

基于网络药理学与分子对接探讨玫瑰花治疗乳腺癌的机制

赵云焘 冯日升 张颖 刘获晞 吴文梅*

广东药科大学生命科学与生物制药学院 广东广州 510006

摘要: 目的: 通过网络药理学与分子对接探讨玫瑰花治疗乳腺癌的机制。方法: 利用 TCMSP 和 Swiss Target Prediction 数据库预测其成分的作用靶点, 并通过 GeneCards 数据库获得乳腺癌相关基因靶点。取交后得到玫瑰花抗乳腺癌的潜在靶点。运用 Cytoscape 3.9.1 分析相关靶点的蛋白互作图, 筛选核心靶点并进行分子对接。在 DAVID 数据库中对靶点进行 GO 和 KEGG 分析。结果: 网络药理学筛选的可能核心成分有 3 种: quercetin、Demethoxycapillarisin 和 vitamin-e, 其作用靶点共 215 个, 通过与乳腺癌相关的 2000 个靶点取交, 确定了 105 个玫瑰花抗乳腺癌的作用靶点。KEGG 分析表明交集靶点主要与 FoxO、AMPK 及 PI3K-Akt 等信号通路相关, 通过分类汇总图发现这些交集靶点与癌症治疗通路密切相关。分子对接表明关键靶点 EGFR、ERBB2、ESR1 和 PIK3R1 与玫瑰花之间具有较强的结合作用。结论: 玫瑰花可能通过多靶点、多通路作用于乳腺癌。

关键词: 乳腺癌; 玫瑰花; 作用机制; 网络药理学; 分子对接

乳腺癌 (BC) 是全球女性最常见的癌症之一^[1], 发病率和病死率均位居第一^[2], 约占女性所有恶性肿瘤的三分之一, 其死亡率约占病例总数的 15%^[3]。乳腺癌其错综复杂的发病机制和多样化的临床表现对有效治疗和预防构成了重大障碍^[4]。随着这种疾病的全球发病率持续上升, 必须揭示乳腺癌的多面性, 以制定有效的治疗策略^[5]。

中草药和天然药物的主要优点在于含有多种活性成分, 这些成分可能通过同时作用于多个靶点而产生累加或协同作用^[6]。玫瑰花是蔷薇科植物玫瑰的干燥花蕾, 其花瓣可入药, 味甘微苦, 性温。具有行气解郁、活血止痛的功效。适用于治疗肝胃气痛、食欲不振、恶心呕吐、月经不调以及跌打损伤等症状^[7]。经网络药理学分析, 初步推断玫瑰花对治疗 TNBC 具有一定疗效。槲皮素是玫瑰花的重要成分之一, 可用于治疗慢性支气管炎, 对冠心病和高血压患者也具有辅助治疗作用^[8]。当前的科研领域中, 关于槲皮素在 BC 中的潜在作用机制, 仍然稀缺。

因此, 本研究剖析玫瑰花的抗肿瘤机制, 并发现槲皮素在 BC 中的核心作用靶点及其相关靶标基因表达水平发挥重要作用。通过这一系统的研究, 期望能够揭示槲皮素及其活性成分在 BC 治疗中的新路径, 为开发副作用小、预后效果佳的联合治疗方案提供有力的科学依据。

1. 方法

1.1 玫瑰花活性成分及其对应靶点的预测

利用 TCMSP 数据库根据 Lipinski's 五原则、口服利用度 (OB) $\geq 30\%$, 类药性 (DL) ≥ 0.18 对玫瑰花的活性成分进行筛选。Pubchem 数据库获取有效化合物的化学结构, 获取 smiles 号。借助 Swiss Target Prediction 数据库获取活性成分对应的潜在靶点。

1.2 乳腺癌靶点的获取

以“Breast cancer”为关键词输入 GeneCards 数据库进行检索, 选择“homo sapiens”, 获取乳腺癌相关的靶点基因。

1.3 玫瑰花 - 乳腺癌共同靶点的筛选

通过 Venny 2.1.0 将玫瑰花相关靶点与疾病靶点分别载入 List 1 和 List 2 取交集, 得到 Venny 图。

1.4 玫瑰花治疗乳腺癌的成分靶点网络图分析

将共同靶点导入 Cytoscape 3.9.1, 构建“药物 - 成分 - 靶点”可视化的网络图。然后使用 Network Analyzer 功能对该网络进行相关特征分析, 并利用 CytoNCA 插件进行网络拓扑学分析。网络中每个节点代表药物、有效成分或靶点基因等属性, 节点之间的交互通过连接线表示。

