

# 基于数据挖掘及网络药理学探讨乳香-没药 治疗乳腺癌的作用机制

姚 慧<sup>1</sup>, 孙 涛<sup>2\*</sup>, 徐君南<sup>2</sup>

<sup>1</sup>辽宁中医药大学研究生院, 沈阳 116620; <sup>2</sup>辽宁省肿瘤医院, 沈阳 110042

**摘要:**本文基于数据挖掘、网络药理学及分子对接方法,探讨乳香-没药药对治疗乳腺癌的分子机制。采用 Apriori 关联规则算法挖掘出乳香-没药药对并建立决策树模型,使用 Cytoscape 软件构建“化合物-靶点”可视化网络及蛋白互作网络,再运用 David 数据库进行 GO 富集分析和 KEGG 通路富集分析,并对“有效成分-靶点”拓扑分析所得度值较高的有效成分进行分子对接。结果表明乳香-没药有效成分 53 个,相关靶点 329 个,拓扑分析出 81 个核心靶点信息,GO 分析共包含 248 条生物过程、60 条细胞组分、94 条分子功能、39 条 KEGG 富集结果。分子对接结果显示与关键靶点对接较好的成分有 3-甲氧基呋喃鸟苷-9-烯-8-酮、乳香酸、槲皮素。本研究初步揭示了乳香-没药通过多成分、多靶点、多通路发挥治疗乳腺癌的作用,为其药物研发提供新的方向及依据。

**关键词:**乳香-没药;乳腺癌;数据挖掘;网络药理学;分子对接;作用机制

中图分类号:R285

文献标识码:A

文章编号:1001-6880(2021)12-2107-12

DOI:10.16333/j.1001-6880.2021.12.016

## Study on the action mechanism of frankincense-myrrh in the treatment of breast cancer based on data mining and network pharmacology

YAO Hui<sup>1</sup>, SUN Tao<sup>2\*</sup>, XU Jun-nan<sup>2</sup>

<sup>1</sup>Graduate School of Liaoning University of Traditional Chinese Medicine, Shenyang 116620, China;

<sup>2</sup>Liaoning Cancer Hospital, Shenyang 110042, China

**Abstract:** Based on data mining, network pharmacology and molecular docking, the molecular mechanism of frankincense-myrrh in the treatment of breast cancer was discussed. The frankincense-myrrh pair is mined by Apriori association rule algorithm and the decision tree model is established. The “compound-target” visual network and protein interaction network were constructed by using Cytoscape software. Then, go enrichment analysis and KEGG pathway enrichment analysis were carried out by using David database, and the active components with high degree value obtained from “active component-target” topology analysis were connected. The results showed that there were 53 active components of frankincense-myrrh, 329 related targets and 81 core target information were obtained by topological analysis. Go analysis included 248 biological processes, 60 cell components, 94 molecular functions, and 39 KEGG enrichment results. Molecular docking results showed that the components with good docking with key targets were 3-methoxyfuraguanosine-9-en-8-one, mastic acid and quercetin. This study preliminarily reveals that frankincense-myrrh plays a role in the treatment of breast cancer through multi-component, multi-target and multi-channel, and provides a new direction and basis for drug research and development.

**Key words:** frankincense-myrrh; breast cancer; data mining; network pharmacology; molecular docking; mechanism of action

据统计,2020年,女性乳腺癌首次超过肺癌,成为全球最常见的癌症,约占新发癌症病例的

11.7%<sup>[1]</sup>。我国乳腺癌的发病率也呈逐年上升且年轻化态势。目前乳腺癌主要治疗手段有手术、化疗、放疗、靶向治疗等,均会产生不同程度的副反应,影响患者预后。而中医治疗以其不良反应小、有效改善患者生存质量、延长患者生存时间的优势,在乳腺癌治疗中起着举足轻重的作用。

收稿日期:2021-06-16 接受日期:2021-10-15

基金项目:CSGO 基金(Y-2019Genecast-019);辽宁省重点攻关项目(2018225076);辽宁省“兴辽计划”青年拔尖人才(XLYC1907160);吴阶平基金(320675018541)

\*通信作者 E-mail: jianong@126.com

乳香 (frankincense) 为橄榄科植物乳香树 *Boswellia carterii* Birdw. 及同属植物 *Boswellia bhaurdajiana* Birdw. 树皮渗出的树脂, 性辛、温, 味苦, 归心、肝、脾经, 具有活血止痛、消肿生肌之效。没药 (myrrh) 是橄榄科 (Burseraceae) 没药属 *Commiphora* Engl. 植物地丁树 *Commiphora myrrh* Engl. 或哈地丁树 *Commiphora molmol* Engl. 的干燥树脂<sup>[2]</sup>, 性平, 味辛、苦, 归心、肝、脾经, 散瘀止痛、消肿生肌之效佳。乳香、没药配伍应用最早出现在《证治准绳》中的“乳香止痛散”中, 临床上二者常以药对形式出现, 相须为用, 协同增效, 为活血化瘀类常用中药。

目前有乳香、没药针对关节炎、乳腺增生、癌性疼痛的网络药理学研究, 这些研究只是直观反映了化合物的靶点及通路之间的生物过程, 尚没有深层次的对乳香、没药治疗乳腺癌进行网络药理学及相关验证研究。本文通过查阅《肿瘤方剂大辞典》等中医书籍, 收集乳腺癌相关临床诊疗数据, 挖掘药物配伍规律、核心药物等关联规则, 得出中医治疗乳腺癌所使用药物中置信度最高的为“乳香-没药”药对。进而对乳香-没药药对展开网络药理学及分子对接研究, 进一步明确乳香、没药活性成分及作用靶点, 并分析其治疗乳腺癌的分子机制。

## 1 资料与方法

### 1.1 中药复方的收集

查阅《中华人民共和国药典临床用药须知·中药成方制剂卷》(中国医药科技出版社)、《肿瘤方剂大辞典》(中医古籍出版社)、《肿瘤良方大全》(安徽科学技术出版社)、《肿瘤良方大全》(山西科学技术出版社)、《肿瘤治疗名方验方》(人民卫生出版社), 筛选出主治乳腺癌的临床医案, 检出乳腺癌内服方共 351 首。

纳入标准: 所选方剂均为中医药治疗乳岩 (即乳腺癌) 的临床药方, 符合《中医内科学》<sup>[3]</sup> 乳岩诊断; 治疗处方为中药内服制剂且有明确的药物组成、剂量; 文献中认定此方剂内服之疗效确切; 治疗方法以中药为主。排除标准: 中医临床诊断非乳岩的医案; 临床症状及方药记录不全者; 以针灸、贴敷等其他治法治疗者。

运用 Apriori 关联规则算法挖掘药对所有数据采用双人原则录入 WPS Excel 表格, 并进行筛选、规范化命名、核查等, 例如将忍冬花、二花、双花统一为金银花, 元胡、元胡索、延胡统一为延胡索, 粉甘草、粉草统一为甘草等, 建立目标数据库。运用 SPSS

