

DOI:10.13288/j.11-2166/r.2024.21.013

基于生物信息学和分子动力学的中药治疗原发性骨髓纤维化作用机制研究

郭家缘¹, 新吉乐², 张曼¹, 刘洺欣¹, 刘景文¹, 苏雅静¹, 石慧慧¹,
郭珏¹, 刘文卿¹, 魏凯路², 宋亚琳^{3,4}, 马秋玲^{1,3,4}✉

1. 河南中医药大学第二临床医学院, 河南省郑州市金水区东风路六号, 450002; 2. 河南中医药大学药学院; 3. 河南省中医院/河南中医药大学第二附属医院; 4. 河南省中西医结合防治血液病工程研究中心

[摘要] **目的** 利用生物信息学和分子动力学方法探讨中药治疗原发性骨髓纤维化(PMF)的分子机制。**方法** 通过检索中国知网数据库(CNKI)、中文科技期刊数据库(CCD)和中国学术期刊数据库(CSPD)相关文献, 发掘整理1985—2024年间用于PMF治疗的高频中药, 利用TCMSP、SwissTargetPrediction数据库及相关文献检索高频中药的主要活性成分及其相关靶点; GEO数据库下载PMF的基因表达谱数据集GSE44426和GSE124281, 并运用R软件进行数据归一化处理 and 差异表达基因筛选; 运用加权基因共表达网络分析(WGCNA)获取关键模块枢纽基因。取中药活性成分靶点、差异表达基因和关键模块枢纽基因的共同交集基因, 运用Cytoscape3.9.2软件生成目标网络, 通过拓扑分析生成核心靶点网络, 通过GO和KEGG通路富集分析选取关键通路, 并从String数据库获得蛋白相互作用关系, 构建药物-成分-靶点网络和蛋白质互作网络(PPI)关系图。利用Discovery Studio 2020软件进行分子对接, GROMACS程序进行分子动力学模拟。**结果** 共搜集21首处方, 涉及121种中药, 用药频次 ≥ 10 次的中药有9种, 由高到低排序依次为丹参、黄芪、白术、当归、党参、甘草、白芍、茯苓、熟地黄。9种高频中药共获得98种活性成分和1125个潜在靶点。GSE44426和GSE124281数据集共筛选24个基因样本, 其中健康对照组(14例)和PMF组(10例), 共鉴定出319个PMF组与健康对照组之间的差异表达基因, 包含122个上调基因, 197个下调基因。WGCNA共筛选24个共表达模块基因, 发现与PMF发病密切相关的5个模块分别为MEpink、MEdarkred、MEblack、MEgrey、MEturquoise, 包含7112个关键模块枢纽基因。GO和KEGG通路富集分析显示, 上述高频中药治疗PMF主要涉及脂质和动脉粥样硬化通路, 包含热休克蛋白90家族(HSP90AA1、HSP90AB1)、酪氨酸激酶(SRC)、丝裂原活化蛋白激酶(MAPK1)、白细胞介素10(IL-10)和白细胞介素1 β (IL-1 β)6个核心靶点。从药物-成分-靶点网络找到靶向于这6个核心靶点的中药成分共7种活性成分, 包括毛蕊花糖苷、毛蕊异黄酮、山柰酚、木犀草素、柚皮素、槲皮素和茯苓酸。分子对接和分子动力学分析显示, 关键中药为熟地黄、黄芪、白芍、丹参、甘草和茯苓, 7个活性成分中毛蕊异黄酮与HSP90AB1结合最稳定。**结论** 治疗PMF的主要中药可能为熟地黄、黄芪、白芍、丹参、甘草和茯苓, 有效成分包括毛蕊花糖苷、毛蕊异黄酮、山柰酚、木犀草素、柚皮素、槲皮素、茯苓酸, 相关靶点为HSP90AA1、HSP90AB1、SRC、MAPK1、IL-10、IL-1 β , 最关键通路为脂质和动脉粥样硬化通路。

[关键词] 原发性骨髓纤维化; 生物信息学; 分子动力学; 分子对接; 中药

原发性骨髓纤维化(primary myelofibrosis, PMF)是恶性骨髓增殖性肿瘤, 其发病更具恶性和

侵袭性^[1]。约20%的PMF患者会出现白血病进展和合并症, 血栓、出血、并发症、心血管事件和白血病转化是患者的主要死因^[2]。根据PMF临床表现, 其可归属于中医学“虚劳”“髓枯”“癥瘕”“积聚”等范畴。已有研究表明, 中医药治疗可以

基金项目: 河南省中医药科学研究专项(2024ZY1010); 河南省2022年科技发展计划(222102310353)

✉ 通讯作者: ling93317@163.com

缓解 PMF 全身症状, 提高患者生存质量^[3], 但其作用机制尚不明确。

生物信息学是结合生物学、计算机科学、信息技术和数学的交叉学科, 不仅能够处理大规模的基因组、转录组和蛋白质组数据, 还可预测生物分子的结构和功能, 为理解生物学现象提供新视角^[4]。分子动力学模拟可在原子层面上模拟生物分子的运动和相互作用, 通过模拟分子在生理条件下的行为, 揭示蛋白质折叠、配体结合、分子组装等过程的动态特性^[5]。加权基因共表达网络分析(WGCNA)是生物信息学领域中一种重要分析方法, 能全面且系统地探索基因与给定特征之间的相关性。生物信息学结合分子动力学可从生物网络的角度解读中药治疗疾病的分子机制^[6]。本研究通过对 1985—2024 年中药治疗 PMF 的文献进行数据挖掘, 并基于生物信息学和分子动力学探索中药治疗 PMF 的可能作用机制, 现报告如下。

1 资料与方法

1.1 文献来源

借助中国知网数据库(CNKI)、中文科技期刊数据库(CCD)和中国学术期刊数据库(CSPD), 设置“原发性骨髓纤维化”和“中医药”为关键词, 检索 1985—2024 年间中药治疗 PMF 的相关文献。