1.5 蛋白质互作 (PPI) 网络图的构建

将共同靶点导入 String 在线工具, 并将相互作用的阈值调整为 0.7, 以获取 PPI 的网络信息。然后保存并下载靶标

基因之间的互作关系结果，并导入 Cytoscape 3.9.1 进行处理，利用 Network Analyzer 进行更精确的分析。

1.6 GO 功能富集及 KEGG 通路富集分析

将共同靶点导入 DAVID 6.8 数据库，设定阈值 $P < 0.05$ ，对靶点进行 GO 和 KEGG 分析，利用微生信网进行可视化，筛选出 degree 值较高的交集靶点。

1.7 靶点基因的生存预测分析

利用 Kaplan–Meier Plotter 生存分析网站，使用 Log–rank 检验，使用工具中的“Breast cancer mRNA (gene chip)”检索交集靶点与乳腺癌患者的总体生存期 (OS) 和无病生存期 (RFS) 之间的关系。

1.8 对 quercetin 和靶点蛋白进行分子对接

分别在 Pubchem 数据库和 RCSB PDB 蛋白质数据库下载 quercetin 的化学结构 MOL2 文件和相关靶点的蛋白质结构 PDB 文件，使用 PyMOL 对蛋白质结构文件进行去配体、去水处理后，把 MOL2 文件和 PDB 文件导入 AutoDock 进行分子对接，得到分子对接结果后使用 PyMOL 软件将其可视化。

2. 结果

2.1 玫瑰花活性成分以及潜在靶点

在 TCMSP 数据库共筛选出 10 种活性成分 (见表 1)。再根据 Lipinski’s 五原则筛选出 3 种活性成分: quercetin、Demethoxycapillarisin 和 vitamin-e。利用 Swiss Target Prediction 数据库预测玫瑰花的作用靶点，共得到玫瑰花潜在靶点有 215 个。

表 1 玫瑰花有效活性成分信息表

活性成分	OB (%)	DL
rugosin D_qt	57.29	0.67
Demethoxycapillarisin	52.33	0.25
cyanin	47.42	0.76
rubixanthin	47.26	0.53
quercetin	46.43	0.28
beta-carotene	37.18	0.58
beta-sitosterol	36.91	0.75
sitosterol	36.91	0.75
LYC	32.57	0.51
vitamin-e	32.29	0.7

2.2 获取抗癌靶点

通过 GeneCards 数据库对乳腺癌靶点进行筛选，共获得 2000 个潜在靶点。然后利用 Venny 2.1.0 将乳腺癌疾病基因与玫瑰花靶点基因作 Venny 图 (见图 1)，获得玫瑰花对乳

腺癌作用的潜在作用靶点一共有 105 个。

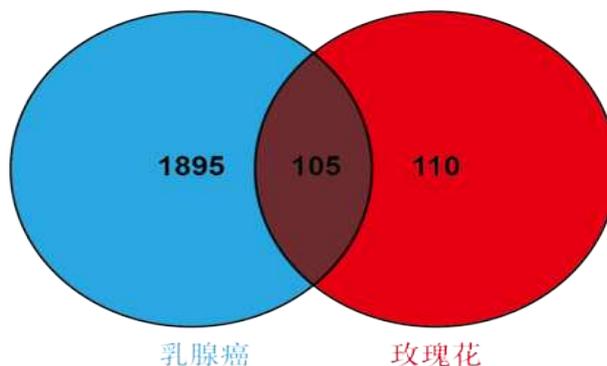


图 1 乳腺癌 – 玫瑰花靶点 Venny 图

2.3 构建药物 – 成分 – 疾病靶点网络图

利用 Cytoscape 3.9.1 制作出“药物 – 成分 – 靶基因”网络图，其中 quercetin、Demethoxycapillarisin 和 vitamin-e 这 3 种活性成分可能是玫瑰花治疗乳腺癌的核心成分 (见图 2)。

