

基于系统的化合物-靶点相互作用预测模型的消栓通络方 抗脑卒中网络药理学研究

郑一夫^{1,2#}, 孔令雷^{1#}, 贾皓¹, 张宝月¹, 王喆¹, 许律捷¹, 刘艾林^{1*}, 杜冠华^{1*}

(1. 中国医学科学院、北京协和医学院药物研究所, 北京 100050; 2. 武汉大学药学院, 湖北 武汉 430072)

摘要: 消栓通络方临床上主治脑血栓引起的精神呆滞、言语迟涩等症状, 疗效显著, 但其作用机制尚不明确。本文通过搜集消栓通络方中的化学成分和治疗脑卒中相关靶点, 获得 1 251 个化学成分和 10 个脑卒中相关靶点, 采用朴素贝叶斯和递归分割等机器学习算法, 基于分子指纹和分子描述符, 结合分子对接方法, 构建了脑卒中 10 个相关靶点的 18 个化合物-靶点相互作用预测模型。应用这些模型预测了消栓通络方的活性化学成分及其作用靶点, 发现了 153 个潜在活性化学成分, 其中 22 个可以与 2 个以上脑卒中药物靶点相互作用。在此基础上, 利用网络构建专业软件, 构建了化学成分-靶点网络, 并通过 Gene-Ontology (GO) 富集分析, 确证了靶点重要的生物过程, 如凝血 (blood coagulation)、血管生成正调控 (positive regulation of angiogenesis) 和离子转运正调控 (positive regulation of ion transport) 等。在此基础上, 构建了靶点-通路网络。本研究利用机器学习、分子对接、虚拟筛选、数据挖掘及网络构建等技术方法, 探索并部分揭示了消栓通络方抗脑卒中的活性物质基础及其化学成分-靶点-通路的网络作用机制, 为消栓通络方的深入研究提供重要信息依据。

关键词: 消栓通络方; 脑卒中; 药物靶点; 机器学习; 分子对接; 虚拟筛选; 网络药理学
中图分类号: R285 文献标识码: A 文章编号: 0513-4870(2020)02-0256-09

Network pharmacology study on anti-stroke of Xiaoshuan Tongluo formula based on systematic compound-target interaction prediction models

ZHENG Yi-fu^{1,2#}, KONG Ling-lei^{1#}, JIA Hao¹, ZHANG Bao-yue¹, WANG Zhe¹,
XU Lü-jie¹, LIU Ai-liu^{1*}, DU Guan-hua^{1*}

(1. Institute of Materia Medica, Chinese Academy of Medical Sciences and Peking Union Medical College,
Beijing 100050, China; 2. College of Pharmacy, Wuhan University, Wuhan 430072, China)

Abstract: Xiaoshuan Tongluo formula is effective in treating mental retardation and speech astringency caused by cerebral thrombosis, but its mechanism remains unclear. In this investigation, by collecting the chemical constituents from Xiaoshuan Tongluo formula and the targets related to stroke, we obtained 1 251 constituents from the formula and 10 drug targets related with stroke. We established 18 prediction models of compound-target interaction for 10 stroke-related targets, using molecular docking method and machine learning methods includes Naive Bayesian and recursive partitioning based on the input of molecular fingerprints and molecular descriptors. Using

收稿日期: 2019-06-30; 修回日期: 2019-09-23.

基金项目: 国家自然科学基金资助项目 (81673480); 北京市自然科学基金项目 (7182113, 7192134); “十三五”重大新药创制专项 (2018ZX09711001-009-009, 2018ZX09711001-003-002); 协和创新工程项目 (2016-I2M-3-007).

#共同第一作者.

*通讯作者 Tel: 86-10-83150885, E-mail: liuailin@imm.ac.cn;

Tel: 86-10-63165184, E-mail: dugh@imm.ac.cn

DOI: 10.16438/j.0513-4870.2019-0521

these models, we predicted the active chemical constituents from Xiaoshuan Tongluo formula and their drug targets, 153 potential active constituents were discovered, 22 among them could interact with at least two drug targets related with stroke. On this basis, the chemical constituent-target network was constructed using network construction software, and then the important metabolic pathways of the targets were identified by using Gene-Ontology (GO) enrichment analysis, such as blood coagulation, positive regulation of angiogenesis, positive regulation of ion transport and so on. On this basis, a target-pathway network was constructed. In conclusion, using machine learning, molecular docking, virtual screening, data mining and network construction, this study explored and partially revealed the active chemical constituents and chemical constituent-target-pathway network action mechanism of Xiaoshuan Tongluo formula against stroke, which will provide important information for its further study.

Key words: Xiaoshuan Tongluo formula; stroke; drug target; machine learning; molecular docking; virtual screening; network pharmacology

脑卒中是由于脑部血管突然破裂或因血管阻塞导致血液不能流入大脑而引起脑组织损伤的一组疾病。据国家卫生部第三次全国死亡原因调查显示, 中国脑卒中年发病率超过 200 万, 每年 150 万人因脑卒中死亡, 死亡率已经是西方发达国家的 4~5 倍, 是日本的 3.5 倍^[1], 跃升为我国第一大致死原因。与此相应的是, 获准脑卒中治疗的药物寥寥无几, 目前临床上仅存在唯一单靶点作用药物组织型纤溶酶原激活剂 (tissue plasminogen activator, tPA)^[2] 治疗脑卒中。但由于其治疗窗狭窄, 对患者发病情况有严苛条件要求, 仅 3%~5% 患者可以接受治疗^[3], 且有导致脑出血的风险^[4]。因此, 有关脑卒中的治疗亟待提高。脑卒中分为缺血性脑卒中和出血性脑卒中, 其中缺血性脑卒中占 70% 左右^[5,6]。因此, 本文将针对缺血性脑卒中开展相关研究。