Modeler 软件处理目标数据库内的中药, 运用 Apriori 关联规则算法建立数据挖掘模型, 得出置信度最高的药对。

### 1.2 建立决策树模型

决策树算法是一种离散函数的逼近方法, 通过归类来解决数据特征。经数据预读, 选择乳香作为因变量, 其余频次  $\geq 12$  的 84 个中药作为自变量进入决策树模型的筛选过程, 运用 SPSS Statistics 22.0 软件中的 CHAID、CRT、QUEST 决策树方法进行决策树模型分析。

### 1.3 药对中所含化学成分的收集与筛选

通过检索中药系统药理学分析平台 (traditional Chinese medicine systems pharmacology database and analysis platform, TCMSP) (<http://lsp.nwsuaf.edu.cn/tcmsp.php>) 检索乳香、没药的化学成分, 结合口服生物利用度 (oral bioavailability, OB)  $\geq 30\%$  和药物相似性 (drug likeness, DL)  $\geq 0.18$  作为筛选条件, 时间截止至 2021 年 5 月 10 日, 筛选出活性成分较高的化合物并收集靶蛋白, 并查阅相关文献作为补充。将上面收集到的靶蛋白在数据库 UniProt (<https://www.uniprot.org/>) 中进行校正, 选择物种为人, 用靶蛋白的简写表示, 建立药物的靶点蛋白数据库。

### 1.4 乳腺癌与药对相关靶点的收集

在人类基因数据库 Genecards (<https://www.genecards.org/>) 中以“breast cancer”为关键词, 检索出乳腺癌的相关靶基因, 并与药对中活性成分靶基因进行映射对比筛选出共同靶点, 则为药对活性成分治疗乳腺癌的靶点。

### 1.5 “化合物-靶点”的网络构建及可视化分析

运用 Cytoscape 3.7.1 软件构建乳香-没药“化合物-靶点”网络并进行可视化分析。网络中各节点 (node) 分别代表活性成分和关键靶点基因; 网络中边 (edge) 用来连接活性成分与关键靶点基因; 连接到网络的节点以度值 (degree) 为单位进行表示。某节点与其他节点连接的边数越多, 则说明该节点在网络中扮演的角色越重要, 度值也越大。

### 1.6 蛋白网络互作构建及拓扑分析

利用 Cytoscape 3.7.1 软件中 Bisogenet 插件分别对乳香-没药与乳腺癌关键靶点进行蛋白网络互作分析 (protein-protein interaction, PPI), 通过 Merge 功能取二者交集, 再用 CytoNCA 插件对网络中所有点的拓扑参数度值 (degree)、接近中心性 (closeness centrality, CC) 和介度中心性 (betweenness centrality,

BC)进行分析,借助网络拓扑理念筛选网络中的关键节点,筛选标准为 degree 大于所有节点的2倍 degree 中位数,CC、BC 大于所有节点对应值的中位数。

### 1.7 GO 功能富集和 KEGG 通路分析

将关键靶点上传至 David 6.8 (database for annotation, visualization and integrated discovery) 数据库 (<https://david.ncicrf.gov/>, Version 6.8), 以人类为研究对象,进行 GO (gene ontology) 分析以了解靶点主要的作用过程,主要包括生物过程 (biological process, BP), 细胞组成 (cellular component, CC), 分子功能 (molecular function, MF) 三方面,并且对关键靶点进行 KEGG (kyoto encyclopedia of genes and genomes, <https://www.kegg.jp/>) 富集分析,并使用微生信在线绘制柱状图和气泡图。

### 1.8 分子对接验证

定义 CytoNCA 筛选后乳香-没药药对和乳腺癌

相关疾病的核心交集靶点为受体,筛选整理后的乳香-没药重要活性成分为配体。检索 RSCB PDB 数据库 (<http://www.rcsb.org/pdb/home/home.do>) 并下载核心靶点的 3D 结构,检索并下载 TCMS 数据库中乳香-没药活性成分的 2D 结构,利用 PyMOL 2.4.0 软件对乳腺癌靶蛋白进行去水、加氢等预处理。应用 AutoDock Vina 软件进行分子对接,最后运用 PyMOL、LigPlot + 软件对对接结果进行三维可视化分析。

## 2 结果

### 2.1 药物频数分析

351 首处方涉及中药 358 味,用药总频次达 4 298 次,频率越高药物种类越少,以 6:3:1 分成高频药 (频率 >40%)、中频药 (频率 10% ~ 40%)、低频药 (频率 <10%),得出高频药 1 种,中频药 29 种,低频药 328 种,高、中频药详见表 1。

表 1 治疗乳腺癌高、中频药物频次分布

Table 1 High and intermediate frequency Chinese medicine in treatment of breast cancer

中药 Chinese medicine	频次 Frequency	频率 Rate (%)	中药 Chinese medicine	频次 Frequency	频率 Rate (%)
当归 Angelicae Sinensis Radix	188	53.56	白花蛇舌草 Spreading Hedyotis Herba	56	15.95
甘草 Glycyrrhizae Radix et Rhizoma	118	33.62	青皮 Citri Reticulatae Pericarpium Viride	55	15.67
夏枯草 Prunellae Spica	100	28.49	没药 Myrrh	55	15.67
黄芪 Astragali Radix	95	27.07	金银花 Lonicerae Japonicae Flos	55	15.67
茯苓 Poria	95	27.07	赤芍 Paeoniae Radix Rubra	54	15.38
白芍 Paeoniae Radix Alba	95	27.07	牡蛎 Ostreae Concha	54	15.38
瓜蒌 Trichosanthis Fructus	84	23.93	郁金 Curcumae Radix	51	14.53
白术 Atractylodis Macrocephalae Rhizoma	84	23.93	生地黄 Rehmanniae Radix	50	14.25
山慈菇 Crematae Pseudobulbus Pleiones	80	22.79	重楼 Paridis Rhizoma	50	14.25
香附 Cyperis Rhizoma	74	21.08	川芎 Chuanxiong Rhizoma	50	14.25
陈皮 Citri Reticulatae Pericarpium	73	20.80	丹参 Salviae Miltiorrhizae Radix et Rhizoma	49	13.96
柴胡 Bupleuri Radix	71	20.23	蜂房 Vespaes Nidus	43	12.25
穿山甲 Manis Squama	70	19.94	党参 Codonopsis Radix	41	11.68
乳香 Frankincense	66	18.80	橘叶 Citrus Reticulata Blanco	40	11.40
蒲公英 Taraxaci Herba	61	17.38	天花粉 Trichosanthis Radix	37	10.54

## 2.2 运用关联规则获取药对

取 358 味中药里频数大于平均数 12 的 84 味中药,运用 SPSS Modeler 中的 Apriori 进行建模,选取支持度  $>0.05$ ,置信度  $>0.85$ ,前项  $>18$ ,进行二项关联分析(见图 1、表 2)。其中,置信度最高的且关

联性强的药对是乳香-没药,在文献中有 64 个实例,置信度为 96.364%,支持度为 15.670%。在此次分析规则中,置信度较高的药对,可作为临床上治疗乳腺癌的常用药对。