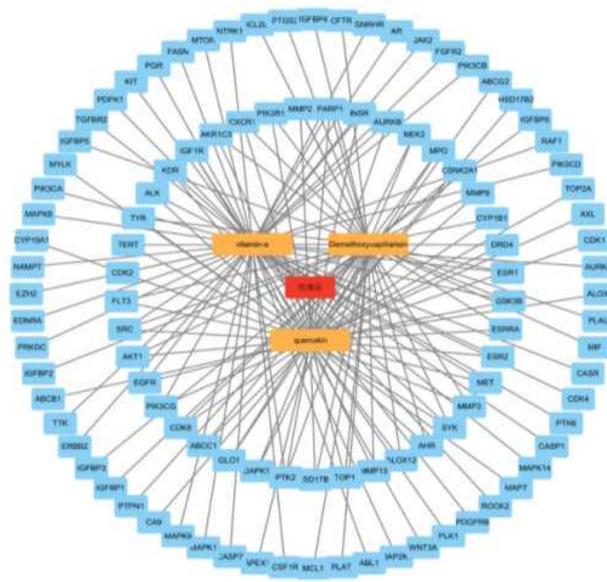


图 2 “药物 – 成分 – 靶点”网络

2.4 PPI 网络的构建

将 105 个交集靶点导入 String 数据库，构建 PPI 网络图，其中基因和相互作用分别作为网络节点和边，去除网络中处于孤立状态的蛋白质。然后下载 String 的 CSV 文件，将其导入 Cytoscape 3.9.1 以获得 PPI 网络图 (见图 3)。筛选出排名前 20 的核心靶点 (见图 4)。