中药消栓通络方为临床常用的经验方, 自 1990 年起为历版中国药典所收录, 主要用于脑卒中恢复期的治疗。消栓通络方包含川芎、丹参、黄芪、泽泻、三七、槐花、桂枝、郁金、木香、冰片、山楂共 11 味中药配伍而成^[6], 具有活血化瘀, 温经通络之功效, 主治血脂增高, 脑血栓引起的精神呆滞、舌质发硬、言语迟涩、发音不清、手足发凉和活动疼痛。但此常用药方鲜有研究, 且其发挥功效的活性成分及作用机制尚未见深入的研究报道。

中药复方使用多味草药配伍而成, 作用成分复杂, 可以通过与脑卒中相关的多个靶点产生药效作用。通过生物学实验手段来确定中药中有效成分的抗脑卒中机制不仅消耗大量的人力、物力和财力, 且任务艰巨^[7]。因此, 急需发展新的、有效的方法来研究中药抗脑卒中的网络作用机制。随着中药现代化发展, 国内外建立了多个中药成分数据库, 如 TCMID、TCMSP、HIT、TCMDB@Taiwan 等^[8]。随着生命科学的发展和蛋白质结构解析技术的进步, 越来越多的蛋白结构得到解析, 在此基础上构建了许多重要的药物靶点数据库, 如 Thomson Reuters Integrity^[9]、Therapeutic Target

Database (TTD) 等^[10]。基于现有数据库和机器学习方法, 进行数据挖掘及相关生物信息学分析, 结合计算模拟方法和网络构建技术, 使中药活性成分发现和网络机制研究成为可能。数据库的快速发展推动了国内外大量对于中药复方有效成分机制的研究^[11,12], 本文采用类似的方法进行消栓通络方治疗脑卒中网络药理学机制研究。

本文通过中药成分数据库及上市药物靶点数据库, 结合文献分析, 搜集消栓通络方中的化学成分及治疗脑卒中相关靶点信息, 基于分子指纹和分子描述符, 采用朴素贝叶斯和递归分割等机器学习算法, 结合分子对接计算, 系统地构建抗脑卒中的化合物-靶点相互作用预测模型。应用预测模型, 预测消栓通络方潜在活性成分及其作用靶点。在此基础上, 利用生物网络构建软件, 构建化学成分-靶点网络和靶点-通路网络, 以揭示消栓通络方抗脑卒中的活性物质基础及其网络作用机制, 为消栓通络方的深入研究提供信息依据。

材料与方法

消栓通络方中药成分的搜集 利用中药药理学数据库 TCMSP (isp.nwu.edu.cn/tcmsp.php)、TCMID (183.129.215.33/tcmid/) 和 Batman-TCM (bionet.ncpsb.org/batman-tcm/), 分别以消栓通络方中 11 味药材名称为关键词, 检索并保存化学成分结构等信息。

脑卒中靶点及靶点预测模型的选择 在 Thomson Reuters Integrity 和 Therapeutic Target Database (TTD) 数据库中, 针对 stroke 搜索临床试验 3 期及上市药物, 进一步检索药物作用靶点, 经过整理, 结合文献分析, 确定脑卒中相关靶点。然后进入 BindingDB 数据库搜索靶点的已知活性化合物, 由于机器学习模型都设置了 9 个描述符, 当活性化合物的个数是描述符个数的 3 倍以上, 建立的模型才具有统计学意义, 因此, 以至少 50 个活性化合物作为基于活性化合物结构建立预测模型的

标准,以保证机器学习模型可靠性。若靶点的活性化合物个数大于50,则采用机器学习方法建模;若活性化合物个数不足50,则采用分子对接方法建模。

化合物-靶点相互作用机器学习模型的建立 对于活性化合物数大于50个的靶点,并以 $IC_{50} < 10\ 000\ \text{nmol}\cdot\text{L}^{-1}$ 为条件筛选活性配体。从中随机挑选25%的配体,上传到DUD-E [a database of useful (docking) decoys — enhanced, <http://dude.docking.org/>] 数据库,以产生非活性配体。

在此基础上,按照活性化合物与非活性化合物1:3的比例,测试集与训练集1:3的比例随机分配所有化合物。然后采用朴素贝叶斯^[13]和递归分隔^[14]算法,基于ECFP₆^[15]和MACCS^[16]分子指纹,以及分子描述符,在Discovery Studio 2018平台上构建模型,建模流程如图1所示,其中MACCS分子指纹利用PaDEL Descriptor软件进行计算。针对每个脑卒中相关靶点,采用不同算法和分子指纹,分别构建了4个相关模型,即每个靶点有4个预测模型NB(ECFP₆)、NB(MACCS)、RP(ECFP₆)和RP(MACCS)。经过前期分析发现,当采用朴素贝叶斯算法时,基于ECFP₆分子指纹的模型表现更好;当使用MACCS分子指纹时,递归分隔算法表现更好^[17]。因此,最后每个靶点采用NB(ECFP₆)和RP(MACCS)的组合预测模型。