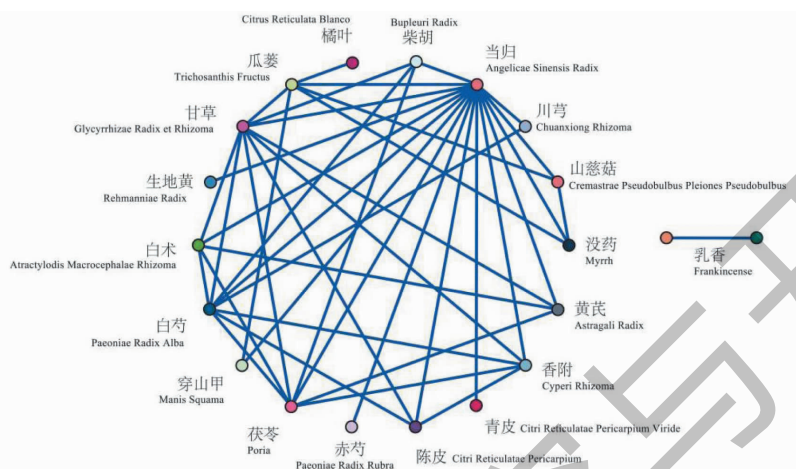


图 1 关联关系可视图

Fig. 1 Visual graph of association relationship

表 2 治疗乳腺癌处方中药物的关联规则

Table 2 Association rules for drugs in the treatment of breast cancer

关联规则 Association rule	频次 Frequency	支持度 Support (%)	置信度 Confidence level (%)	提升度 Lifting degree
乳香-没药 Frankincense-Myrrh	53	15.670	96.364	5.125
甘草-桔梗 Glycyrrhizae Radix et Rhizoma-Platycodonis Radix	30	8.832	93.548	2.783
当归-川芎 Angelicae Sinensis Radix-Chuanxiong Rhizoma	46	14.245	92.000	1.718
当归-白芥子 Angelicae Sinensis Radix-Sinapis Semen	17	5.413	89.474	1.670
没药-牛黄 Myrrh-Bovis Calculus	16	5.128	88.889	5.673
乳香-牛黄 Frankincense-Bovis Calculus	16	5.128	88.889	4.727
当归-栀子 Angelicae Sinensis Radix-Gardeniae Fructus	21	6.838	87.500	1.634
当归-白芍 Angelicae Sinensis Radix-Paeoniae Radix Alba	83	27.066	87.368	1.631
当归-桔梗 Angelicae Sinensis Radix-Platycodonis Radix	27	8.832	87.097	1.626
海藻-昆布 Sargassum-Laminariae Thallus Eckloniae Thallus	25	8.262	86.207	10.086

## 2.3 决策树模型分析

经数据预读,选择乳香作为因变量,其余频次  $\geq 12$  的 84 个中药作为自变量进入决策树模型的筛选过程。采用 SPSS Statistics 22.0 软件中的 CHAID、

CRT、QUEST 决策树方法进行识别规律的挖掘,使用 10 倍交叉验证方法进行模式验证,统一将父节点数设定为 100,子节点为 50。三种决策树模型的最佳识别中药均为没药,摘要见表 3。

表3 CHAID、CRT、QUEST 决策树模型摘要

Table 3 Summary of CHAID, CRT and Quest decision tree models

决策树模型 Decision tree mode	组成的中药 Composition of TCM	模型深度 Model depth	终结点/节点 Endpoint / node	对乳香有/无的识别路线 Identification route of frankincense	最佳识别中药 Best identification of TCM	10倍交叉验证准确率 10x cross validation accuracy
CHAID	乳香、没药、穿山甲、当归 Frankincense, Myrrh, Manis Squama, Angelicae Sinensis Radix	3	4/11	4	没药 Myrrh	95.7%
CRT	乳香、没药 Frankincense, Myrrh	1	2/5	2	没药 Myrrh	95.7%
QUEST	乳香、没药 Frankincense, Myrrh	1	2/5	2	没药 Myrrh	95.7%

## 2.4 乳香-没药活性成分的筛选

在 TCMSMP 中检索乳香-没药所有成分数据, 设置口服生物利用度  $OB \geq 30\%$ , 药物相似性  $DL \geq$

0.18, 同时筛去无对应靶点的成分, 最终筛选出乳香 8 个主要活性成分, 没药 45 个主要活性成分, 见表 4。

表4 乳香-没药的有效活性成分

Table 4 Active components of frankincense myrrh

化合物 ID Mol ID	化合物名称 Molecule name	口服生物利用度 OB	药物相似性 DL	草药 Herb
MOL001215	甘遂醇 Tirucalol	42.12	0.75	FR1
MOL001241	o-乙酰基- $\alpha$ -乳香酸 O-acetyl- $\alpha$ -boswellic acid	42.73	0.7	FR2
MOL001243	3 $\alpha$ -羟基-齐墩果-12-烯-24-酸 3 $\alpha$ -Hydroxy-olean-12-en-24-oic-acid	39.32	0.75	FR3
MOL001255	乳香酸 Boswellic acid	39.55	0.75	FR4
MOL001263	3-氧代-铁铝酸 3-Oxo-tirucallic, acid	42.86	0.81	FR5
MOL001265	乙酰- $\alpha$ -乳香酸 Acetyl- $\alpha$ -boswellic, acid	42.73	0.7	FR6
MOL001272	燃烧物 Incensole	45.59	0.22	FR7
MOL001295	扁枝杉烯 Phyllocladene	33.4	0.27	FR8
MOL001001	槲皮素-3-O- $\beta$ -D 葡萄糖醛酸苷 Quercetin-3-O- $\beta$ -D-glucuronide	30.66	0.74	MY1
MOL001002	鞣花酸 Ellagic acid	43.06	0.43	MY2
MOL001004	天竺葵色素 Pelargonidin	37.99	0.21	MY3
MOL001006	7,22E-迪恩-3 $\beta$ -醇 Poriferasta-7,22E-dien-3 $\beta$ -ol	42.98	0.76	MY4
MOL001009	Guggulsterol-VI	54.72	0.43	MY5
MOL001013	甘露醇酸 Mansumbinoic acid	48.1	0.32	MY6
MOL001019	(7S,8R,9S,10R,13S,14S,17Z)-17-亚乙基-7-羟基-10,13-二甲基-1,2,6,7,8,9,11,12,14,15-十氢环戊二烯[a]菲-3,16-二酮 (7S,8R,9S,10R,13S,14S,17Z)-17-Ethylidene-7-hydroxy-10,13-dimethyl-1,2,6,7,8,9,11,12,14,15-decahydrocyclopenta[a]phenanthrene-3,16-dione	35.75	0.48	MY7
MOL001021	7 $\beta$ ,15 $\beta$ -二羟基孕烷-4-烯-3,16-二酮 7 $\beta$ ,15 $\beta$ -Dihydroxypregn-4-ene-3,16-dione	43.11	0.51	MY8
MOL001022	11 $\alpha$ -羟基-4,17(20)-反式-二烯-3,16-二酮 11 $\alpha$ -Hydroxypregna-4,17(20)-trans-diene-3,16-dione	36.62	0.47	MY9
MOL001026	没药醇 C Myrrhanol C	39.96	0.58	MY10
MOL001027	没药酮 A Myrrhanone A	40.25	0.63	MY11
MOL001028	(8R)-3-氧代-8-羟基-聚足-13E,17E,21-三烯 (8R)-3-Oxo-8-hydroxy-polypoda-13E,17E,21-triene	44.83	0.59	MY12
MOL001029	没药酮 B Myrrhanones B	34.39	0.67	MY13
MOL001031	淫羊藿醇 Epimansumbinol	61.81	0.4	MY14