1.2 文献纳入标准

1) 研究对象为明确诊断为 PMF 的患者; 2) 干预措施为中药治疗 PMF, 包括具有中医理论支持的经典方剂、自拟经验方或无具体方名但须包含完整的药物组成; 3) 文献类型为临床研究及经验类; 4) 文献中有数据支撑此方剂临床用药有效, 且可使症状得到改善。

1.3 文献排除标准

1) 研究对象合并其他疾病; 2) 同一方剂进行不同的研究或重复发表的文献, 仅保留发表年限距今最近的 1 篇; 3) 干预措施仅为单味中药或其提取物, 或合并中药外用、针灸、推拿等治疗方式。

1.4 中药规范化及高频中药筛选

参照《中华人民共和国药典》^[7]对方剂中中药名称统一进行规范化处理, 如“山棱”规范为“三棱”, “田七”规范为“三七”等。双人双机分别独立检索并输入数据, 建立数据库, 运用 Microsoft Excel 2020 软件进行统计, 高频中药定义为用药频次 ≥ 10 的中药。

1.5 高频中药活性成分和潜在靶点的获取

高频中药的活性成分来源于中药系统药理学数据库和分析平台(TCMSP, <https://old.tcmsp-e.com/tcmsp.php>), 设置口服生物利用度(OB) $\geq 30\%$ 、类药性(DL) ≥ 0.18 作为活性成分的筛选标准。从 PubChem 数据库(<https://pubchem.ncbi.nlm.nih.gov>)中获取中药活性成分的 Isomeric SMILES 结构格式, 并上传到 SwissTargetPrediction 数据库(<http://swisstargetprediction.ch>), 预测活性成分的潜在靶点。利用 Cytoscape 3.9.2 软件对高频中药、活性成分和潜在靶点构建药物-成分-靶点网络。

1.6 基于生物信息学的数据分析

1.6.1 数据采集 从高通量基因表达数据库(GEO, <https://www.ncbi.nlm.nih.gov/geo>)中检索 PMF 患者的两个微阵列数据集 GSE44426 (包括 PMF 患者 6 例和健康对照者 6 例) 和 GSE124281 (包括 PMF 患者 6 例和健康对照者 8 例), 两个数据集的样本均来自骨髓。将两个数据集的基因数据进行合并, 其中 GSE124281 数据集 PMF 样本 5 和样本 6 的基因表达量过少, 故 GSE124281 数据集中仅收录 PMF 样本 1-4 的第一次测量数据; GSE44426 数据集收录全部 PMF 样本的测量数据。

1.6.2 数据归一化处理和差异表达基因筛选 对微阵列数据集 GSE44426 (12 例) 和 GSE124281 (12 例) 进行标准化处理, 将标准化后的数据合并, 合并后的基因表达谱数据包括健康对照组 (14 例) 和 PMF 组 (10 例), 通过 R 软件“sva”包的“ComBat”函数去除数据的批次效应, 进行数据归一化处理, 排除不合格样品。使用 R 软件“limma”包处理基因表达基质, 识别 PMF 组和健康对照组之间的差异表达基因(DEGs)。选取 $P < 0.05$ 且 $|\log_2(\text{FoldChange})| > 0.5$ 的基因标记为上调或下调基因。

1.6.3 WGCNA WGCNA 可以引入加权的相关系数, 构建共表达的基因模块, 广泛应用于候选生物标志物的识别或治疗靶点的筛选中。利用 R 语言的“WGCNA”包计算基因之间的加权相关系数, 构建加权基因共表达网络; 使用选定的软阈值参数和动态树切割算法, 对具有相似表达模式的基因进行层次聚类分析, 识别出模块。采用皮尔逊相关系数, 将模块特征基因与 GEO 数据库收集的疾病分期之间的相关性进行计算, 绘制聚类模块特征热图。利用拓扑重叠矩阵计算模块之间的相似性, 筛选关键模块枢纽基因。

1.6.4 核心靶点的筛选 取中药活性成分靶点、DEGs 和关键模块枢纽基因的共同交集基因, 上传到 String 数据库 (<https://cn.string-db.org>), 采用 Cytoscape 3.9.2 软件的 CytoNCA 内置模块生成目标网络, 通过拓扑分析生成核心靶点网络。

1.6.5 运用 GO 和 KEGG 通路富集分析 取中药活性成分靶点、DEGs 和关键模块枢纽基因的共同交集基因, 通过 GO 和 KEGG 通路富集分析探索共同交集基因的生物学途径和潜在功能。利用微生信平台 (<http://www.bioinformatics.com.cn>) 对共同交集基因的 GO 和 KEGG 通路富集分析结果进行可视化呈现, 并根据分析结果选取关键通路。

1.6.6 构建蛋白质互作网络 (PPI) 将关键通路的相关靶点上传到 String 数据库, 生成目标 PPI 网络。然后将生成的 PPI 网络导入 Cytoscape 3.9.2 软件中, 进一步利用 CytoHubba 内置模块生成主要目标网络, 通过拓扑分析从主要目标网络生成核心目标网络。

1.7 分子对接

为验证核心靶点的准确性, 将作用于关键通路上游的核心靶点与其对应的活性成分进行分子对接。从 PubChem 数据库确定活性成分的化合物名称、分子量和 2D 结构, 从 RCSB PDB 数据库 (<https://www1.rcsb.org>) 下载候选靶点的蛋白质晶体结构。运用 Discovery Studio 2020 软件进行分子对接, 使用 PyMol 2.4.1 软件观察和分析对接结果, 并将结果可视化处理。以结合分数值作为分析亲和力大小的指标, 一般认为结合分数值越低, 配体与受体两者结合越紧密、越稳定。

1.8 分子动力学研究

分子动力学模拟是一种计算和预测化合物稳定性的方法, 是研究分子体系的重要手段之一, 对于了解蛋白质-配体复合物和分子间相互作用的稳定性有重要作用。利用 GROMACS 程序进行分子动力学模拟轨迹, 水模型选择 TIP3P 模型。对模拟体系进行溶剂化、离子平衡后, 分别将体系升温至 300 K 和 1.01325 bar, 再对该体系进行 50 ns 的分子动力学模拟, 以了解其偏差、波动和分子间相互作用。