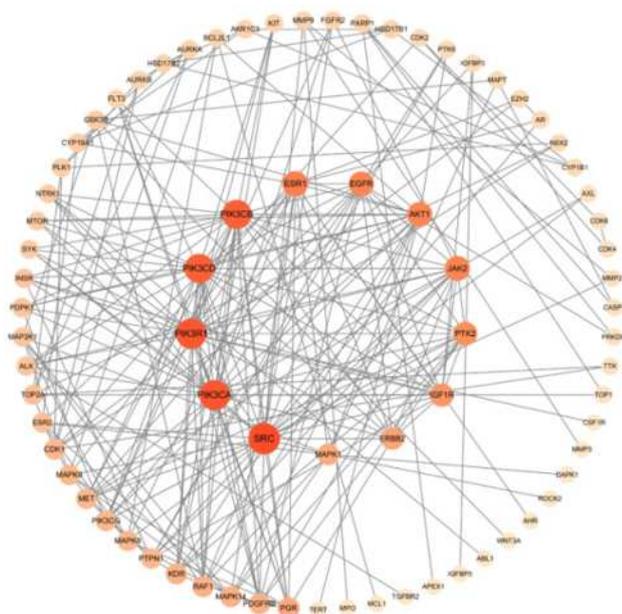


图3 玫瑰花蛋白互作网络图

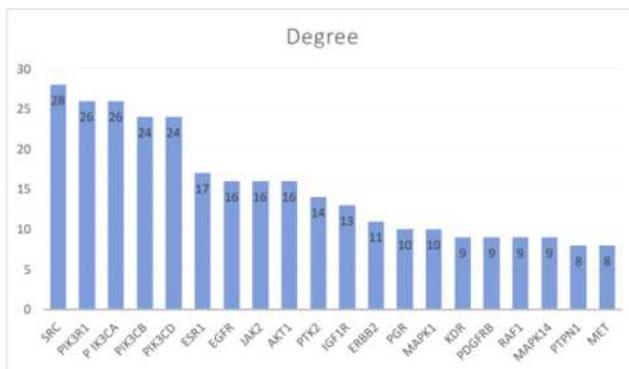


图4 PPI网络的核心靶点

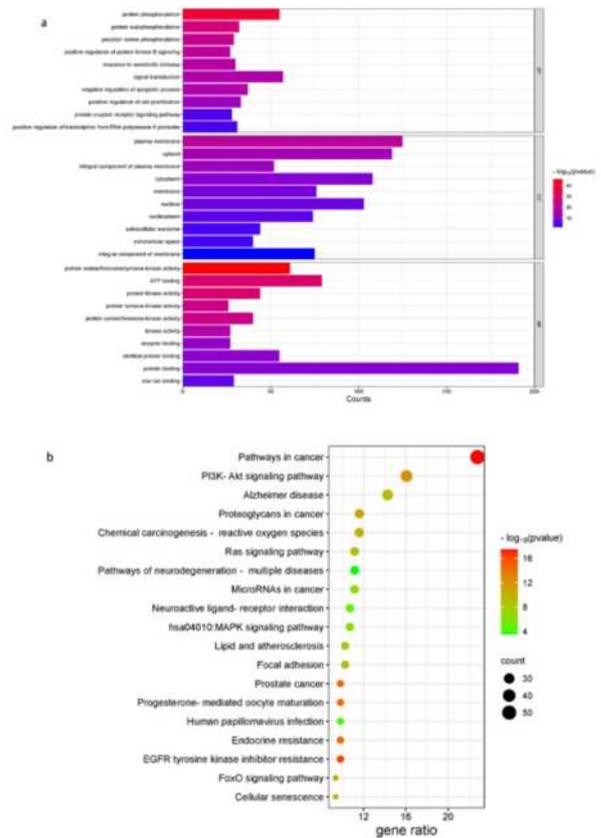
2.5 对靶基因进行 GO 功能和 KEGG 通路富集分析

将交集靶点导入 DAVID 数据平台进行 GO 和 KEGG 通路分析, 得到的数据以“FDR、P-value 均 <0.05”为前提进行筛选, 导入微生信生物平台, 得到共同靶点的 GO 功能富集和 KEGG 通路富集 (见图 5)。通过 KEGG 通路富集分析, 确定了玫瑰花在治疗乳腺癌中起作用的主要通路, 包括 FoxO、AMPK 及 PI3K-Akt 信号通路、人乳头瘤病毒感染与孕酮介导的卵母细胞成熟。

2.6 四个靶点在乳腺癌与正常组织的表达情况与生存预后分析

基于 quercetin、Demethoxycapillarisin、vitamin-e 治疗乳腺癌的核心靶点分析结果, 利用 GEPIA2 在线网站, 筛选出 4 个度值较高的靶点 PIK3R1、ESR1、EGFR 和 ERBB2, 相对于正常乳腺组织, PIK3R1 和 EGFR 在 BC 中都表现出低

水平, 而 ESR1 和 RBB2 在 BRCA 中表现高水平, 均差异显著 (图 6)。

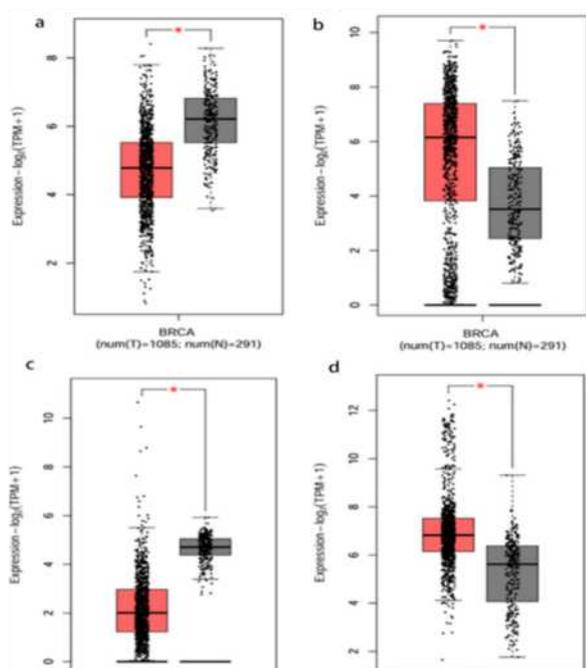


(a) GO 功能富集; (b) KEGG 通路富集

图5 交集靶点的 GO 功能富集条形图与交集靶点的 KEGG 通路富集气泡图

为了更深入地探讨 PIK3R1、ESR1、EGFR 和 ERBB2 这四个靶点与乳腺癌之间的关系, 分别对这四个靶点在乳腺癌中的 OS 和 RFS 进行了分析, 结果 (如图 7 所示)。研究发现, PIK3R1 和 ESR1 的高表达与乳腺癌患者的 OS 和 RFS 均呈明显正相关。具体而言, PIK3R1 高表达的乳腺癌患者显示出更好的总体生存期和无病生存期 (OS HR=0.84, 95% CI=0.69-1.01, p=0.061; RFS HR=0.7, 95% CI=0.63-0.78, p=6.1e-12), (见图 7, a-b)。同样, ESR1 高表达的患者也表现出更好的生存预后 (OS HR=0.73, 95% CI=0.6-0.88, p=0.00084; RFS HR=0.71, 95% CI=0.64-0.78, p=1.3e-11), (见图 7, c-d)。