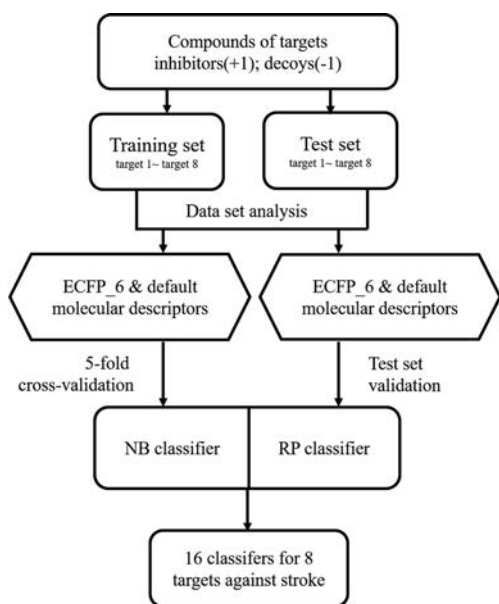


Figure 1 Workflow of the construction of machine learning models

构建的模型需经过训练集的五倍交叉验证^[18]与测试集验证,以评价预测模型预测能力。验证指标基于真阳性 (true positive, TP)、真阴性 (true negative, TN)、假阳性 (false positive, FP) 和假阴性 (false negative, FN)

的基础数据进行计算,包括敏感度 (sensitivity, SE)、精确度 (specificity, SP)、总体预测精度 (overall prediction accuracy, Q)、马修斯相关系数 (Matthews correlation coefficient, MCC) 和受试集工作特征曲线的曲线下面积 (area under the receiver operating characteristic curve, AUC)。各指标计算公式如下:

$$SE = \frac{TP}{TP + FN} \quad (1)$$

$$SP = \frac{TN}{TN + FP} \quad (2)$$

$$Q = \frac{TP + TN}{TP + TN + FP + FN} \quad (3)$$

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FN)(TP + FP)(TN + FN)(TN + FP)}} \quad (4)$$

另外, AUC为ROC曲线下与坐标轴围成的面积,最大不会超过1,数值越大表示预测效果越好。

分子对接模型的建立 对于活性化合物小于50个的靶点,在PDB (Protein Data Bank) 数据库下载蛋白晶体结构,选择分辨率 $< 2.5\ \text{\AA}$ 并具有配体的复合物,以保证对接的精度。在对接前对蛋白分子进行预处理,去除结构周围的水分子,加上氢原子,提取所有配体,进行逐一重新对接,其中对接的活性口袋由原始配体位置决定。使用Discovery Studio 2018中的CDocker分子对接方法,对蛋白与蛋白配体进行对接^[19]。对接结果通过RMSD^[18] (root-mean-square deviation) 进行验证,RMSD值小于2.5则认为分子对接结果有效。从所有配体中选择RMSD值最小的配体对应的活性口袋进行中药化学成分的对接。

中药化学成分-靶点相互作用预测 基于上述建立的所有机器学习模型和分子对接模型对消栓通路方中的化学成分进行靶点预测。对于机器学习模型,预测结果为True或False,True表示化合物与靶点可以相互作用,False则相反。对于分子对接模型,根据-CDocker energy计算预测结果,-CDocker energy可以说明配体与靶点之间结合强度,对于-CDocker energy值比原配体大的中药化学成分,认为其对接蛋白为化学成分的潜在靶点。然后,对预测结果进行整合分析,统计消栓通路方中活性化学成分与各个脑卒中相关靶点的作用关系。

作用网络分析 统计潜在活性化学成分及其作用靶点,利用网络构建软件Cytoscape^[20]构建化学成分-靶点作用网络。之后,在DAVID数据库中,对脑卒中靶点进行Gene-Ontology (GO)^[21]富集分析,选择 $P < 0.05$ 的靶点代谢通路,并构建靶点-通路网络^[22,23]。

结果

1 消栓通路方中药成分的搜集

通过中药药理学数据库TCMSP、TCMID和Batman-TCM, 分别以药材名称为关键字, 检索并保存化合物。搜集整理得到各单味药的化学成分数目如下: 冰片1个、川芎189个、丹参202个、桂枝220个、槐花27个、黄芪87个、木香106个、三七119个、山楂32个、郁金222个和泽泻46个, 总计1251个化学成分。

2 脑卒中靶点的搜集与分类

利用Thomson Reuters Integrity和TTD数据库, 并结合文献分析, 共确定10个脑卒中相关靶点, 分别是F3、F10、F11、MMP9、P2RY1、TBXA2R、VEGFA、SERPINE1、THBD和P2RY12。其中包含凝血系统相关靶点4个, 血小板相关靶点3个、保护血管内皮靶点2个及纤溶系统靶点1个, 具体名称、UniProt序列号及其功能如表1所示。其中THBD和P2RY12, 由于BindingDB数据库中缺少其配体相关信息, 采用分子对接方法建模, 其余8个靶点采用机器学习方法建模。

3 机器学习模型

从BindingDB数据库和DUD-E数据库分别得到靶点活性配体及非活性配体后, 按照活性化合物与非活性化合物1:3比例, 测试集与训练集1:3的比例随机分配所有化合物, 各个模型中化合物分配情况如表2所示。其中Tanimoto index为评价数据集中化合物空间多样性的指数, 化合物空间多样性越大, 其数值越小, 构建模型预测能力则越好。由表中数据可知, 所有靶点的数据集的Tanimoto index^[24]均小于0.150, 说明

活性化合物与非活性化合物相似度较低, 数据集中化合物空间多样性较高。

利用Discovery Studio 2018平台, 基于ECFP₆和MACCS分子指纹, 以及分子描述符, 采用朴素贝叶斯和递归分隔算法构建模型。经过训练集的五倍交叉验证及测试集验证, 评价模型的有效性和精确度, 其结果如表3~6所示, 其中马修斯相关系数MCC和受试集工作特征曲线的曲线下面积AUC作为评价模型预测能力的主要指标。由表可知, NB(ECFP₆)模型训练集的五倍交叉验证中AUC均大于0.996, MCC均大于0.977; 测试集验证中AUC均大于0.992, MCC均大于0.778。RP(MACCS)模型训练集的五倍交叉验证中AUC均大于0.976, MCC均大于0.869; 测试集验证中除F11靶点AUC为0.746外, AUC均大于0.928; 除F11靶点MCC为0.583外, MCC均大于0.710。因此, 剔除F11靶点的递归分隔模型, 所有模型均具有较好的有效性和精确度, 可以用于消栓通路方中化学成分作用靶点的预测。