2 结果

2.1 高频中药筛选结果

共搜集 21 首方剂, 涉及 121 种中药, 总使用频次为 384 次, 用药频次 ≥ 10 次的中药有 9 种, 由高到低排序依次为丹参 (16 次, 4.17%)、黄芪 (15

次, 3.91%)、白术 (14 次, 3.65%)、当归 (14 次, 3.65%)、党参 (13 次, 3.39%)、甘草 (13 次, 3.39%)、白芍 (11 次, 2.86%)、茯苓 (11 次, 2.86%)、熟地黄 (10 次, 2.60%)。

2.2 药物-成分-靶点网络的构建结果

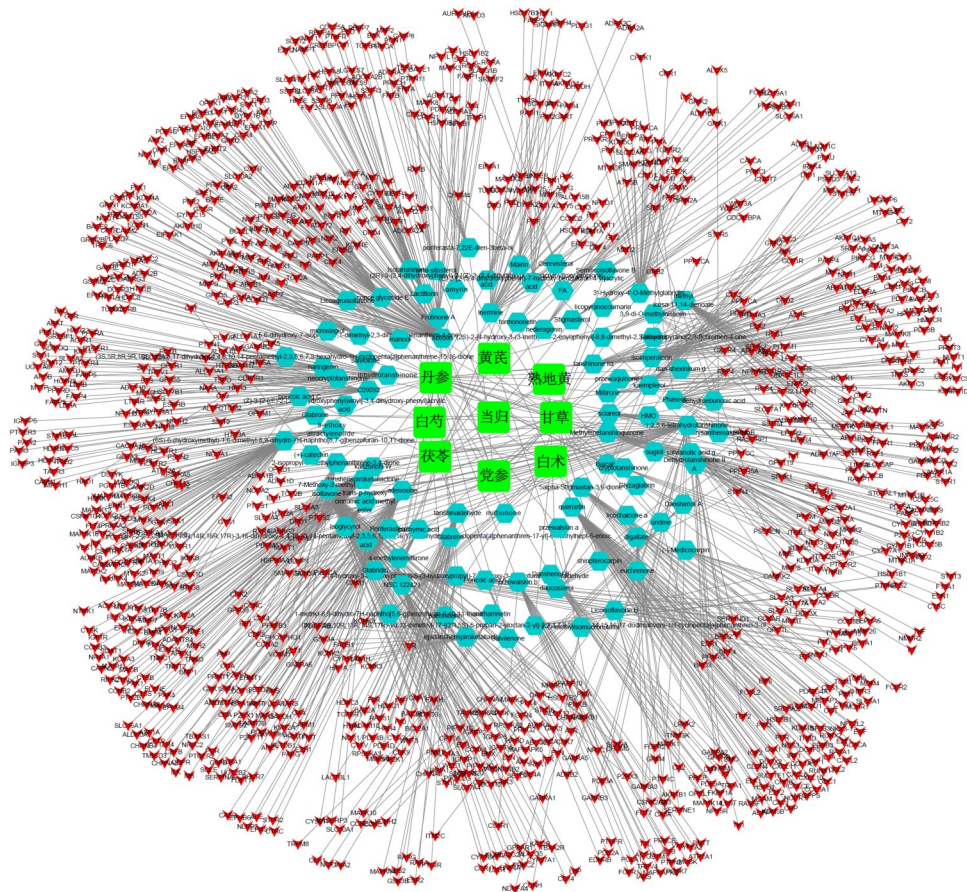
9 种高频中药共获得 98 种活性成分和 1125 个潜在靶点, 将高频中药的活性成分与潜在靶点相对应, 构建药物-成分-靶点网络图 (见图 1)。

2.3 PMF 的生物信息学研究结果

2.3.1 基因表达谱数据的预处理 对健康对照组 (14 例) 和 PMF 组 (10 例) 的基因表达谱数据进行归一化处理, 结果显示这 24 个样本均具有良好的芯片质量, 可用于进一步研究。通过 “ComBat” 函数算法对两个数据集的数据进行合并, 共得到 13240 个基因, 去除数据间的批次效应, 绘制归一化和背景校正值的箱线图。图 2 示, 每个样本中基因表达的中位数基本在一条直线上, 表明预处理后的数据作为标准数据, 可以在后续研究中进行分析。

2.3.2 DEGs 识别结果 共鉴定出 319 个 PMF 组与健康对照组之间的 DEGs, 包括 122 个上调基因, 197 个下调基因。图 3 展示了前 20 个上调基因 (包括 GDF15、TNS1、GSPT1、ADAMTS12、PITX2、PTPRR、EN1、MGAM、PTGS1、MYL4、SCD、RAP1GAP、WFDC1、S1PR5、PTPRB、TRPC4、PDE11A、OXTR、IL1B、IFI27) 和下调基因 (包括 HAND2、HLA-DRB5、CHI3L2、CD70、BARX1、SMOC2、C11orf21、GSC、NPTX2、CP、STEAP4、GLDN、OGN、FGF13、SOX5、CX3CL1、CHI3L1、FGL2、HOXA13、DIO3)。

