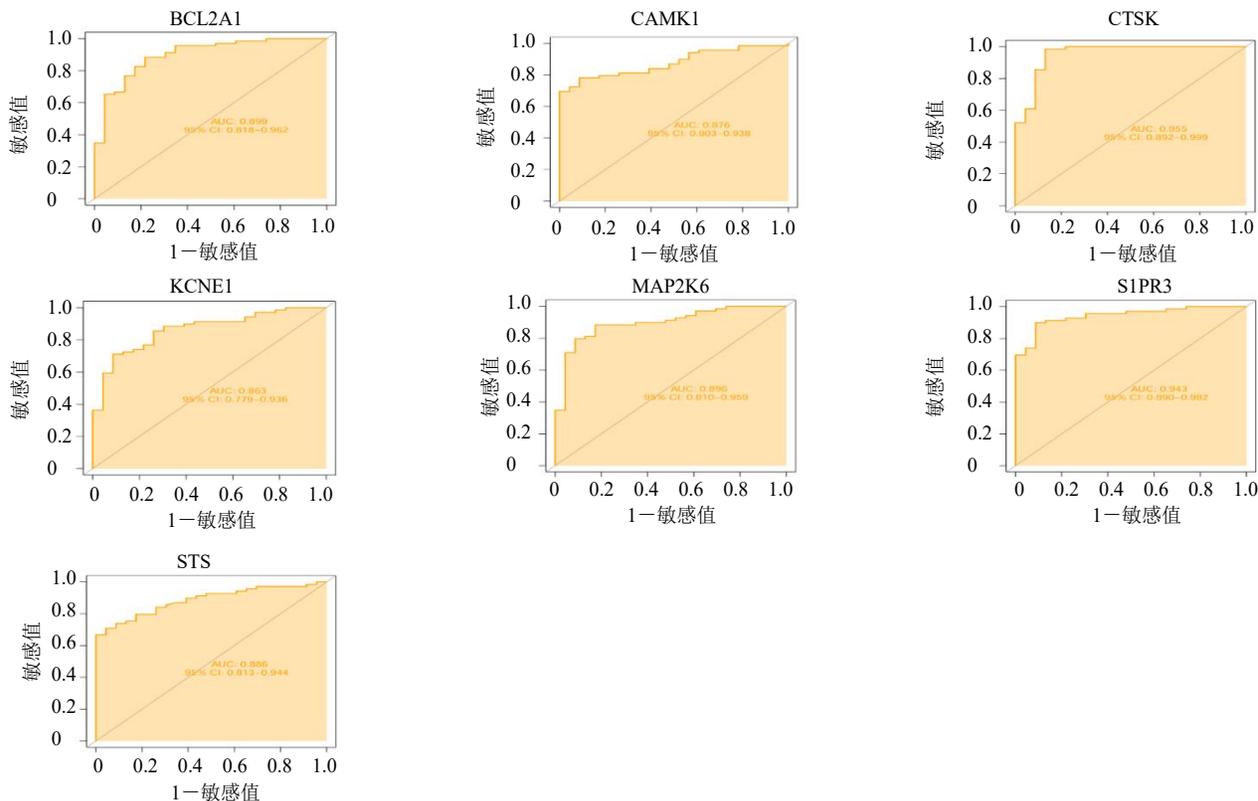


图 10 ROC 曲线分析
Fig. 10 ROC curve analysis

2.11 核心靶点与药物活性成分的分子对接

筛选出丹参中度值最高的活性成分丹参酮 II_A 为配体，并将上述筛选出的 7 个核心靶点 CTSK、S1PR3、BCL2A1、MAP2K6、STS、CAMK1、KCNE1 作为受体进行分子对接，结合能分别为 -8.339、-7.743、-6.934、-8.274、-7.959、-8.673、-8.996 kcal/mol (1 kcal=4.2 kJ)，均 ≤ -6.0 kcal/mol，表明核心靶点与关键活性成分均呈现显著结合亲和力。进一步对结合度最优的前 3 组合 (KCNE1-丹参酮 II_A、CAMK1-丹参酮 II_A、CTSK-丹参酮 II_A) 进行可视化分析 (图 11)。

3 讨论

本研究创新性整合网络药理学、机器学习及分子对接技术，从多维度解析丹参治疗脑梗死的作用机制，不仅验证了丹参“多成分-多靶点-多通路”的经典药理特征，更通过精准筛选与功能验证，为其临床转化与机制深化提供了新视角。