续表 4(Continued Tab. 4)

化合物 ID Mol ID	化合物名称 Molecule name	口服生物 利用度 OB	药物 相似性 DL	草药 Herb
MOL001033	Diayangambin	63.84	0.81	MY15
MOL001040	(2 <i>R</i> )-5,7-二羟基-2-(4-羟基苯基)苯并二氢吡喃-4-酮 (2 <i>R</i> )-5,7-Dihydroxy-2-(4-hydroxyphenyl) chroman-4-one	42.36	0.21	MY16
MOL001045	(13 <i>E</i> ,17 <i>E</i> ,21 <i>E</i> )-8-羟基多羟基-13,17,21-三烯-3-酮 (13 <i>E</i> ,17 <i>E</i> ,21 <i>E</i> )-8-Hydroxypolypodo-13,17,21-trien-3-one	44.34	0.58	MY17
MOL001046	(13 <i>E</i> ,17 <i>E</i> ,21 <i>E</i> )-聚-13,17,21-三烯-3,18-二醇 (13 <i>E</i> ,17 <i>E</i> ,21 <i>E</i> )-Polypodo-13,17,21-triene-3,18-diol	39.96	0.58	MY18
MOL001049	16-过氧曼苏宾-13(17)-烯-3β-ol 16-Hydroperoxymansumbin-13(17)-en-3β-ol	41.05	0.49	MY19
MOL001052	曼苏宾-13(17)-烯-3,16-二酮 Mansumbin-13(17)-en-3,16-dione	41.78	0.45	MY20
MOL001061	(16 <i>S</i> ,20 <i>R</i> )-二羟基达玛-24-烯-3-酮 (16 <i>S</i> ,20 <i>R</i> )-Dihydroxydammar-24-en-3-one	37.34	0.78	MY21
MOL001062	15α-羟基曼森宾酮 15α-Hydroxymansumbinone	37.51	0.44	MY22
MOL001063	28-乙酰氧基-15α-羟基曼森宾酮 28-Acetoxy-15α-hydroxymansumbinone	41.85	0.67	MY23
MOL001069	3β-乙酰氧基-16β,20( <i>R</i> )-二羟基达玛-24-烯 3β-Acetoxy-16β,20( <i>R</i> )-dihydroxydammar-24-ene	38.72	0.81	MY24
MOL001088	1α-Acetoxy-9,19-cyclolanost-24-en-3β-ol [(3 <i>R</i> ,5 <i>R</i> ,8 <i>R</i> ,9 <i>R</i> ,10 <i>R</i> ,13 <i>R</i> ,14 <i>R</i> ,17 <i>S</i> )-17-[(2 <i>S</i> ,5 <i>S</i> )-5-(2-羟基丙基-2-基)-2-甲基氧代-2-基]-4,4,8,10,14-五甲基-2,3,5,6,7,9,11,12,13,15,16,17-十二氢-1 <i>H</i> -环戊[ <i>a</i> ]菲-3-基]乙酸酯	44.4	0.78	MY25
MOL001092	[(3 <i>R</i> ,5 <i>R</i> ,8 <i>R</i> ,9 <i>R</i> ,10 <i>R</i> ,13 <i>R</i> ,14 <i>R</i> ,17 <i>S</i> )-17-[(2 <i>S</i> ,5 <i>S</i> )-5-(2-Hydroxypropan-2-yl)-2-methylloxolan-2-yl]-4,4,8,10,14-pentamethyl-2,3,5,6,7,9,11,12,13,15,16,17-dodecahydro-1 <i>H</i> -cyclopenta[ <i>a</i> ]phenanthren-3-yl] acetate 卡巴拉酮 Cabraleone	33.07	0.8	MY26
MOL001093	卡巴拉酮 Cabraleone	36.21	0.82	MY27
MOL001095	异芳醚酮 Isofouquierone [(5 <i>aS</i> ,8 <i>aR</i> ,9 <i>R</i> )-8-氧代-9-(3,4,5-三甲氧基苯基)-5,5 <i>a</i> ,6,9-四氢异苯并呋喃并[6,5- <i>f</i> ][1,3]苯并二氧杂环戊烯-8 <i>a</i> -基]乙酸酯	40.95	0.78	MY28
MOL001126	[(5 <i>aS</i> ,8 <i>aR</i> ,9 <i>R</i> )-8-Oxo-9-(3,4,5-trimethoxyphenyl)-5,5 <i>a</i> ,6,9-tetrahydroisobenzofurano[6,5- <i>f</i> ][1,3]benzodioxol-8 <i>a</i> -yl] acetate Phellamurin_qt	44.08	0.9	MY29
MOL001131	Phellamurin_qt	56.6	0.39	MY30
MOL001138	(3 <i>R</i> ,20 <i>S</i> )-3,20-二羟基达玛-24-烯 (3 <i>R</i> ,20 <i>S</i> )-3,20-Dihydroxydammar-24-ene	37.49	0.75	MY31
MOL001145	(20 <i>S</i> )-3β-乙酰氧基-12β,16β,25-四羟基达玛-23-烯 (20 <i>S</i> )-3β-Acetoxy-12β,16β,25-tetrahydroxydammar-23-ene	34.89	0.82	MY32
MOL001146	(20 <i>S</i> )-3β,12β,16β,25-五羟基达玛-23-烯 (20 <i>S</i> )-3β,12β,16β,25-Pentahydroxydammar-23-ene	37.94	0.75	MY33
MOL001147	(20 <i>R</i> )-3β-乙酰氧基-16β-二羟基达玛-24-烯 (20 <i>R</i> )-3β-Acetoxy-16β-dihydroxydammar-24-ene	40.36	0.82	MY34
MOL001148	3β-羟基达玛-24-烯 3β-Hydroxydammar-24-ene	40.27	0.82	MY35
MOL001156	3-甲氧基呋喃酮-9-烯-8-酮 3-Methoxyfuranoguaia-9-en-8-one [(5 <i>S</i> ,6 <i>R</i> ,8 <i>R</i> ,9 <i>Z</i> )-8-甲氧基-3,6,10-三甲基-4-氧代-6,7,8,11-四氢-5 <i>H</i> -环癸基[ <i>b</i> ]呋喃-5-基]乙酸酯	35.15	0.18	MY36
MOL001164	[(5 <i>S</i> ,6 <i>R</i> ,8 <i>R</i> ,9 <i>Z</i> )-8-Methoxy-3,6,10-trimethyl-4-oxo-6,7,8,11-tetrahydro-5 <i>H</i> -cyclodeca[ <i>b</i> ]furan-5-yl] acetate 没药甾酮 Guggulsterone	34.76	0.25	MY37
MOL001175	没药甾酮 Guggulsterone	42.45	0.44	MY38
MOL000358	β-谷甾醇 β-Sitosterol	36.91	0.75	MY39
MOL000449	豆甾醇 Stigmasterol	43.83	0.76	MY40
MOL000490	牵牛花色素 Petunidin	30.05	0.31	MY41
MOL000979	2-甲氧基呋喃酮-9-烯-8-酮 2-Methoxyfuranoguaia-9-ene-8-one	66.18	0.18	MY42
MOL000098	槲皮素 Quercetin	46.43	0.28	MY43
MOL000988	4,17(20)-(顺式)-孕二烯-3,16-二酮 4,17(20)-(cis)-Pregnadiene-3,16-dione	51.42	0.48	MY44
MOL000996	谷丙转氨酶 IV Guggulsterol IV	33.59	0.74	MY45