4 分子对接模型

应用Discovery Studio 2018中的CDocker分子对接方法, 对蛋白与蛋白配体进行对接。对接结果通过RMSD进行验证, 两个靶点THBD和P2RY12所构建模型RMSD值分别为0.7580和0.6566, RMSD值均小于2.5, 认为分子对接模型有效。蛋白晶体结构与蛋白原始配体相互作用如图2所示。

5 靶点预测结果

基于上述建立的所有机器学习模型和分子对接模

Table 1 The name and classification of stroke related targets

Target symbol	Target name	UniProt accession	Function
Tissue factor	F3	P13726	Coagulation system
Coagulation factor X	F10	P00742	Coagulation system
Coagulation factor XI	F11	P03951	Coagulation system
Matrix metalloproteinase-9	MMP9	P14780	Vascular endothelium
P2Y purinoceptor 1	P2RY1	P47900	Platelet-related target
Thromboxane A2 receptor	TBXA2R	P21731	Platelet-related target
Vascular endothelial growth factor A	VEGFA	P15692	Vascular endothelium
Plasminogen activator inhibitor 1	SERPINE1	P05121	Fibrinolytic system
Thrombomodulin	THBD	P07204	Coagulation system
P2Y purinoceptor 12	P2RY12	Q9H244	Platelet-related target

Table 2 Diversity analysis of compounds datasets

Target	Training set				Test set			
	Inhibitor	Decoy	Total	Tanimoto index	Inhibitor	Decoy	Total	Tanimoto index
F3	65	195	260	0.098	21	63	84	0.097
F10	351	1053	1404	0.120	117	351	468	0.107
F11	58	168	226	0.124	19	57	76	0.114
P2Y1R	569	1707	2276	0.127	190	570	760	0.121
SERPINE1	235	705	940	0.121	78	234	312	0.112
TBXA2R	377	1131	1508	0.134	125	375	500	0.125
VEGFA	65	195	260	0.089	22	66	88	0.069
MMP9	1400	4200	5600	0.150	466	1398	1864	0.133

Table 3 Performance of NB (ECFP₆) model's training set by 5-fold cross-validation. TP: True positive; FN: False negative; FP: False positive; TN: True negative; SE: Sensitivity; SP: Specificity; MCC: Matthews correlation coefficient; AUC: The area under the receiver operating characteristic curve

Target	TP	FN	FP	TN	SE	SP	MCC	AUC
F3	65	0	0	195	1.000	1.000	1.000	1.000
F10	351	0	2	1 051	1.000	0.998	0.996	0.999
F11	58	0	2	166	1.000	0.988	0.977	0.996
MMP9	1 389	11	3	4 197	0.992	0.999	0.993	0.999
P2RY1	561	8	0	1 707	0.986	1.000	0.991	0.999
SERPINE1	235	0	1	704	1.000	0.999	0.997	1.000
TBXA2R	376	1	1	1 130	0.997	0.999	0.996	1.000
VEGFA	65	0	0	195	1.000	1.000	1.000	0.998

Table 4 Performance of NB (ECFP₆) model's test set validation

Target	TP	FN	FP	TN	SE	SP	MCC	AUC
F3	18	3	0	63	0.857	1.000	0.905	0.995
F10	116	1	8	343	0.991	0.977	0.950	0.999
F11	19	0	8	49	1.000	0.860	0.778	1.000
MMP9	465	1	17	1 381	0.998	0.988	0.975	1.000
P2RY1	190	0	3	567	1.000	0.995	0.990	0.999
SERPINE1	77	1	2	232	0.987	0.991	0.975	1.000
TBXA2R	119	6	0	375	0.952	1.000	0.968	0.992
VEGFA	22	0	0	66	1.000	1.000	1.000	1.000

Table 5 Performance of RP (MACCS) model's training set by 5 fold cross-validation

Target	TP	FN	FP	TN	SE	SP	MCC	AUC
F3	62	3	0	195	0.954	1.000	0.969	0.976
F10	339	12	62	991	0.966	0.941	0.869	0.988
F11	57	1	8	160	0.983	0.952	0.902	0.995
MMP9	1 379	21	65	2 635	0.985	0.976	0.954	0.996
P2RY1	563	6	61	1 646	0.989	0.964	0.926	0.995
SERPINE1	226	9	22	683	0.962	0.969	0.914	0.992
TBXA2R	374	3	30	1 101	0.992	0.973	0.944	0.996
VEGFA	65	0	9	186	1.000	0.954	0.915	0.998

Table 6 Performance of RP (MACCS) model's test set validation

Target	TP	FN	FP	TN	SE	SP	MCC	AUC
F3	18	3	0	63	0.857	1.000	0.905	0.928
F10	110	7	26	325	0.940	0.926	0.826	0.960
F11	10	9	2	55	0.526	0.965	0.583	0.746
MMP9	457	9	32	1 366	0.981	0.977	0.943	0.992
P2RY1	183	7	19	551	0.963	0.967	0.911	0.981
SERPINE1	73	5	36	198	0.936	0.846	0.710	0.961
TBXA2R	120	5	10	365	0.960	0.973	0.921	0.973
VEGFA	22	0	8	58	1.000	0.879	0.803	0.943

型, 分别对消栓通路方中的 1 251 个化学成分进行虚拟筛选, 以期预测发现与靶点相互作用的活性成分。对预测结果进行统计和整理, 共得到 153 个活性化学成分, 其中 22 个化学成分可以与 2 个以上靶点相互作用。