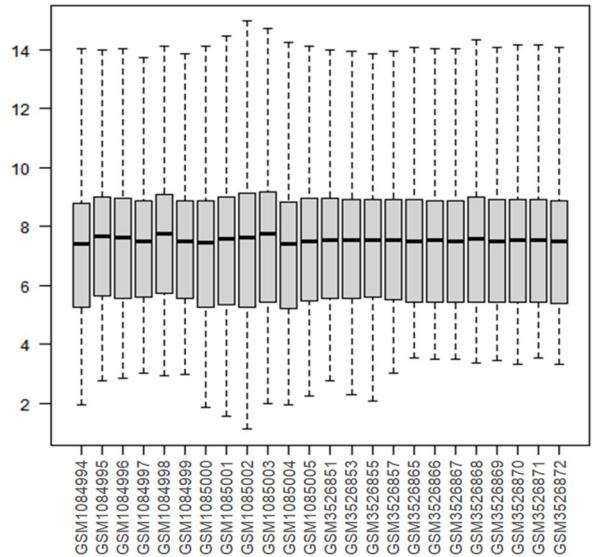
2.3.3 WGCNA 网络的构建 将 24 个样本共计 13240 个基因构建共表达网络。分层聚类分析显示未离群样本。当软阈值 (β) 达到 18 时, 无标度拓扑拟合指数 (R^2) 超过 0.90, 表明拓扑模型的拟合程度较高, 且平均连通性趋近于 0, 故最佳 β 选择 18。通过基于共表达关系的层次聚类分析, 得到 WGCNA 网络。对具有相似表达模式的基因进行聚类, 得到 24 个不同的共表达模块基因, 包括 MElightcyan、MEgrey、MEagenta、MEpurple、MEgreenyellow、MElightyellow、MEblue、Darkturquoise、MEpink、MEDarkred、MERoyalblue、METan、MEblack、MEcyan、MEgreen、MERed、METurquoise、Midnightblue、MESalmon、MEDarkgrey、MEDarkgreen、MEyellow、MEbrown、MElightgreen。利用模块-性状关联分析, 检测模块与疾病分期之间的相关性, 通过拓扑



注：红色箭头代表潜在作用靶点，蓝色正六边形代表中药活性成分，绿色矩形代表高频中药。

图1 中药治疗原发性骨髓纤维化文献中高频中药的药物-成分-靶点网络图

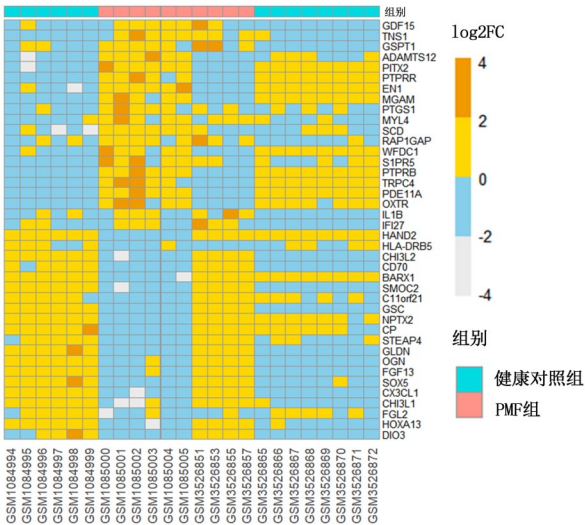
Figure 1 Drug-ingredient-target network diagram of high-frequency Chinese medicinal herbs in the literature on the treatment of primary myelofibrosis



注：横坐标 GSM 编号代表 GSE 数据集内单个样本；纵坐标代表基因表达量。

图2 健康对照组和原发性骨髓纤维化组的基因表达谱归一化处理结果

Figure 2 Normalization of gene expression profiles in healthy control group and primary myelofibrosis group

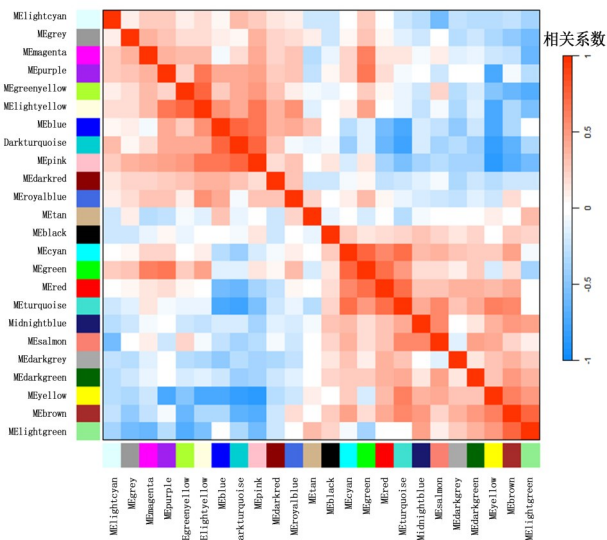


注：横坐标 GSM 编号代表 GSE 数据集内单个样本；右侧为排名前 20 个上调基因和下调基因名称，蓝色表示下调基因，红色表示上调基因；PMF，原发性骨髓纤维化。

图3 健康对照组和原发性骨髓纤维化组差异表达基因分析结果

Figure 3 Differentially expressed genes between the healthy control group and primary myelofibrosis group

重叠矩阵模块邻接热图（见图 4）分析模块相关性，热图中颜色越深表示模块之间的相关性越高。共发现与 PMF 发病密切相关的 5 个模块分别为 MEpink、MEdarkred、MEblack、MEgrey、MEturquoise，包含 7112 个基因。因此，通过 WGCNA 确定这 7112 个基因作为关键模块枢纽基因。



注：热图颜色代表皮尔逊相关系数，红色表示正相关，蓝色表示负相关。

图 4 健康对照组和原发性骨髓纤维化组共表达基因的拓扑重叠矩阵模块邻接热图

Figure 4 Adjacency Heatmap of the topological overlap matrix of co-expressed genes in the healthy control group and primary myelofibrosis group

2.3.4 核心靶点筛选结果 通过对 1125 个活性成分的潜在靶点、319 个 DEGs 和 7112 个关键模块枢纽基因的共同交集基因分析，确定治疗 PMF 的高频中药共有 496 个共同交集基因（即潜在靶点）。其中，有 30 个靶点同时出现在 3 个部分中，分别为 FGR、MAPK10、LDLR、NCF1、IL10、E2F2、CXCL10、CDKN1A、BRD3、MAP2K3、CCNB3、MPO、HK1、QPCT、MAOB、SYK、TGFB2、JUN、OXTR、CD14、NAAA、LYZ、IL1B、P2RX7、FASN、LBP、CTSS、SERPINE1、EPHB6、PTGS1。其中 MAPK10、BRD3、QPCT、MAOB、OXTR、NAAA、EPHB6 为游离靶点，予删除。剩余 23 个核心靶点 PPI 网络图见图 5，PPI 网络中节点越大、颜色越深表示度值越大，在网络中的重要性也越高。由图 5 可以看出，IL1B、IL10 和 JUN 可能是高频中药作用于 PMF 的重要靶点。