从活性成分与靶点筛选来看，依托 TCMSP 与 Swiss Target Prediction 数据库，本研究共获取丹参相关作用靶点 610 个，结合 OMIM、GeneCards 等数据库筛选的 11 520 个脑梗死疾病靶点，通过交集

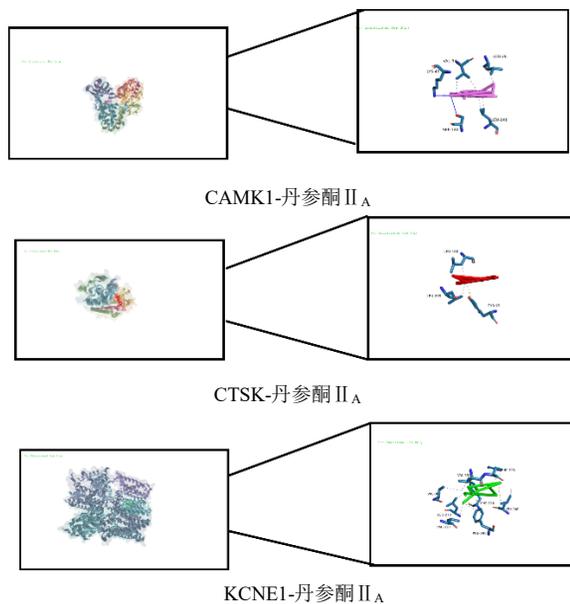


图 11 分子对接可视化

Fig. 11 Molecular docking visualization

分析锁定 610 个共同作用靶点，构建了丹参干预脑梗死的潜在分子网络。丹参-活性成分-靶点-通路网络拓扑分析显示，丹参酮 II_A 为核心活性成分 (度值最高)，该成分作为丹参脂溶性成分的代表，

有研究证实其可通过改善脑微循环、抑制神经细胞凋亡及减轻氧化应激损伤,在脑梗死模型中发挥显著神经保护作用,与本研究网络预测结果形成呼应^[15-17]。进一步结合 GSE58294 数据集差异分析、WGCNA 共表达网络构建,以及逻辑回归、支持向量机、随机森林 3 种机器学习模型的交叉验证,最终筛选出 CTSK、S1PR3、BCL2A1、MAP2K6、STS、CAMK1、KCNE1 共 7 个核心靶点,其中 S1PR3 与脑梗死的直接关联已明确,其余新发现靶点为深入解析疾病机制提供了精准的分子锚点^[18-19]。

GO 与 KEGG 通路富集分析从功能层面揭示了丹参干预脑梗死的关键调控方向。GO 富集结果显示,核心靶点主要参与肌肉系统过程、平滑肌收缩正调控、蛋白酪氨酸激酶活性及氧化还原酶活性等生物学过程,其中平滑肌收缩调控与脑缺血后血管舒缩功能恢复密切相关,而蛋白酪氨酸激酶活性与氧化还原酶活性则分别涉及细胞信号转导与氧化应激平衡,提示丹参可能通过多环节调节脑血管功能与细胞稳态,减轻脑梗死后的脑组织损伤。KEGG 通路富集分析中,色氨酸代谢、破骨细胞分化、NF- κ B 信号通路、造血细胞谱系及急性髓系白血病等通路显著富集,其中 NF- κ B 信号通路作为脑梗死病理进程中炎症反应的核心通路,其过度激活可导致白细胞介素-6、肿瘤坏死因子- α 等促炎因子大量释放,加剧血脑屏障破坏与神经炎症^[20];色氨酸代谢紊乱则会导致 5-羟色胺等神经递质失衡,进一步加重脑缺血后的神经功能障碍。丹参核心靶点对这些通路的调控,提示其可能通过抑制炎症级联反应、改善代谢失衡及调节免疫细胞功能,实现对脑梗死的多维度治疗效应。

SHAP 分析与 ROC 评估为核心靶点的可靠性与临床价值提供了量化依据。特征重要性分析显示,CTSK、S1PR3 与 BCL2A1 为最关键的调控基因。CTSK 作为半胱氨酸蛋白酶家族成员,可通过降解炎症基质蛋白抑制脑梗死后的炎症细胞浸润^[21];S1PR3 作为鞘氨醇-1-磷酸受体,其表达水平与血管内皮完整性及免疫细胞迁移密切相关,低表达可减少中性粒细胞向缺血脑组织募集^[22];BCL2A1 作为抗凋亡基因,高表达能通过抑制线粒体凋亡通路改善神经元存活^[23]。特征依赖分析显示,CTSK、S1PR3 表达量与 SHAP 值呈负相关,BCL2A1 则呈正相关,且 CTSK 与 S1PR3 存在表达交互作用(CTSK 高表达+S1PR3 低表达时丹参疗效增强),提示丹参可能