### 2.5 “化合物-靶点”的网络构建及可视化分析

在 TCMSP 中以  $OB \geq 30\%$  和  $DL \geq 0.18$  作为筛选条件,收集到乳香有效成分 8 个,没药有效成分 45 个,删除重复并去除假阳性,利用 UniProt (ht-

tp://www. uniprot. org) 将蛋白质名称转换成 Gene symbol,通过 Cytoscape 3. 7. 1 软件构建乳香-没药“化合物-靶点”的可视化网络图(见图 2)。

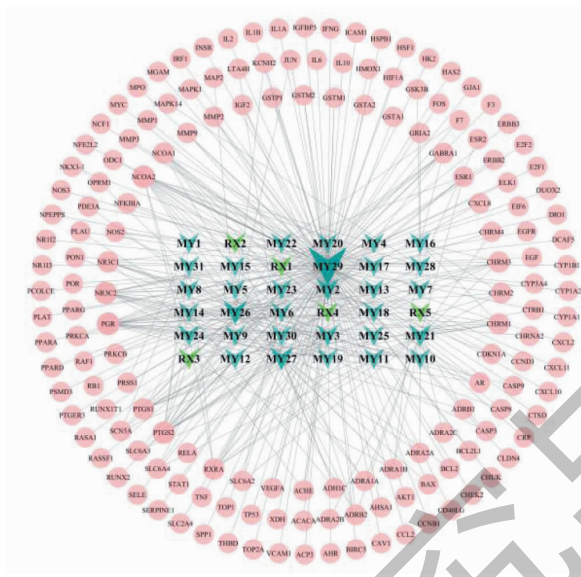


图 2 乳香-没药“化合物-靶点”的可视化网络图

Fig. 2 Visualization network diagram of “compound-target” of frankincense-myrrh

注:浅绿色为乳香;墨绿色为没药;粉色为基因靶点。Note:Light green is frankincense;Dark green is myrrh;Pink is the gene target.

### 2.6 蛋白网络互动构建及拓扑分析

利用 GeneCards (<http://www. genecards. org>) 收集乳腺癌靶点,通过多次截取参数 Relevance score 中位数法最终选取 1 123 个乳腺癌靶点。利用 Cytoscape 3. 7. 1 软件中 Bisogenet 插件分别构建药物和疾病靶基因 PPI 网络,并取交集进行蛋白网络互做分析,得到由 5 633 个节点和 141 839 个边组成的网络(见图 3),再用 CytoNCA 插件对网络进行拓扑分析,选取 degree 大于所有节点的 2 倍 degree 中位数,

CC、BC 大于所有节点对应值的中位数,提取 Hithubs 网络,得到 1 459 个节点和 63 918 个边组成的网络(见图 3),再次使用截取中位数法筛选,最终构建了 1 个具有 81 个节点和 1 524 条边的靶点网络(见图 3)。这 81 个核心基因为药物与疾病中共同基因的核心部分,其中 degree 值评分较高的有酪氨酸 3-单加氧酶/色氨酸 5-单加氧酶活化蛋白(YWHAE)、cAMP 反应元件结合蛋白(CREBBP)、生长因子受体结合蛋白 2(GRB2)、酪氨酸 3-单氧合

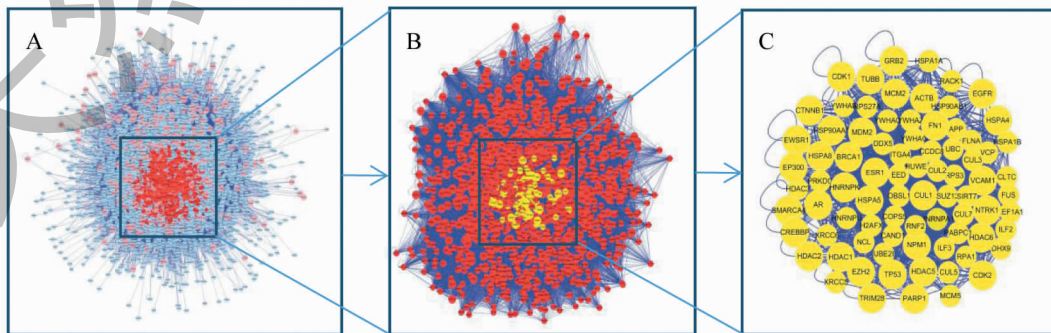


图 3 蛋白网络互动构建及拓扑分析

Fig. 3 Protein network interaction construction and topology analysis

注:A 有 5 633 个节点、141 839 条边;B 有 1 459 个节点、63 918 条边;C 有 81 个节点、1 524 条边。Note:A contains 5 633 nodes and 141 839 edges;B contains 1 459 nodes and 63 918 edges;C contains 81 nodes and 1 524 edges.



酶/色氨酸 5-单氧合酶激活蛋白  $\gamma$  肽(YWHAG)、核内不均一性核糖核蛋白 K(HNRNPK)等,为后续进行 GO 功能富集、KEGG 通路分析以及分子对接奠定基础。

## 2.7 GO 功能富集分析

将“2.6”中 81 个核心基因利用 DAVID 平台进行 GO 功能富集分析,包括 3 个分支:BP、CC 和 MF。设定阈值  $P < 0.05$  (见图 4)。BP 分析可以看出这

些靶点主要涉及 positive regulation of transcription from RNA Pol-II、negative regulation of transcription from RNA Pol-II、transcription/DNA-template 等 248 条生物过程。CC 分析可以看出靶点主要涉及 nucleus、nucleoplasm、cytoplasm 等 60 条细胞组分。MF 分析中可以看出,靶点主要涉及 protein binding、poly(A) RNA binding、enzyme binding 等 94 条分子功能。

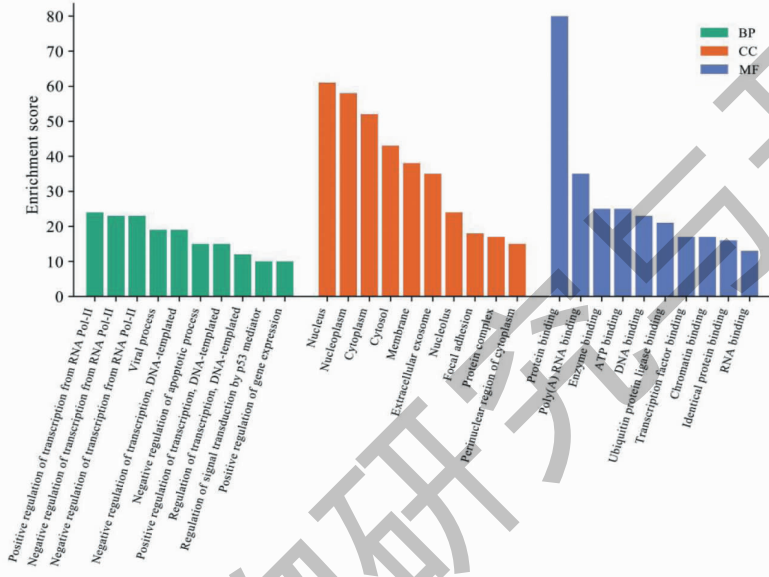


图 4 乳香-没药活性成分治疗乳腺癌靶点的 GO 富集分析

Fig. 4 GO enrichment analysis of frankincense-myrrh active ingredients for breast cancer target