2.3.5 GO 和 KEGG 通路富集分析结果 对 496 个共同交集基因进行 GO 通路富集分析，PMF 发病涉及 825 个分子功能、5345 个生物过程和 465 个细胞

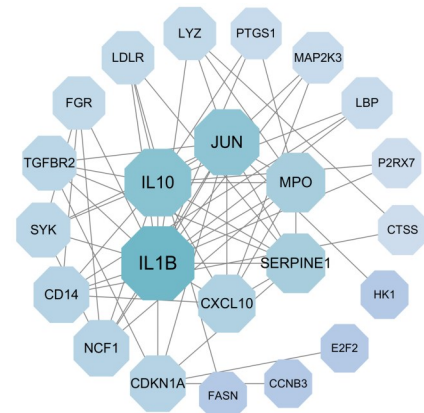


图 5 高频中药治疗原发性骨髓纤维化的 23 个核心靶点的蛋白质互作网络（PPI）图

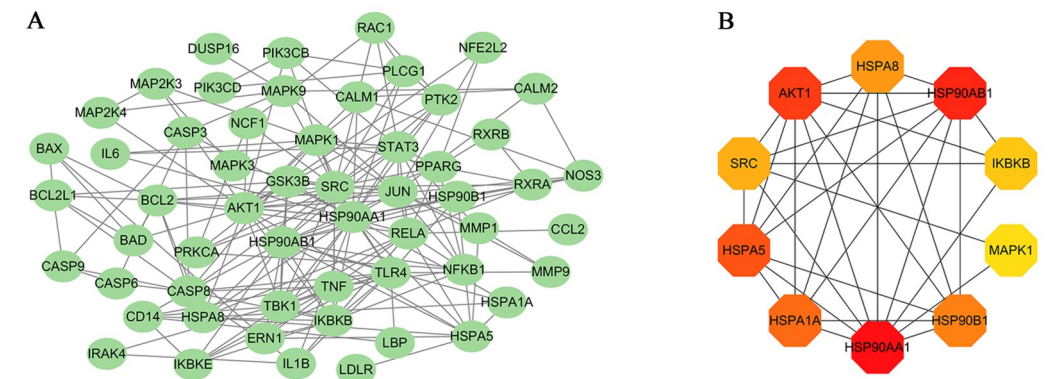
Figure 5 Protein interaction network (PPI) of 23 core targets of high-frequency Chinese medicinal herbs in the treatment of primary myelofibrosis

组成，排名前 3 位的分子功能包括蛋白丝氨酸/苏氨酸激酶活性、蛋白酪氨酸激酶活性、蛋白丝氨酸/苏氨酸/酪氨酸激酶活性；排名前 3 位的生物过程包括蛋白质自磷酸化、肽基丝氨酸修饰、肽基丝氨酸磷酸化；排名前 3 位的细胞组成集中在膜区、膜筏、膜微域。PMF 发病还涉及 305 个 KEGG 通路，排名前 3 位的通路包括脂质和动脉粥样硬化、乙型肝炎、缺氧诱导因子 1（HIF-1）信号通路。

2.3.6 最关键通路蛋白 PPI 构建结果 根据 KEGG 通路富集分析结果，对 PMF 发病涉及的最关键通路脂质和动脉粥样硬化通路的 56 个蛋白绘制 PPI 图（图 6A）。其中排名前 10 位的基因包括 HSP90AA1、HSP90AB1、HSPA1A、HSPA5、SRC、AKT1、HSPA8、IKBKB、MAPK1 和 HSP90B1（图 6B）。有研究^[8-9]显示，热休克蛋白 90α 家族 A 类成员 1（HSP90AA1）、热休克蛋白 90α 家族 B 类成员 1（HSP90AB1）、酪氨酸激酶（SRC）、丝裂原活化蛋白激酶（MAPK1）、白细胞介素 1β（IL-1β）和白细胞介素 10（IL-10）与 PMF 发病密切相关。因此，从药物-成分-靶点网络找到靶向于这 6 个核心靶点的中药成分共 7 种活性成分，包括毛蕊花糖苷、毛蕊异黄酮、山柰酚、木犀草素、柚皮素、槲皮素和茯苓酸。进一步通过分子对接分析这些活性成分与核心靶点之间的相互作用关系。

2.4 分子对接结果

表 1 示，共涉及关键中药 6 种，分别为黄芪、熟地黄、白芍、丹参、甘草、茯苓，有 7 个活性成分和核心靶点结合较好（结合分数 < -5.6 kcal/mol），其中 HSP90AB1 与毛蕊异黄酮的结合活性最好，而



注：A，脂质和动脉粥样硬化通路的 56 个蛋白 PPI 网络；B，脂质和动脉粥样硬化通路的核心目标网络，颜色由黄变红表示度值增大，在网络中的重要性升高。

图 6 中药治疗原发性骨髓纤维化的脂质和动脉粥样硬化通路蛋白质互作网络（PPI）图
Figure 6 Protein interaction network (PPI) of lipid and atherosclerosis pathways in the treatment of primary myelofibrosis with Chinese medicinal herbs