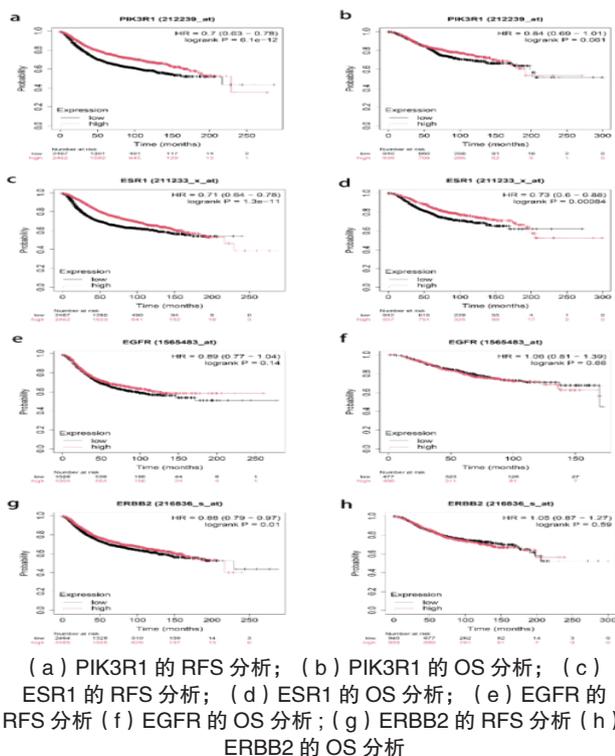
有趣的是, 当评估 EGFR 和 ERBB2 靶点的表达与乳腺癌患者生存的关系并不大, 高表达 EGFR 和 ERBB2 的患者与正常无差。



(a) PIK3R1 的表达情况; (b) ESR1 的表达情况; (c) EGFR 的表达情况; (d) ERBB2 的表达情况

图 6 靶点在 BRCA 肿瘤组织和正常组织中的表达情况

通过上述分析可以推测,在正常组织中,PIK3R1 和 ESR1 基因的高表达与乳腺癌患者的良好预后相关,其高表达可能抑制肿瘤进展,可能成为治疗乳腺癌的一个重要因素之一。

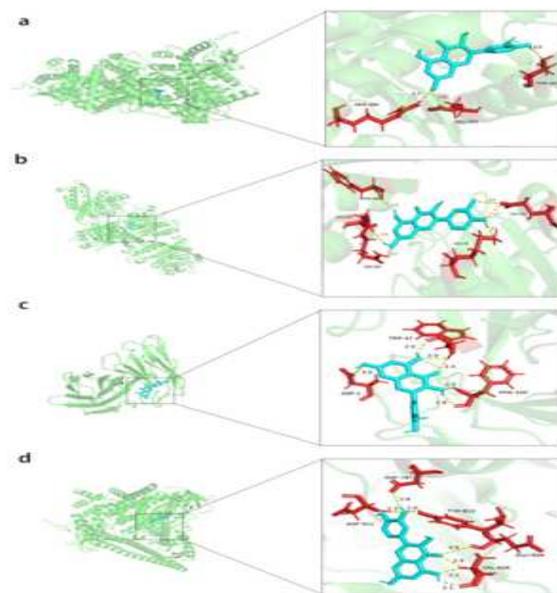


(a) PIK3R1 的 RFS 分析; (b) PIK3R1 的 OS 分析; (c) ESR1 的 RFS 分析; (d) ESR1 的 OS 分析; (e) EGFR 的 RFS 分析 (f) EGFR 的 OS 分析; (g) ERBB2 的 RFS 分析 (h) ERBB2 的 OS 分析