## 2.8 KEGG 通路富集分析

利用 DAVID 平台对 81 个核心基因进行 KEGG 通路分析,设定阈值为  $P < 0.05$ ,得到 39 条 KEGG 通路富集,并绘制气泡图,结合表 5 和图 5,可

知基因主要集中在病毒致癌作用(viral carcinogenesis)、癌症的途径(pathways in cancer)、PI3K-Akt 信号通路(PI3K-Akt signaling pathway)、癌症中的微小核糖核酸(microRNAs in cancer)等通路上。

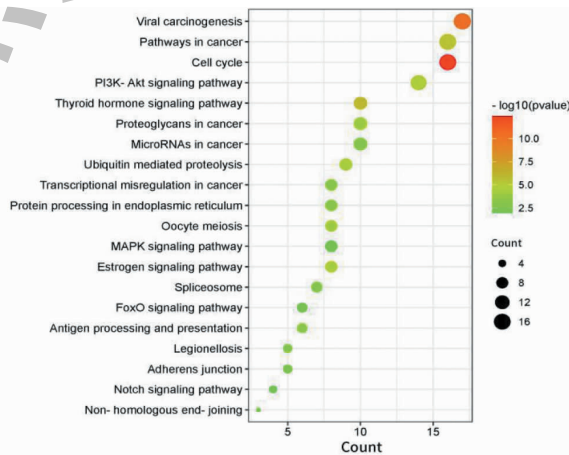


图 5 KEGG 通路富集分析气泡图

Fig. 5 Bubble diagram of KEGG pathway enrichment analysis



表 5 KEGG 通路富集分析(前 20 条)

Table 5 KEGG pathway enrichment analysis (top 20)

基因 ID Gene ID	通路 Pathway	基因 Gene
hsa05203	Viral carcinogenesis	YWHAE、HDAC5、CREBBP、HDAC2、HDAC3、HDAC1、YWHAZ、HDAC6、HNRNPK、YWHAQ、CDK2、MDM2、CDK1、EP300、GRB2、TP53、YWHAG
hsa05200	Pathways in cancer	NTRK1、CREBBP、HSP90AA1、HDAC2、HSP90AB1、HDAC1、CUL2、FN1、EGFR、AR、CDK2、MDM2、EP300、CTNNB1、GRB2、TP53
hsa04110	Cell cycle	YWHAE、CREBBP、HDAC2、PRKDC、HDAC1、CUL1、YWHAZ、YWHAQ、CDK2、MDM2、CDK1、EP300、MCM5、TP53、YWHAG、MCM2
hsa04151	PI3K-Akt signaling pathway	YWHAE、HSP90AA1、HSP90AB1、ITGA4、FN1、BRCA1、YWHAZ、EGFR、YWHAQ、CDK2、MDM2、GRB2、TP53、YWHAG
hsa04919	Thyroid hormone signaling pathway	HDAC2、CREBBP、HDAC3、HDAC1、MDM2、EP300、CTNNB1、ESR1、TP53、ACTB
hsa05205	Proteoglycans in cancer	DDX5、MDM2、FN1、CTNNB1、FLNA、GRB2、ESR1、TP53、EGFR、ACTB
hsa05206	MicroRNAs in cancer	CREBBP、UBE2I、HNRNPK、MDM2、EP300、GRB2、BRCA1、TP53、EGFR、EZH2
hsa04120	Ubiquitin mediated proteolysis	CUL7、UBE2I、CUL5、CUL3、CUL2、MDM2、CUL1、HUWE1、BRCA1
hsa05202	Transcriptional misregulation in cancer	NTRK1、DDX5、HDAC2、EWSR1、FUS、HDAC1、MDM2、TP53
hsa04141	Protein processing in endoplasmic reticulum	HSPA8、VCP、HSP90AA1、HSP90AB1、HSPA5、CUL1、HSPA1B、HSPA1A
hsa04114	Oocyte meiosis	YWHAE、AR、YWHAQ、CDK2、CUL1、CDK1、YWHAZ、YWHAG
hsa04010	MAPK signaling pathway	NTRK1、HSPA8、FLNA、GRB2、TP53、EGFR、HSPA1B、HSPA1A
hsa04915	Estrogen signaling pathway	HSPA8、HSP90AA1、HSP90AB1、GRB2、ESR1、EGFR、HSPA1B、HSPA1A
hsa03040	Spliceosome	HSPA8、DDX5、HNRNPK、HNRNPU、HNRNPA1、HSPA1B、HSPA1A
hsa04068	FoxO signaling pathway	CREBBP、CDK2、MDM2、EP300、GRB2、EGFR
hsa04612	Antigen processing and presentation	HSPA8、HSP90AA1、HSP90AB1、HSPA4、HSPA1B、HSPA1A
hsa05134	Legionellosis	EEF1A1、HSPA8、VCP、HSPA1B、HSPA1A
hsa04520	Adherens junction	CREBBP、EP300、CTNNB1、EGFR、ACTB
hsa04330	Notch signaling pathway	HDAC2、CREBBP、HDAC1、EP300
hsa03450	Non-homologous end-joining	XRCC6、XRCC5、PRKDC

## 2.9 分子对接分析

在 Cytoscape 中使用 cytoNCA 插件对乳香、没药的有效成分进行分析,根据 degree 结果得出评分较高的活性成分有没药中的 3-甲氧基呋喃酮-9-烯-8 酮、槲皮素和乳香中的乳香酸等。运用 PyMOL、AutoDock Vina 等软件对关键成分及主要靶点进行分子对接(见图 6),同时对主要有效成分作用于

CREBBP、GRB2、HNRNPK、YWHAE、YWHAG 的结合能进行了测定,一般认为配体与受体结合的构象稳定时能量越低,发生的作用可能性越大。本研究中与 GRB2 结合能最低的化合物为 Boswellic acid (结合能为-9.1 kcal/mol),由此可见乳香-没药中的核心化学成分乳香酸与受体蛋白 GRB2 形成构象能量低,结构稳定,结合活性较高。