通过调控这些基因的表达模式重塑脑梗死病理状态下的分子网络。ROC 曲线分析显示,7 个核心基因的 AUC 值均高于 0.85,其中 CTSK 与 S1PR3 诊断效能最优,证实其可作为丹参治疗脑梗死的潜在分子标志物,为临床疗效评估提供客观指标。

分子对接验证进一步证实丹参核心成分与靶点的作用基础。以丹参酮 II_A 为配体,与 7 个核心靶点进行分子对接,结合能均 ≤ -6.0 kcal/mol,其中 KCNE1-丹参酮 II_A、CAMK1-丹参酮 II_A 和 CTSK-丹参酮 II_A 复合物结合能最低,提示这些组合具有优异的结合稳定性。KCNE1 作为电压门控钾通道调节蛋白,功能异常会导致脑梗死后神经元兴奋性紊乱^[24];CAMK1 参与钙信号通路调控,对维持神经元突触可塑性及存活至关重要,丹参酮 II_A 与这些靶点的稳定结合,可能是其发挥神经保护作用的关键分子机制,也验证了网络药理学预测结果的可靠性^[25]。

本研究仍存在局限,活性成分与靶点筛选依赖现有数据库,可能存在未被收录的潜在成分或靶点;生物信息学与机器学习分析结果需通过细胞实验(如神经元细胞缺氧/复氧模型)与动物模型(如大鼠大脑中动脉栓塞模型)进一步验证;未深入探讨核心靶点在脑梗死不同病程阶段(急性期、恢复期)的作用差异,也未分析丹参与其他药物联用的协同效应。未来研究可围绕核心成分-靶点组合开展体外验证,结合临床样本分析优化诊断标志物体系,为丹参的临床应用提供更充分的实验支撑。

综上所述,本研究通过多技术整合,明确丹参可通过丹参酮 II_A 等核心成分,靶向调控 CTSK、S1PR3、BCL2A1 等关键靶点,干预 NF- κ B 信号通路、色氨酸代谢等核心通路,发挥抗炎、抗凋亡、调节免疫微环境等作用,不仅为脑梗死治疗提供了新的机制阐释,也为中药的现代化研究提供了“数据库筛选-计算模拟-功能验证”的完整技术框架。

利益冲突 所有作者均声明不存在利益冲突

参考文献

- [1] Feigin V L, Norrving B, Mensah G A. Global burden of stroke [J]. *Circ Res*, 2017, 120(3): 439-448.
- [2] Virani S S, Alonso A, Benjamin E J, et al. Heart disease and stroke statistics-2020 update: A report from the American heart association [J]. *Circulation*, 2020, 141(9): e139-e596.
- [3] 中华医学会神经病学分会, 中华医学会神经病学分会脑血管病学组. 中国脑血管病一级预防指南 2019 [J].