6 化学成分-靶点和靶点-通路作用网络

选择活性化学成分最多的一味中药丹参, 利用 Cytoscape 软件构建其化学成分-靶点作用网络, 如图 3 所示。在 DAVID 数据库中, 对脑卒中靶点进行 GO 富

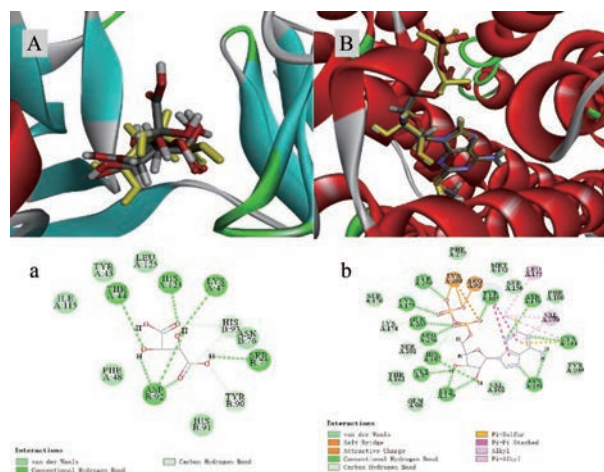


Figure 2 The crystal structure and the interaction between proteins and their original ligands. A: Interaction between THBD and its ligand in their three dimensional (3D) structures, Graph a shows their interaction in two dimensional (2D) structures; B: Interaction between P2RY12 and its ligand in their 3D structures, Graph b shows their interaction in 2D structures

集分析, 得到靶点的代谢通路, 选择 P 值小于 0.05 的通路, 在此基础上, 构建靶点-通路网络, 如图 4 所示。

讨论

缺血性脑卒中的理想治疗模式包括血管再通和神经保护, 去除脑血栓的同时, 保护神经减轻缺血再灌注造成的损害^[25], 这也意味着脑卒中的治疗需要多靶点介入。因此本研究开展网络药理学机制分析, 力图发现治疗脑卒中的多靶点有效成分组, 多个有效成分通过多靶点多通路作用于脑卒中, 以期得到更好的疗效。

本文的创新之处在于首次针对脑卒中相关药物靶点, 系统地建立了化合物-药物靶点相互作用预测模型, 并通过验证具有很好的预测能力。此外, 应用构建的化合物-靶点相互作用预测模型, 得到了消栓通路方中的活性成分, 并利用网络药理学方法, 建立了活性成分-药物靶点-通路网络, 从而揭示了消栓通路方抗脑卒中的药效物质基础和作用机制。

本研究通过机器学习算法朴素贝叶斯和递归分隔以及 Cdocker 分子对接方法, 对脑卒中临床应用经典处方消栓通路方中的化学成分进行作用靶点预测, 得到不错的结果。其中有 3 个小分子值得特别留意: ①姜黄素, 是一种植物多酚, 其具有广泛的生物学效应而引起广泛关注。本研究构建的预测模型产生的活性化合物中, 共有 4 个姜黄素类化合物入选, 分别是姜黄素、1,2-二氢姜黄素、去甲氧基姜黄素^[26]和双去甲氧基姜黄素, 对 MMP9 具有潜在活性, 提示姜黄素类化合物对脑卒中可能的治疗作用。许多研究表明, 姜黄素