表 1 中药治疗原发性骨髓纤维化的核心靶点及中药活性成分分子对接结果

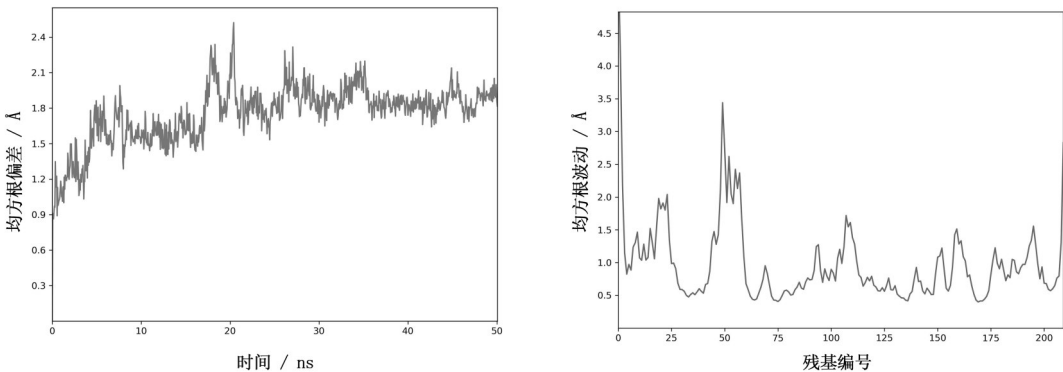
核心蛋白 靶点	PDB 数据库 编号	活性成分	中药	结合分数 /kcal·mol ⁻¹
HSP90AA1	2YI7	毛蕊花糖苷	熟地黄	-9.714
HSP90AB1	1UYM	毛蕊异黄酮	黄芪	-11.200
SRC	1Y57	山柰酚	白芍	-5.886
MAPK1	7NR8	木犀草素	丹参	-6.496
MAPK1	7NR8	柚皮素	甘草	-6.471
MAPK1	7NR8	槲皮素	黄芪	-9.573
IL-1β	6Y8M	槲皮素	甘草	-3.821
IL-1β	6Y8M	茯苓酸	茯苓	3.614
IL-10	2H24	木犀草素	丹参	-7.858
IL-10	2H24	槲皮素	黄芪	-1.995

注：HSP90AA1，热休克蛋白 90α 家族 A 类成员 1；HSP90AB1，热休克蛋白 90α 家族 B 类成员 1；SRC，酪氨酸激酶；MAPK1，丝裂原活化蛋白激酶 1；IL-1β，白细胞介素 1β；IL-10，白细胞介素 10。

IL-1β 与茯苓酸不结合。毛蕊异黄酮主要与 HSP90AB1 上的色氨酸 162 (Trp162) 和亮氨酸 103 (Leu103) 残基形成 2 个氢键，共有 9 个残基通过疏水作用与蛋白结合。上述结构可能是受体与活性成分发挥防治 PMF 作用的分子机制，需进一步进行分子动力学模拟分析。

2.5 分子动力学模拟结果

分子对接结果显示，毛蕊异黄酮和 HSP90AB1 的亲合力较强，故进一步通过分子动力学模拟对毛蕊异黄酮-HSP90AB1 复合物的稳定性进行研究。使用均方根偏差 (RMSD) 和均方根波动 (RMSF) 评估复合物的稳定性。RMSD 曲线可反映蛋白质构象的波动。图 7A 示，毛蕊异黄酮-HSP90AB1 复合物的 RMSD 曲线在前期有一定的波动，30 ns 后趋于稳定，表明两者结合相对稳定。RMSF 曲线可反映蛋白质氨基酸残基的波动。图 7B 示，毛蕊异黄酮与 HSP90AB1 在溶液状态下依然能够稳定结合。



注：A，均方根偏差曲线；B，均方根波动曲线。

图 7 毛蕊异黄酮与热休克蛋白 90α 家族 B 类成员 1 结合的分子动力学模拟图
Figure 7 Molecular dynamics simulation of the binding of calycosin and heat shock protein 90α family class B member 1

3 讨论

中医学认为, PMF 属本虚标实之证, 与肝、脾、肾三脏关系密切, 血瘀为致病关键因素^[3]。PMF 患者早期多因胁下积块就诊, 表现为胁肋胀痛、胁下积块、烦躁易怒、善太息等, 病位在肝^[10]。《金匮翼·积聚统论》云:“凡忧思郁怒, 久不得解者, 多成此疾。”情志不畅, 肝气不舒, 气机阻滞, 气血津液运行失常, 内生痰瘀互结, 戕害机体, 日久可发为癥瘕、积聚等病证。PMF 后期出现病态造血, 髓外造血导致脾肿大、出血、骨痛等症状。脾为后天之本, 肝木横克脾土而致脾虚湿盛, 久则可致虚劳。肾为先天之本, 诸邪郁久化热, 热瘀血积, 日久伤阴, 损伤肾精则骨枯髓虚。本研究最终筛选出黄芪、熟地黄、丹参、茯苓、白芍和甘草 6 味治疗 PMF 的关键中药。其中黄芪益气健脾; 熟地黄补血滋阴, 益精填髓; 丹参活血祛瘀, 凉血消痛; 茯苓健脾祛湿; 白芍养血柔肝; 甘草缓急止痛, 解癌毒。较契合 PMF 患者肝郁脾虚、血瘀髓虚的病机。

本研究利用生物信息学技术分析了 GEO 数据库的 GSE44426 和 GSE124281 数据集, 通过网络药理学方法检索有 30 个靶点与中药治疗 PMF 相关, 通过构建“中药-成分-靶点”网络、PPI 网络以及 GO、KEGG 通路富集分析, 可以预测中药通过调控 HSP90AA1、HSP90AB1、SRC、MAPK1、IL-10 和 IL-1B 基因发挥治疗 PMF 的作用。其中 HSP90AA1、HSP90AB1 都属于热休克蛋白家族, 在血液系统恶性肿瘤中多呈高表达, 且高表达的热休克蛋白 90 (HSP90) 与急性髓系白血病及骨髓增生异常综合征的预后不良相关^[11]。有研究报道, HSP90 可以通过激活肿瘤坏死因子 α (TNF- α) 通路来诱导炎症反应^[12], 活化的核因子 κ B (NF- κ B) 通过促进白细胞介素、TNF- α 和其他促炎细胞因子的表达来调节炎症反应^[13]。炎症信号的过度激活和炎症细胞因子的过度产生, 可以抑制良性造血并且促进骨髓的纤维化转化^[14]。在 PMF 患者中, 血浆 IL-1 β 、IL-10 水平升高是预后不良的标志物^[15], IL-10 与其受体结合会激活非受体酪氨酸激酶 1 (JAK1) 和酪氨酸激酶 2 (JAK2), 从而导致信号转导和转录激活因子 3 (STAT3) 的招募和磷酸化, STAT3 同源二聚化并转位到细胞核, 起到转录激活剂的作用^[16]。除了 JAK 突变以外, PMF 患者中丝裂原活化蛋白激酶 (MAPK) 和 NF- κ B 信号通路也常被异