图 7 四个核心靶点在乳腺癌中的生存分析

2.7 分子对接结果

通过与潜在靶标的分子对接,研究 quercetin 与蛋白质的相互作用。结果显示,quercetin 分别与 EGFR、ERBB2、ESR1 和 PIK3R1 蛋白结合能力分别为 -8.3 kcal/mol、 -7.9 kcal/mol、 -8.1 kcal/mol、 -8.8 kcal/mol,表明其具有较强的结合作用。通过 Pymol 软件对 quercetin 与蛋白质形成的复合物进行可视化后,得到了结合模式,可以清晰地观察到 quercetin 与蛋白质口袋之间的氨基酸残基相互作用。例如,EGFR 蛋白相互作用的氨基酸残基包括 THR-347、GLU-353、ARG-394 (见图 8a),ERBB2 蛋白相互作用的氨基酸残基包括 GLU-112、LYS-175、GLU-225、GLN-226、PHE-268 (见图 8b),ESR1 蛋白相互作用的氨基酸残基包括 ASP-1、TRP-47、PHE-103 (见图 8c),PIK3R1 蛋白相互作用的氨基酸残基包括 ASP-787、TYR-813、GLU-826、VAL-828、ASP-911 (见图 8d)。quercetin 能够与蛋白质口袋的氨基酸形成强氢键和疏水相互作用,例如,与 ASP 和 GLU 氨基酸形成氢键,且氢键距离远小于传统氢键的 3.5\AA 距离,这种强结合将对小分子锚定在蛋白质口袋中起着重要作用。综上所述,这些相互作用有助于增强 quercetin 在 EGFR、ERBB2、ESR1 和 PIK3R1 蛋白的结合口袋内的稳定性,并形成稳定的复合物。因此,这种化合物可能具有针对 EGFR、ERBB2、ESR1 和 PIK3R1 的潜在活性。



(a) quercetin 与 EGFR 分子对接; (b) quercetin 与 ERBB2 分子对接; (c) quercetin 与 ESR1 分子对接; (d) quercetin 与 PIK3R1 分子对接

图 8 quercetin 与 EGFR、ERBB2、ESR1、PIK3R1 的分子对接模式

3. 讨论

网络药理学是一门综合性学科, 基于系统生物学、生物信息学和传统药理学, 通过多源数据库 (TCMSP、Pubchem 及 Swiss Target Prediction 等) 驱动靶点发现, 其生物信息整合机制对于识别小分子药物的靶点至关重要。通过利用网络药理学中的数据库, 能够分析所研究中药和小分子物质的 3D 立体结构和作用靶点, 进而找到肿瘤细胞的受体靶点。将小分子作用靶点和肿瘤相关细胞的受体靶点进行交集分析, 并构建蛋白质相互作用网络, 筛选出核心靶标。通过实验验证这些核心靶标对肿瘤的作用, 从而深入了解药物的作用机制。

本研究旨在探讨网络药理学和分子对接在玫瑰花治疗乳腺癌机制中的应用。通过网络药理学, 确定了玫瑰花核心主要成分为 quercetin、Demethoxycapillarisin、vitamin-e。而关于 quercetin 在乳腺癌的研究鲜有报道, 其对细胞机制亦缺乏。因此, 计划从 quercetin 这一活性成分入手, 研究其与乳腺癌的关系, 探索 quercetin 靶向治疗乳腺癌的关键分子, 并揭示 quercetin 抑制肿瘤恶性进展的分子机制, 这对于改善患者的生存预后具有重要的临床意义。

KEGG 分析发现, 玫瑰花干预乳腺癌的关键分子机制涉及多个信号传导途径。FoxO 在胰岛素 / 胰岛素样生长因子通路下游扮演着关键的调节角色, 与细胞功能中的葡萄糖、甘油三酯和胆固醇稳态等相关^[9]。AMPK 主要调控蛋白质代谢、脂质代谢、糖类代谢以及自噬和线粒体稳态等代谢过程, 而脂质代谢紊乱可能是乳腺癌发生和发展的原因之一。激活 AMPK 可以改善脂质代谢紊乱, 通过降低脂肪生成相关基因的表达来预防和减轻各种引起的乳腺癌^[10,11]。研究表明, 雌激素可以通过膜上的雌激素相关受体激活 MAPK/ERK 信号通路, 从而促进肿瘤细胞的生长和增殖^[12]。这一途径在乳腺癌中可能尤为关键。