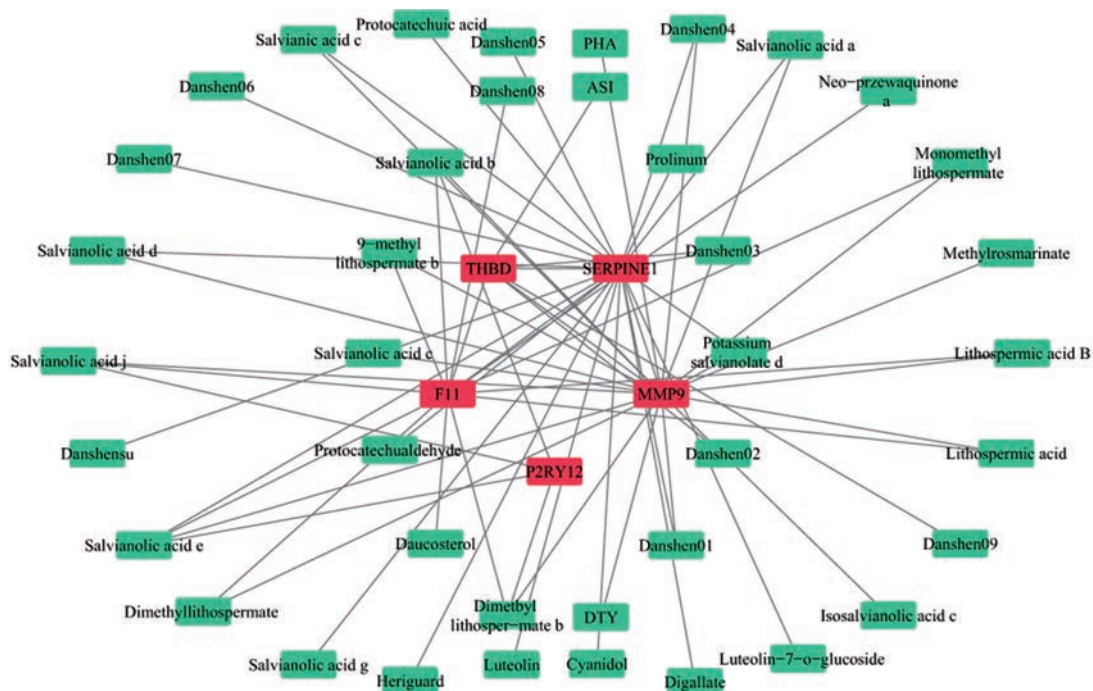


Figure 3 The chemical constituent-target network. The rectangle means active constituents from Radix Salviae, and the oval means stroke related targets. Danshen01: (2*R*)-3-(3,4-Dihydroxyphenyl)-2-[(*Z*)-3-(3,4-dihydroxyphenyl)acryloyl]oxy-propionic acid; Danshen02: (2*S*)-2-Amino-3-[(2*R*)-2-amino-3-hydroxy-3-oxopropyl]disulfanylpropanoic acid; Danshen03: (2*S*)-3-(3,4-Dihydroxyphenyl)-2-hydroxypropanoic acid; Danshen04: (2*S*,3*S*)-2-(3,4-Dihydroxyphenyl)-7-hydroxy-4-[(*E*)-3-hydroxy-3-oxoprop-1-enyl]-2,3-dihydrobenzofuran-3-carboxylic acid; Danshen05: (*E*)-3-[2-(3,4-Dihydroxyphenyl)-7-hydroxy-benzofuran-4-yl]acrylic acid; Danshen06: (*Z*)-3-[2-[(*E*)-2-(3,4-Dihydroxyphenyl)vinyl]-3,4-dihydroxy-phenyl]acrylic acid; Danshen07: 1-(3,4-Dihydroxyphenyl)-2-hydroxyethanone; Danshen08: 6-*O*-Syringyl-8-*O*-acetyl shanzhiside methyl ester; Danshen09: (2*S*)-2-Amino-5-(diaminomethylideneamino) pentanoic acid

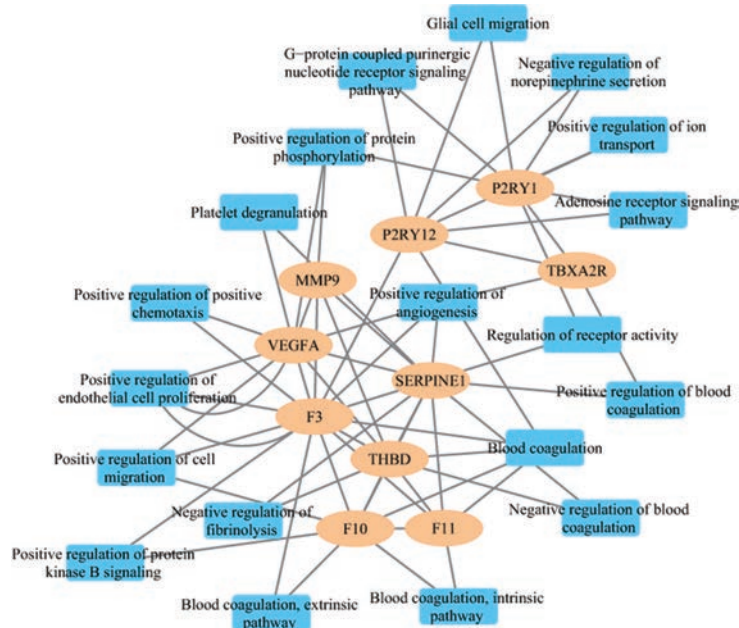


Figure 4 The target-pathway network based on the enrichment analysis of Gene-Ontology (GO) biological pathway

对脑卒中有治疗作用, 其作用机制包括抗氧化、清除氧自由基的作用、抗炎作用、降低胆固醇作用和抗凝作用^[3]。此外, 还有研究表明, 姜黄素可能通过对表观遗传

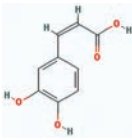
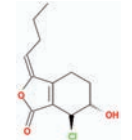
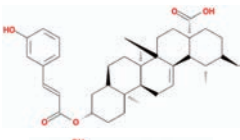
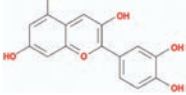
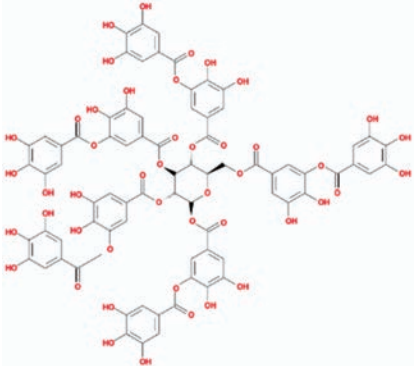
的影响达到抗脑卒中的作用^[27]。另外, Li等^[28]研究显示姜黄素通过下调炎症 (NF- κ B、ICAM-1、MMP-9) 和凋亡 (caspase-3) 相关标志物, 保护脑缺血再灌注造

成的损伤,达到治疗作用。上述文献^[26,27]报道进一步确证了本研究建模预测结果的可靠性,也说明姜黄素类化学成分可以作为脑卒中治疗的重要组成部分。② 肉桂醛大量存在于肉桂等植物体内,研究表明其具有抗炎、抗氧化、抗血小板聚集等多种生物学作用。模型预测其对 F11、MMP9、TBXA2R 和 VEGFA 四个靶点具有潜在活性,提示肉桂醛对脑卒中潜在的治疗作用。Zhao 等^[29]研究显示,肉桂醛可以通过抑制 TLR4/ TRAF6/NF- κ B 炎症通路,减轻缺血脑组织的炎症反应,从而在缺血再灌注损伤后对脑产生保护作用。Chen 等^[30]类似研究显示,肉桂醛通过减弱 iNOS, COX-2 表达和 NF- κ B 信号通路,抑制神经炎症,从而对脑产生保护作用。此外, Ballabeni 等^[31]以 *Ocotea quixos* 精油为实验药物,其主要成分为反式肉桂醛,结果显示肉桂醛具有抗血小板作用和血管舒张作用,抑制小鼠血栓的形成。由文献^[29-31]可以发现,肉桂醛可以通过血管再通和神经保护两种方式治疗脑卒中,是潜在优异的抗脑卒中的活性成分。③ 丹酚酸为丹参水溶性主要有效成分^[32],为一类结构相似的酚酸类化合物。本研究