常激活^[14]。IL-10 还可以诱导 STAT3 和 NF- κ B 通路之间的串扰, STAT 家族与 NF- κ B 共同定位在 SnaI 启动子区域, IL-10 可能会激活 NF- κ B 信号通路, 激活的 NF- κ B 驱动靶基因表达, 释放 IL-10 进而增强免疫反应^[17]。此外, TNF- α 可以激活 SRC, 同时 SRC 会激活 STAT3 和 MAPK 信号通路^[18]。NF- κ B 的激活也受 MAPK 信号通路的调控, MAPK/NF- κ B 途径被认为是调控炎症反应的潜在重要靶点^[19]。

综上所述, 我们推测在 PMF 中 HSP90AA1、HSP90AB1、SRC、MAPK1、IL-10 和 IL-1 β 的关系网络如下: 1) 高表达的 HSP90 激活 NF- κ B 信号通路, 促进 IL-10、TNF- α 等炎症因子的表达; 2) IL-10 通过 JAK-STAT 途径激活 NF- κ B, TNF- α 通过 SRC 和 MAPK 信号通路激活 STAT3 和 NF- κ B, 进一步促进炎症因子表达。这些通路的异常激活会引起炎症反应和免疫异常, 促进 PMF 的发生和进展。中药熟地黄、黄芪、白芍、丹参、甘草和茯苓的活性成分毛蕊花糖苷、毛蕊异黄酮、山柰酚、木犀草素、柚皮素、槲皮素和茯苓酸可能通过调节 HSP90AA1、HSP90AB1、SRC、MAPK1、IL-10 和 IL-1 β 抑制炎症和免疫反应治疗 PMF。分子对接和分子动力学模拟的结果进一步支持了这一作用机制, 其中毛蕊异黄酮通过与 HSP90AB1 蛋白上的色氨酸 162 (Trp162) 和亮氨酸 103 (Leu103) 残基形成稳定的氢键, 实现了在溶液中的稳定结合。

本研究通过文献数据挖掘及生物信息学相结合的方式, 对治疗 PMF 高频中药与疾病核心蛋白靶点进行研究, 初步阐释了中药治疗 PMF 的主要药物为熟地黄、黄芪、白芍、丹参、甘草和茯苓, 有效成分包括毛蕊花糖苷、毛蕊异黄酮、山柰酚、木犀草素、柚皮素、槲皮素、茯苓酸, 相关靶点为 HSP90AA1、HSP90AB1、SRC、MAPK1、IL-10、IL-1 β , 最关键通路为脂质和动脉粥样硬化通路。此结果为中医药寻找和发现 PMF 的有效治疗和预防策略提供了依据。

参考文献

- [1] MESA RA, VERSTOVSEK S, CERVANTES F, et al. Primary myelofibrosis (PMF), post polycythemia vera myelofibrosis (post-PV MF), post essential thrombocythemia myelofibrosis (post-ET MF), blast phase PMF (PMF-BP): Consensus on terminology by the international working group for myelofibrosis research and treatment (IWG-MRT) [J]. Leuk Res, 2007, 31(6): 737-740.
- [2] TEFFERI A. Primary myelofibrosis: 2023 update on diag-

- nosis, risk-stratification, and management [J]. *Am J Hematol*, 2023, 98(5):801–821.
- [3] 王梦亚, 程志. 原发性骨髓纤维化的中医药治疗探析[J]. *中医研究*, 2016, 29(8):9–11.
- [4] UESAKA K, OKA H, KATO R, et al. Bioinformatics in bioscience and bioengineering: recent advances, applications, and perspectives [J]. *J Biosci Bioeng*, 2022, 134(5):363–373.
- [5] DURRANT JD, MCCAMMON JA. Molecular dynamics simulations and drug discovery[J]. *BMC Biol*, 2011, 9:71. doi:10.1186/1741-7007-9-71.
- [6] LI W, YUAN B, ZHAO Y, et al. Transcriptome profiling reveals target in primary myelofibrosis together with structural biology study on novel natural inhibitors regarding JAK2[J]. *Aging (Albany NY)*, 2021, 13(6):8248–8275.
- [7] 国家药典委员会. 中华人民共和国药典:一部[M]. 北京:中国医药科技出版社, 2020.
- [8] CHORZALSKA A, MORGAN J, AHSAN N, et al. Bone marrow-specific loss of ABL1 induces myeloproliferative neoplasm with features resembling human myelofibrosis [J]. *Blood*, 2018, 132(19):2053–2066.
- [9] WANG F, ZHANG H, HE B, et al. Heat shock protein 90 inhibitors induce cell differentiation via the ubiquitin-dependent aurora kinase A degradation in a MPLW515L mouse model of primary myelofibrosis[J]. *Hematol Oncol*, 2023, 41(3):474–486.
- [10] 刘清池, 马传宝, 李建英, 等. 鳖甲生血丸治疗原发性骨髓纤维化患者 120 例临床观察[J]. *中医杂志*, 2014, 55(8):677–680.
- [11] MJAHEH H, GIRODON F, FONTENAY M, et al. Heat shock proteins in hematopoietic malignancies [J]. *Exp Cell Res*, 2012, 318(15):1946–1958.
- [12] MIAO Z, ZHANG K, BAO R, et al. Th1/Th2 imbalance and heat shock protein mediated inflammatory damage triggered by manganese via activating NF- κ B pathway in chicken nervous system in vivo and in vitro [J]. *Environ Sci Pollut Res Int*, 2021, 28(32):44361–44373.
- [13] XU P, HUANG MW, XIAO CX, et al. Matairesinol suppresses neuroinflammation and migration associated with Src and ERK1/2-NF- κ B pathway in activating BV2 microglia[J]. *Neurochem Res*, 2017, 42(10):2850–2860.
- [14] FISHER DAC, MINER CA, ENGLE EK, et al. Cytokine production in myelofibrosis exhibits differential responsiveness to JAK-STAT, MAP kinase, and NF κ B signaling[J]. *Leukemia*, 2019, 33(8):1978–1995.
- [15] BARONE M, CATANI L, RICCI F, et al. The role of circulating monocytes and JAK inhibition in the infectious-driven inflammatory response of myelofibrosis [J]. *Oncoimmunology*, 2020, 9(1):1782575.
- [16] CORDOBA-MORENO MO, SANTOS GC, MUXEL SM, et al. IL-10-induced STAT3/NF- κ B crosstalk modulates pineal and extra-pineal melatonin synthesis [J]. *J Pineal Res*, 2024, 76(1):e12923.
- [17] BARBOSA LIMA LE, MUXEL SM, KINKER GS, et al. STAT1 - NF κ B crosstalk triggered by interferon gamma regulates noradrenaline-induced pineal hormonal production[J]. *J Pineal Res*, 2019, 67(3):e12599.
- [18] YANG CM, LEE IT, LIN CC. et al. c-Src-dependent MAPKs/AP-1 activation is involved in TNF-alpha-induced matrix metalloproteinase-9 expression in rat heart-derived H9c2 cells[J]. *Biochem Pharmacol*, 2013, 85(8):1115–1123.
- [19] WANG HW, WU T, QI JY, et al. Salidroside attenuates LPS-stimulated activation of THP-1 cell-derived macrophages through down-regulation of MAPK/NF- κ B signaling pathways [J]. *J Huazhong Univ Sci Technolog Med Sci*, 2013, 33(4):463–469.