- 中华神经科杂志, 2019, 52(9): 684-709.
- [4] Chang Y, Kim B M, Bang O Y, *et al.* Rescue stenting for failed mechanical thrombectomy in acute ischemic stroke: A multicenter experience [J]. *Stroke*, 2018, 49(4): 958-964.
- [5] Grossberg J A, Rebello L C, Haussen D C, *et al.* Beyond large vessel occlusion strokes: Distal occlusion thrombectomy [J]. *Stroke*, 2018, 49(7): 1662-1668.
- [6] 马鹏珍, 谢晓霞, 庞华鑫, 等. 基于数据整合及网络医学的参麦注射液治疗低灌注性脑梗死的有效性评价 [J]. *中草药*, 2025, 56(20): 7485-7497.
- [7] Kaesmacher J, Abdullayev N, Maamari B, *et al.* Safety and angiographic efficacy of intra-arterial fibrinolytics as adjunct to mechanical thrombectomy: Results from the INFINITY registry [J]. *J Stroke*, 2021, 23(1): 91-102.
- [8] Levine S R, Khatri P, Broderick J P, *et al.* Review, historical context, and clarifications of the NINDS rt-PA stroke trials exclusion criteria: Part 1: Rapidly improving stroke symptoms [J]. *Stroke*, 2013, 44(9): 2500-2505.
- [9] Zhang R L, Chopp M, Chen H, *et al.* Temporal profile of ischemic tissue damage, neutrophil response, and vascular plugging following permanent and transient (2H) middle cerebral artery occlusion in the rat [J]. *J Neurol Sci*, 1994, 125(1): 3-10.
- [10] Wu D H, Lyu Y, Zhong P, *et al.* Human urinary kallidinogenase promotes good recovery in ischemic stroke patients with level 3 hypertension [J]. *Brain Behav*, 2017, 7(8): e00752.
- [11] Yoshimura S, Uchida K, Daimon T, *et al.* Randomized controlled trial of early versus delayed statin therapy in patients with acute ischemic stroke: ASSORT trial (administration of statin on acute ischemic stroke patient) [J]. *Stroke*, 2017, 48(11): 3057-3063.
- [12] 任雨贺, 田静, 刘淑莹, 等. 2015 年版《中国药典》(一部)含丹参成方制剂的归类分析 [J]. *中成药*, 2018, 40(12): 2770-2775.
- [13] 单晓晓, 洪帮振, 刘洁, 等. 丹参化学成分、药理作用、临床应用的研究进展及质量标志物的预测分析 [J]. *中国中药杂志*, 2021, 46(21): 5496-5511.
- [14] Xu M M, Zhou H, Hu P, *et al.* Identification and validation of immune and oxidative stress-related diagnostic markers for diabetic nephropathy by WGCNA and machine learning [J]. *Front Immunol*, 2023, 14: 1084531.
- [15] Xu J Z, Zhang P, Chen Y, *et al.* Sodium tanshinone II_A sulfonate ameliorates cerebral ischemic injury through regulation of angiogenesis [J]. *Exp Ther Med*, 2021, 22(4): 1122.
- [16] Cai M, Guo Y X, Wang S Q, *et al.* Tanshinone II_A elicits neuroprotective effect through activating the nuclear factor erythroid 2-related factor-dependent antioxidant response [J]. *Rejuvenation Res*, 2017, 20(4): 286-297.
- [17] Wang L, Xiong X, Zhang X, *et al.* Sodium tanshinone II_A sulfonate protects against cerebral ischemia-reperfusion injury by inhibiting autophagy and inflammation [J]. *Neuroscience*, 2020, 441: 46-57.
- [18] Fan X H, Chen H P, Xu C, *et al.* S1PR3, as a core protein related to ischemic stroke, is involved in the regulation of blood-brain barrier damage [J]. *Front Pharmacol*, 2022, 13: 834948.
- [19] Yu F, Feng X J, Li X, *et al.* Association of plasma metabolic biomarker sphingosine-1-phosphate with cerebral collateral circulation in acute ischemic stroke [J]. *Front Physiol*, 2021, 12: 720672.
- [20] Sifi N, Martin-Eauclaire M F, Laraba-Djebari F. K⁺ channel blocker-induced neuroinflammatory response and neurological disorders: Immunomodulatory effects of astaxanthin [J]. *Inflamm Res*, 2016, 65(8): 623-634.
- [21] Lin S Q, Wang T, Zuo C J. Cathepsin K (CTSK) in inflammatory and immune-mediated diseases [J]. *Immunol Invest*, 2025: 1-23.
- [22] Matuskova H, Porschen L T, Matthes F, *et al.* Spatiotemporal sphingosine-1-phosphate receptor 3 expression within the cerebral vasculature after ischemic stroke [J]. *iScience*, 2024, 27(6): 110031.
- [23] Vogler M. BCL2A1: The underdog in the BCL2 family [J]. *Cell Death Differ*, 2012, 19(1): 67-74.
- [24] Clancy S M, Chen B H, Bertaso F, *et al.* KCNE1 and KCNE3 beta-subunits regulate membrane surface expression of Kv12.2 K⁺ channels *in vitro* and form a tripartite complex *in vivo* [J]. *PLoS One*, 2009, 4(7): e6330.
- [25] Li S Y, Ding H F, Li Q, *et al.* Association between plasma proteome and pulmonary heart disease: A two-stage Mendelian randomization analysis [J]. *Clin Respir J*, 2024, 18(6): e13775.

【责任编辑 金玉洁】