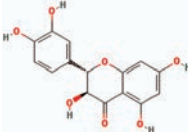
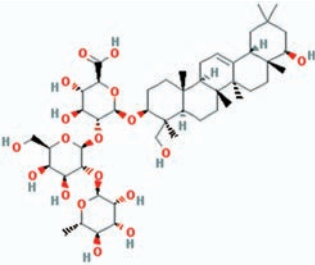
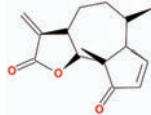
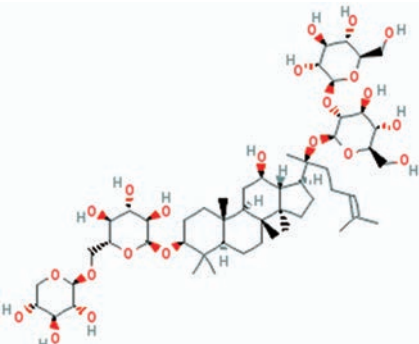
构建模型预测结果中,多种丹酚酸入选,包括丹酚酸 A、B、C、D、E、G、J、N 和异丹酚酸 C,共同作用于 MMP9、SERPINE1、F11 和 P2RY12 四个靶点,提示丹酚酸对脑卒中可能的治疗作用。目前已有许多研究报告丹酚酸对脑损伤的保护作用,且临床上已应用丹酚酸注射液用于脑卒中的治疗。其中, Mahmood 等^[33]研究显示,丹酚酸 A 通过抑制 eNOS 解偶联和 ONOO⁻ 形成神经血管保护作用。另外,丹酚酸 A 可减弱血影蛋白和钙调神经磷酸酶的分解,从而保护大脑免受缺血/再灌注损伤,进而达到治疗脑卒中的作用。Fan 等^[34]研究显示,丹酚酸 B 可以通过降低 GFAP、Iba1、IL-1 β 、IL-6、TNF- α 和 cleaved-caspase 3 的表达,抑制星形胶质细胞和小胶质细胞的过度活化,达到脑卒中治疗中对神经细胞的保护作用。由此可知,多种丹酚酸化合物都存在治疗脑卒中的潜在能力,预测得到的多种丹酚酸化合物值得同样关注。

除受到关注广为研究的化学成分外,预测所得活性化学成分中还包含其他一些典型化学成分结构,经整合分析,如表 7 示。

Table 7 The typical examples of active constituents

Name	Source	Structure	Target
Caffeic acid	Chuanxiong		SERPINE1
Senkyunolide-L	Chuanxiong		F11
Xiongterpene	Chuanxiong		F11
Cyanidol	Danshen		SERPINE1
Tannin	Danshen		F11, SERPINE1

Continued

Name	Source	Structure	Target
(-)-Taxifolin	Guizhi		SERPINE1
Soyasaponin I	Huaihua		F11
Ambrosin	Muxiang		TBXA2R
Ginsenoside Rg3	Sanqi		THBD

综上所述,本研究利用机器学习等算法,借助网络药理学研究方法和技术,构建了化学成分-靶点-通路网络,对消栓通络方的药效物质基础和网络作用机制进行了较为深入的研究。本研究发现了许多潜在的多靶点活性成分通过血管再通和神经保护多种方式治疗脑卒中,为消栓通络方的临床应用和后续研究提供重要信息,同时还发现了许多潜在的多靶点活性成分用于缓解脑卒中的症状并治疗脑卒中,也为后续开展的中药复方抗脑卒中研究提供参考。

References

- [1] Jin BX. Clinical study on the prognostic factors of thrombolysis with alteplase in acute ischemic stroke [J]. *Chin Heal Ind (中国卫生产业)*, 2013, 10: 101-102.
- [2] National institute of neurological disorders and stroke rt-PA stroke study group. Tissue plasminogen activator for acute ischemic stroke [J]. *New Engl J Med*, 1995, 333: 1581-1588.
- [3] Liao CY. Study on therapeutic targets and drugs of stroke [J]. *Chin J Pharm Toxicol (中国药理学与毒理学杂志)*, 2016, 30: 1264-1272.
- [4] Tan CB, Gao M, Wang HQ, et al. Study on the selection of the effective components of Xiaoshuan Tongluo Recipe [J]. *Pharmacol Clin Chinese Mater Med (中药药理与临床)*, 2011, 27: 75-78.
- [5] Sorensen SV, Dewilde S, Singer DE, et al. Cost-effectiveness of warfarin: trial versus "real-world" stroke prevention in atrial fibrillation [J]. *Am Heart J*, 2009, 157: 1064-1073.
- [6] Grysiwicz RA, Thomas K, Pandey DK. Epidemiology of ischemic and hemorrhagic stroke: incidence, prevalence, mortality, and risk factors [J]. *Neurol Clin*, 2008, 26: 871-895.
- [7] Zhang BY, Zheng YF, Pang XC, et al. Study on the material basis and network mechanism of anti-tumor effect of traditional Chinese Medicine [J]. *World J Tradit Chin Med (世界中医药)*, 2018, 13: 1997-2009.
- [8] Xie T, Song S, Li S, et al. Review of natural product databases [J]. *Cell Prolif*, 2015, 48: 398-404.
- [9] Bachurin SO, Bovina EV, Ustyugov AA. Drugs in clinical trials for Alzheimer's disease: the major trends [J]. *Med Res Rev*. 2017, 37: 1186-1225.
- [10] Chen X, Ji ZL, Chen YZ. TTD: therapeutic target database [J]. *Nucleic Acids Res*, 2002, 30: 412-415.