Mechanism of Action of Chinese Medicinal Herbs in the Treatment of Primary Myelofibrosis based on Bioinformatics and Molecular Dynamics

GUO Jiayuan¹, XIN Jile², ZHANG Man¹, LIU Mingxin¹, LIU Jingwen¹, SU Yajing¹, SHI Huihui¹, GUO Jue¹, LIU Wenqing¹, WEI Kailu², SONG Yalin^{3,4}, MA Qiuling^{1,3,4}

1. The Second Clinical Medical College, Henan University of Chinese Medicine, Zhengzhou, 450002; 2. School of Pharmacy, Henan University of Chinese Medicine; 3. Henan Province Hospital of Traditional Chinese Medicine / The Second Affiliated Hospital of Henan University of Chinese Medicine; 4. Henan Engineering Research Center for the Prevention and Treatment of Hematological Diseases with Integrated Traditional Chinese and Western Medicine

ABSTRACT Objective To explore the molecular mechanism implicated in the treatment of primary myelofibrosis (PMF) using Chinese medicinal herbs (CMH) by bioinformatics and molecular dynamics. **Methods** Data mining was performed to find the high-frequency CMH in treating PMF between the year of 1985 and 2024 by searching CNKI, Chinese Science and Technology Journal Database (CCD), and China Academic Journal Database (CSPD). TCMSP, SwissTargetPrediction and related reports were used to collect the main active ingredients of high-frequency

CMH and their targets. The PMF datasets GSE44426 and GSE124281 were downloaded from GEO database, and R software was used for data normalization and differentially expressed genes (DEGs) screening. Key module hub genes were obtained by weighted gene co-expression network analysis (WGCNA) analysis. The common intersection genes of active ingredient targets, DEGs and key module hub genes of CMH were selected, and the target network was generated using Cytoscape 3.9.2 software. The core target network was generated by topological analysis, while key pathways were selected by GO and KEGG pathway enrichment analysis, and protein interaction relationships were obtained from the String database, so as to construct drug-ingredient-target network and protein interaction network (PPI) relationship diagrams. Discovery Studio 2020 software was used to perform molecular docking, and the GROMACS program was used to perform molecular dynamics simulation. **Results** A total of 21 prescriptions were collected involving 121 herbs. There were 9 herbs with a frequency ≥ 10 times, which were Danshen (*Radix et Rhizoma Salviae Miltiorrhizae*), Huangqi (*Radix Astragali*), Baizhu (*Rhizoma Atractylodis Macrocephalae*), Danggui (*Radix Angelicae Sinensis*), Dangshen (*Radix Codonopsis*), Gancao (*Radix et Rhizoma Glycyrrhizae*), Baishao (*Radix Paeoniae Alba*), Fuling (*Poria*) and Shudihuang (*Radix Rehmanniae Praeparata*) from high- to low-frequency. A total of 98 active ingredients and 1125 potential targets were obtained from 9 high-frequency CMH. GSE44426 and GSE124281 data sets screened out 24 gene samples, including 14 of the healthy control group and 10 of the PMF group, and identified 319 DEGs between the two groups, including 122 up-regulated genes and 197 down-regulated genes. WGCNA screened out 24 co-expression module genes and found that the five modules closely related to the onset of PMF were MEpink, MEdarkred, MEblack, MEgrey, and MEturquoise, involving 7112 key module hub genes. The GO and KEGG enrichment analyses indicated that lipids and the atherosclerosis pathways were mainly involved in the mechanism of above high-frequency CMH in treating PMF, which included six hub protein targets: HSP90AA1, HSP90AB1, SRC, MAPK1, IL1B and IL10. From the drug-ingredient-target network, seven active ingredients of CMH targeting at these six hub targets were found, including verbascoside, verbascos isoflavone, kaempferol, luteolin, naringenin, quercetin and pachymic acid. The molecular docking and molecular dynamics analyses showed that the key CMH were Shudihuang, Huangqi, Baishao, Danshen, Gancao and Fuling, and among the seven active ingredients, calycosin had the highest binding affinity with HSP90AB1. **Conclusion** The main CMH for the treatment of PMF may be Shudihuang, Huangqi, Baishao, Danshen, Gancao and Fuling, and the active ingredients include verbascoside, verbascos isoflavones, kaempferol, luteolin, naringenin, quercetin and pachymic acid. The relevant targets are HSP90AA1, HSP90AB1, SRC, MAPK1, IL-10, and IL-1 β , and the most critical pathways are lipid and atherosclerosis pathways.

Keywords primary myelofibrosis; bioinformatics; molecular dynamics; molecular docking; Chinese medicinals herbs

(收稿日期: 2024-06-25; 修回日期: 2024-08-27)

[编辑: 邓媛]