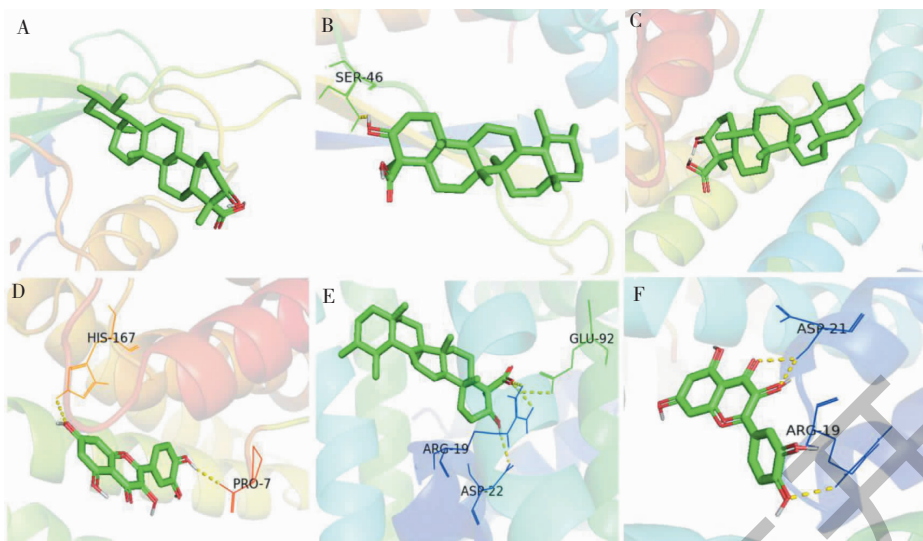


图6 乳香-没药活性成分与靶蛋白对接三维图(结合能  $\geq -9.1$  kcal/mol)

Fig. 6 Three dimensional docking diagram of the active component of frankincense-myrrh and target protein (binding energy  $\geq -9.1$  kcal/mol)

注:A~F分别为GRB2与乳香酸、HNRNPK与乳香酸、YWHAE与乳香酸、YWHAE与槲皮素、YWHAG与乳香酸、YWHAG与槲皮素的分子对接图。Note:A-F are the molecular docking diagrams of GRB2 and boswellic acid, HNRNPK and boswellic acid, YWHAE and boswellic acid, YWHAE and quercetin, YWHAG and boswellic acid, and YWHAG and quercetin, respectively.

表6 乳香-没药活性成分与靶点的结合能

Table 6 Binding energy of active components of frankincense-myrrh with target

活性成分 Active ingredient	靶点 PDB ID Target PDB ID	结合能 Binding energy (kcal/mol)
3-甲氧基呋喃酮-9-烯-8-酮 3-Methoxyfuranogaia-9-en-8-one	CREBBP-4tqn	-5.6
	GRB2-1jyq	-7.5
	HNRNPK-1zzk	-7.0
	YWHAE-2br9	-6.7
	YWHAG-4e2e	-7.2
乳香酸 Boswellic acid	CREBBP-4tqn	-7.0
	GRB2-1jyq	-9.1
	HNRNPK-1zzk	-9.0
	YWHAE-2br9	-8.0
	YWHAG-4e2e	-8.6
槲皮素 Quercetin	CREBBP-4tqn	-5.7
	GRB2-1jyq	-7.6
	HNRNPK-1zzk	-7.4
	YWHAE-2br9	-7.8
	YWHAG-4e2e	-7.7

### 3 讨论与结论

通过 Apriori 关联规则分析目标数据库里的用药规律,得到乳香与没药同时配伍治疗乳腺癌置信

度最高,置信度越高,没药与乳香配伍出现的可能性就越大。CHAID、CRT、QUEST 是三种常用决策树预测方法。CHAID 默认情况下使用多路拆分,使节点

中的样本大小变薄,从而导致树的深度减少。CRT 会执行二进制拆分(每个节点都拆分为两个子节点)。CHAID 旨在与分类/离散的目标一起使用,CRT 可以进行回归和分类。与 CART 相比,CHAID 中的拆分变量和拆分点选择的混淆程度较小。QUEST 节点可提供用于构建决策树的二元分类法,此方法的设计目的是减少大型 CRT 决策树分析所需的处理时间,同时减小分类树方法中常见的偏向类别较多预测变量的趋势。运算过程比 CRT 更简单有效,三种方法优势互补。决策树算法以乳香为因变量,CHAID、CRT、QUEST 三种决策树算法均筛选出没药与乳香的配伍可能性更大,表明乳香、没药这一活血化瘀药对参与树模型构建组方意义较大。乳香最早记载于《明医别录》,性温味淡,气香而走窜,善理气活血。没药最早记载于宋代的《开宝本草》,性平,气薄味苦,重在化瘀理血<sup>[4]</sup>,二者常相须为用,协同增效。

通过 Cytoscape 中的插件 cytoNCA 进一步挖掘乳香与没药中 53 个活性成分,根据参数 degree 排名,结合临床研究文献,发现 3-甲氧基呋喃酮-9-烯-8-酮、乳香酸、槲皮素等成分发挥着重要的作用。现代研究表明,乳香中最具特征性、研究最为深入的成分为乳香酸类(例如乳香酸)成分,属于五环三萜类化合物,对肿瘤细胞有抗增殖、分化诱导和细胞凋亡等作用,特别是其中的乙酰乳香酸毒性低,是一个有前景的抗肿瘤药物和肿瘤转移抑制剂<sup>[5]</sup>。呋喃酮类也能在一定程度上能增强肿瘤细胞增殖抑制活性<sup>[6]</sup>。槲皮素是存在于多种水果与植物中的多酚类物质,具有抗氧化、抗炎、促凋亡等作用<sup>[7]</sup>。研究发现,槲皮素通过抑制 TLR4/NF- $\kappa$ B 信号通路,间接抑制炎症因子白介素-1 $\beta$ (IL-1 $\beta$ )和肿瘤坏死因子(TNF- $\alpha$ )的表达<sup>[8]</sup>。

对关键靶点进行 GO 功能和 KEGG 通路富集分析,进一步探究乳香-没药治疗乳腺癌的作用机制,GO 包含了基因参与的生物过程、所具有的细胞组分、发挥的分子功能三方面功能信息,结果发现,乳香-没药可能通过参与 RNA 聚合酶 II 启动子转录的正/负调控、转录/DNA 模板等 248 条生物过程,细胞核、核质、细胞质等 60 条细胞组分,蛋白质结合、聚(A)RNA 结合、酶结合等 94 条分子反应,进而调控 PI3K-Akt 信号通路、Estrogen 信号通路、MicroRNAs in cancer、MAPK 信号通路、P53 信号通路等发挥抗肿瘤作用。关键基因在乳香-没药治疗乳腺癌