通过与潜在靶标的分子对接, 研究了活性化合物 quercetin 与蛋白质的相互作用。结果显示, quercetin 分别与 EGFR、ERBB2、ESR1 和 PIK3R1 蛋白具有较强的结合作用。通过 Pymol 软件对 quercetin 与蛋白进行分子可视化, 获得其三维结合构象。结构分析显示, 该黄酮类化合物通过氢键和疏水作用与蛋白活性口袋内的关键氨基酸残基形成稳定互作。quercetin 能够与蛋白质口袋的氨基酸形成强氢键和疏水相互作用。综上所述, 这些相互作用有助于增强 quercetin

在 EGFR、ERBB2、ESR1 和 PIK3R1 蛋白的结合口袋内的稳定性, 并形成稳定的复合物。因此, 这种化合物可能具有针对 EGFR、ERBB2、ESR1 和 PIK3R1 的潜在活性。

综上所述, 深入探索网络药理学与分子对接技术, 并结合众多研究, 揭示玫瑰花治疗乳腺癌的潜在机制。quercetin 的效能体现在其对多种信号通路的调控, 这些通路直接参与了乳腺癌的代谢过程。其独特的作用机制在于, quercetin 不仅能加速癌细胞的增殖与凋亡的平衡, 还能刺激与炎症相关的因子活性, 从而对乳腺癌产生显著的抑制作用。进一步通过 GO 和 KEGG 分析, 乳腺癌与 FoxO、AMPK 和 PI3K/Akt 信号通路之间存在联系。这一发现为玫瑰花在治疗乳腺癌方面的机制提供了新见解, 也为未来的临床和实验研究奠定了基础。研究不仅揭示了玫瑰花中 quercetin 的潜在作用机制, 更为玫瑰花在治疗乳腺癌领域的进一步应用提供了科学依据。随着研究的深入, 期待玫瑰花在癌症治疗领域展现出更为广阔的前景。

参考文献:

- [1] Arnold M, Morgan E, Rungay H, et al. Current and future burden of breast cancer: Global statistics for 2020 and 2040[J]. *The Breast*, 2022, 66: 15–23.
- [2] 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: a cancer journal for clinicians*, 2021, 71(3): 209–249.
- [3] Xu Y, Gong M, Wang Y, et al. Global trends and forecasts of breast cancer incidence and deaths[J]. *Scientific data*, 2023, 10(1): 334.
- [4] Giaquinto A N, Sung H, Miller K D, et al. Breast cancer statistics, 2022[J]. *CA: a cancer journal for clinicians*, 2022, 72(6): 524–541.
- [5] Siegel R L, Giaquinto A N, Jemal A. Cancer statistics, 2024[J]. *CA: a cancer journal for clinicians*, 2024, 74(1): 12–49.
- [6] Yang Z, Zhang Q, Yu L, et al. The signaling pathways and targets of traditional Chinese medicine and natural medicine in triple-negative breast cancer[J]. *Journal of ethnopharmacology*, 2021, 264: 113249.
- [7] 马小娜. 花红百日养佳人[J]. *中医健康养生*, 2022, 8(01): 14–17.

[8] 高嘉, 丁宣凯, 胡嘉欢, 等. 共价有机框架材料在药物递送中的应用进展 [J]. 化学试剂, 2023, 45(01): 62-74.

[9] Dong X C. FOXO transcription factors in non-alcoholic fatty liver disease[J]. Liver research, 2017, 1(3): 168-173.

[10] Göransson O, Kopietz F, Rider M H. Metabolic control by AMPK in white adipose tissue[J]. Trends in Endocrinology & Metabolism, 2023, 34(11): 704-717.

[11] Wu L, Zhang L, Li B, et al. AMP-activated protein kinase (AMPK) regulates energy metabolism through modulating thermogenesis in adipose tissue[J]. Frontiers in physiology, 2018,

9: 122.

[12] 刘学红, 王帅, 关永格, 等. 基于 MAPK/ERK 信号通路探讨加味芍药甘草汤对 AM 病灶细胞的干预机制 [J]. 时珍国医国药, 2024, 35(02): 315-319.

作者简介: 赵云焘 (1999—), 男, 汉族, 四川省成都市, 本科, 广东药科大学, 2022 级基础医学专业研究生, 研究方向: 中药提取物抗肿瘤的机制研究。

通讯作者: 吴文梅, 博士, 硕士生导师

基金项目: 广东省中医药科研项目 (编号 20251216), 大学生创新创业训练项目 (编号 S202410573014)