- [11] Liu L, Du B, Zhang H, et al. A network pharmacology approach to explore the mechanisms of Erxian decoction in polycystic ovary syndrome [J]. *Chin Med*, 2018, 13: 46.
- [12] Lee AY, Park W, Kang TW, et al. Network pharmacology-based prediction of active compounds and molecular targets in Yijin-Tang acting on hyperlipidaemia and atherosclerosis [J]. *J Ethnopharmacol*, 2018, 221: 151-159.
- [13] Friedman N, Geiger D, Goldszmidt M. Bayesian network classifiers [J]. *Mach Learn*, 1997, 29: 131-163.
- [14] Stone M. Cross-validation and multinomial prediction [J]. *Biometrika*, 1974, 61: 509-515.
- [15] Rogers D, Brown RD, Hahn M. Using extended-connectivity fingerprints with Laplacian-modified Bayesian analysis in high-throughput screening follow-up [J]. *J Biomol Screen*, 2005, 10: 682-686.
- [16] Cereto-Massagué A, Ojeda MJ, Valls C, et al. Molecular fingerprint similarity search in virtual screening [J]. *Methods*, 2015, 71: 58-63.
- [17] Fang J, Li Y, Liu R, et al. Discovery of multitarget-directed ligands against Alzheimer's disease through systematic prediction of chemical-protein interactions [J]. *J Chem Inform Model*, 2015, 55: 149-164.
- [18] Kirchmair J, Markt P, Distinto S, et al. Evaluation of the performance of 3D virtual screening protocols: RMSD comparisons, enrichment assessments, and decoy selection — what can we learn from earlier mistakes? [J]. *J Comput Aid Mol Design*, 2008, 22: 213-228.
- [19] Hothorn T, Hornik K, Zeileis A. Unbiased recursive partitioning: a conditional inference framework [J]. *J Comput Graph Statist*, 2006, 15: 651-674.
- [20] Shannon P, Markiel A, Ozier O, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks [J]. *Genome Res*, 2003, 13: 2498-2504.
- [21] Ashburner M, Ball CA, Blake JA, et al. Gene Ontology: tool for the unification of biology [J]. *Nat Gene*, 2000, 25: 25-29.
- [22] Pang XC, Wang Z, Fang JS, et al. Network pharmacology study of effective constituents of traditional Chinese medicine for Alzheimer's disease treatment [J]. *Acta Pharm Sin (药学报)*, 2016, 51: 725-731.
- [23] Xu L, Jiang W, Pang X, et al. Network pharmacology study of the effective constituents in the Compound Yizhihao against influenza disease [J]. *Acta Pharm Sin (药学报)*, 2017, 52: 745-752.
- [24] Bajusz D, Rácz A, Héberger K. Why is Tanimoto index an appropriate choice for fingerprint-based similarity calculations [J]. *J Cheminform*, 2015, 7: 20.
- [25] Nour M, Scalzo F, Liebeskind DS. Ischemia-reperfusion injury in stroke [J]. *Int Neurol*, 2012, 1: 185-199.
- [26] Ahmad N, Umar S, Ashafaq M, et al. A comparative study of PNIPAM nanoparticles of curcumin, demethoxycurcumin, and bisdemethoxycurcumin and their effects on oxidative stress markers in experimental stroke [J]. *Protoplasma*, 2013, 250: 1327-1338.
- [27] Chen YH, Tang YH. The epigenetic effect of curcumin on stroke prevention [J]. *Mod Med J China (中国现代医药杂志)*, 2016, 18: 98-101.
- [28] Li W, Suwanwela NC, Patumraj S. Curcumin prevents reperfusion injury following ischemic stroke in rats *via* inhibition of NF- κ B, ICAM-1, MMP-9 and caspase-3 expression [J]. *Mol Med Rep*, 2017, 16: 4710-4720.
- [29] Zhao J, Zhang X, Dong L, et al. Cinnamaldehyde inhibits inflammation and brain damage in a mouse model of permanent cerebral ischaemia [J]. *Br J Pharmacol*, 2015, 172: 5009-5023.
- [30] Chen YF, Wang YW, Huang WS, et al. Trans-cinnamaldehyde, an essential oil in cinnamon powder, ameliorates cerebral ischemia-induced brain injury *via* inhibition of neuroinflammation through attenuation of iNOS, COX-2 expression and NF- κ B signaling pathway [J]. *Neuromol Med*, 2016, 18: 322-333.
- [31] Ballabeni V, Tognolini M, Bertoni S, et al. Antiplatelet and antithrombotic activities of essential oil from wild *Ocotea quixos* (Lam.) Kosterm. (Lauraceae) calices from Amazonian Ecuador [J]. *Pharmacol Res*, 2007, 55: 23-30.
- [32] Du GH, Zhang JT. Research progress of salvianolic acid, an effective water-soluble component of *Salvia miltiorrhiza* [J]. *Basic Clin Med (基础医学与临床)*, 2000, 20: 10-14.
- [33] Mahmood Q, Wang GF, Wu G, et al. Salvianolic acid A inhibits calpain activation and eNOS uncoupling during focal cerebral ischemia in mice [J]. *Phytomedicine*, 2017, 25: 8-14.
- [34] Fan Y, Luo Q, Wei J, et al. Mechanism of salvianolic acid B neuroprotection against ischemia/reperfusion induced cerebral injury [J]. *Brain Res*, 2018, 1679: 125-133.