中具有重要意义,利用 Cytoscape 中的 cytoHubba 插件筛选出 degree 值排名前十的核心基因,根据参考相关文献和 KEGG 富集分析中的关键通路,筛选出 2 条有效作用通路 MicroRNAs in cancer 和 MAPK 信号通路,核心基因富集于这两条通路的有 CREBBP、GRB2、HNRNPK、YWHAЕ、YWHAG。MicroRNAs in cancer 通路中 MicroRNAs 的异常表达与乳腺癌血管形成、肿瘤细胞的转移、侵袭及肿瘤耐药性等过程有密切的关系<sup>[9]</sup>,miRNAs 可以用于乳腺癌的早期诊断指标<sup>[10]</sup>,也有希望成为用于乳腺癌的有效治疗药物。研究证明,没药中的没药甾酮也可能通过调节 miRNAs 的表达进而影响某些关键靶蛋白的活性从而起到抗癌作用<sup>[11]</sup>。MAPK 是细胞内广泛存在的丝/苏氨酸蛋白激酶,MAPK signaling pathway(丝裂原活化蛋白激酶信号转导通路)参与细胞增殖、凋亡、自噬过程及调控耐药相关基因和蛋白的表达<sup>[12]</sup>,干预 MAPK 信号通路,可以提高肿瘤对化疗药物的敏感性,从而逆转耐药<sup>[13]</sup>,有研究表明,没药中的  $\beta$ -榄香烯具有很好的抗肿瘤疗效,其对胶质母细胞瘤的抗增殖作用是通过激活 MAPK 信号通路实现<sup>[14]</sup>。YWHAЕ 与肿瘤的发生、发展密切相关,主要在肿瘤细胞周期、生长凋亡、扩散迁移及信号转导等途径发挥作用<sup>[15]</sup>。CREBBP 作为转录辅激活因子可参与多种细胞功能,同时有抑制肿瘤的作用,其突变能够导致卵巢癌、滤泡性淋巴瘤及乳腺癌的发病率增加<sup>[16]</sup>。GRB2 是一种在细胞中广泛表达的衔接蛋白,其异常表达或活化与肿瘤的发生息息相关,研究表明,它可以作为抗肿瘤治疗的分子靶点<sup>[17]</sup>,因此 GRB2 抑制剂也就成了抗肿瘤药物的研发热点。YWHAG 为酪氨酸 3-单氧合酶/色氨酸 5-单氧合酶激活蛋白  $\gamma$  肽(14-3-3 $\gamma$ ),参与蛋白定位及运输、细胞内信号传导、细胞周期调控和细胞凋亡<sup>[18]</sup>。HNRNPK 基因在细胞内可分布于细胞质和(或)细胞核,能与不同亚细胞结构中的多种蛋白相互作用<sup>[19]</sup>,在肿瘤中参与调控多种癌基因及抑癌基因的表达。

为了进一步验证网络药理学结果的可靠性,利用分子对接技术将核心成分和关键靶点分别进行分子对接,结果发现,核心成分 3-甲氧基呋喃酮-9-烯-8-酮、乳香酸、槲皮素与靶点 CREBBP、GRB2、HNRNPK、YWHAЕ、YWHAG 的结合能均小于-5.0 kcal/mol,表明它们有较好的结合活性。其中乳香酸与 GRB2 的结合能为-9.1 kcal/mol,表明它们具

有强烈的结合活性,提示 GRB2 是乳香-没药可能发挥治疗乳腺癌作用的重要靶点。

综上,本文采用网络药理学的方法构建“化合物-靶点-疾病”相关网络,从整体上分析乳香-没药成分、靶点、信号通路之间的相互作用关系,采用分子对接技术验证了乳香、没药抗乳腺癌作用机制,推测了乳香、没药可能通过 3-甲氧基呋喃萜-9-烯-8-酮、乳香酸、槲皮素等成分协同作用于 CREBBP、GRB2、HNRNPK、YWHAE、YWHAG 等靶点,通过 MicroRNAs in cancer 通路、MAPK 信号通路参与细胞增殖、凋亡、自噬过程及调控耐药相关基因等,侧面展示了乳香、没药通过多成分、多靶点、多通路发挥抗癌作用的机制,后续可对乳香-没药药对进行实验研究,对本文分子靶向对接虚拟筛选研究的结果进行验证,深入探讨关键作用机制,为治疗乳腺癌药物的研发提供新的方向及依据。

#### 参考文献

- Sung H, Ferlay J, Siegel RL, 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.
- Chinese Pharmacopoeia Commission. Pharmacopoeia of the People's Republic of China: Vol 1 (中华人民共和国药典:第一部)[M]. Beijing: China Medical Science Press, 2020, 186-187.
- Zhang BL, Wu MH. Internal Medicine of Traditional Chinese Medicine(中医内科学)[M]. Beijing: China Traditional Chinese Medicine Press, 2017:132-133.
- Wu JR, Jin YP, Zhang B, et al. Study on the formulation law of Chinese patent medicines for removing blood stasis based on association rules and entropy clustering[J]. Chin Med Herald(中国医药导报), 2015, 12(13):117.
- Lv ZL, Gao Y, Li J, et al. Synthesis and antitumor activity of 3-arylbenzofuranone compounds[J]. Chem J Chin Univ(高等学校化学学报), 2013, 11:2531-2539.
- Ma ZW, Wang HJ, Cheng Y, et al. Study on the mechanism of frankincense myrrh in the treatment of knee osteoarthritis based on network pharmacology[J]. J Liaoning Univ Tradit Chin Med(辽宁中医药大学学报), 2021, 23(6):72-78.
- Zhu XQ, Liu ZQ, Wang BL. Study on the mechanism of Quercetin against cervical cancer based on biomolecular network and molecular docking[J]. Nat Prod Res Dev(天然产物研究与开发), 2020, 32:23-31, 69.
- She WW, Huang QR, Li YM, et al. The role of MicroRNAs in the occurrence, development and metastasis of breast cancer[J]. J Ludong Univ: Natural Sci(鲁东大学学报:自科版), 2020, 36:338-352.
- Aggarwal V, et al. Emergence of circulating microRNAs in breast cancer as diagnostic and therapeutic efficacy biomarkers[J]. Mol Diagn Ther, 2020, 24(2):153-173.
- Wang FH, Han JC, Li DF, et al. Aqueous extract of Dendrobium huoshanense mediated NF- $\kappa$ B/p65 and p38 MAPK attenuate alcoholic liver injury in mice[J]. Nat Prod Res Dev(天然产物研究与开发), 2017, 29:569-574.
- Zhong F, Tong ZT, Fan LL, et al. Effects of myxosterone on the growth and miRNA expression profile of transplanted RBE cells of cholangiocarcinoma[J]. J Anhui Med Pharm(安徽医药), 2015, 19:640-642.
- Wang B, Elsa BK, Andrew JA, et al. ATXN1L, CIC, and ETS transcription factors modulate sensitivity to MAPK pathway inhibition[J]. Cell Rep, 2017, 18:1543-1557.
- Yao YQ, Ding X, Jia YC, et al. Anti-tumor effect of  $\beta$ -elemene in glioblastoma cells depends on p38 MAPK activation[J]. Cancer Lett, 2008, 264(1):127-134.
- Chai DY, Yuan JQ, Zhou YP. Research progress of MAPK signaling pathway on tumor multidrug resistance[J]. Chin J New Drug(中国新药杂志), 2019, 28:948-953.
- Wu XY, Li LH, Wang JY, et al. Expression and clinical significance of YWHAE in colorectal cancer[J]. Chin J Lab Diagn(中国实验诊断学), 2019, 23:886-888.
- Garcia C, Ruiz L, Sarmenteroj, et al. CREBBP/EP300 bromodomain inhibition affects the proliferation of AF-positive breast cancer cell lines[J]. Mol Cancer Res, 2019, 17:720-730.
- Zhang XY, An S, Guo XX, et al. Physiological function of growth factor receptor binding protein 2(Grb2)[J]. Chin J Cell Biol(中国细胞生物学学报), 2015, 37:1029-1035.
- Aghazadeh Y, Papadopoulos V. The role of the 14-3-3 protein family in health, disease, and drug development[J]. Drug Discov Today, 2016, 21(2):278-287.
- Chen Y, Zeng Y, Xiao Z, et al. Role of heterogeneous nuclear ribonucleoprotein K in tumor development[J]. J Cell Biochem, 2019, 120:14296-14305.