

· 数据挖掘 ·

基于网络药理学方法预测银杏叶治疗心脑血管类疾病的有效成分及其潜在靶点

苑婕, 李晓杰, 陈超*, 宋向岗, 王淑美*
(广东药学院中药学院, 广州 510006)

[摘要] 目的:采用网络药理学方法,研究银杏叶“成分-靶点-心脑血管相关疾病”复杂网络,预测银杏叶治疗心脑血管类疾病的有效成分及其潜在靶点。方法:建立药物分子-靶点相互作用的随机森林预测模型,对银杏叶化学成分进行靶点预测,然后构建和分析银杏叶“成分-靶点-心脑血管相关疾病”复杂网络。结果:对银杏叶的34个化学成分预测出多个作用靶点,结果得到了文献验证;进一步网络分析的结果初步揭示了银杏叶治疗心脑血管类疾病的有效成分和潜在靶点。结论:本方法可用于研究中药的物质基础和分子作用机制。

[关键词] 网络药理学; 银杏叶; 靶点预测; 心脑血管疾病

[中图分类号] R285.5 **[文献标识码]** A **[文章编号]** 1005-9903(2014)11-0208-05

[doi] 10.13422/j.cnki.syfjx.2014110208

Predicting Effective Components and Potential Targets of *Ginkgo biloba* on Cardio-cerebral Vascular Diseases Based on Network Pharmacology

YUAN Jie, LI Xiao-jie, CHEN Chao*, SONG Xiang-gang, WANG Shu-mei*

(School of Traditional Chinese Medicine, Guangdong Pharmaceutical University, Guangzhou 510006, China)

[Abstract] **Objective:** To construct the complicated compound-target-cardio-cerebral-vascular-disease network of *Ginkgo biloba* and to predict effective components and potential targets of *G. biloba* on cardio-cerebral vascular diseases by network pharmacology. **Method:** The drug-target interaction model was established by random forest algorithm, which was then applied to predict the potential targets interacted with compounds of *G. biloba*. The complicated compound-target-disease network of *G. biloba* was then constructed and analyzed. **Result:** The predicted targets of 34 compounds of *G. biloba* were validated by literatures. Furthermore, network analysis was performed, preliminarily revealing the active ingredients and potential targets of *G. biloba* in treatment of cardio-cerebral vascular diseases. **Conclusion:** The present approach can be used to study the material base and the molecular mechanism of traditional Chinese medicine.

[Key words] network pharmacology; *Ginkgo biloba*; targets prediction; cardio-cerebral vascular diseases

网络药理学是在高通量组学数据分析、计算机模拟计算

及网络数据库检索的基础上,进行生物网络的构建及分析,进而进行药物作用机制的研究及创新药物的发现。网络药理学从药物、靶点与疾病间相互作用的整体性和系统性出发,采用复杂网络模型表达和分析研究对象的药理学性质,特别适宜研究中药多成分-多靶点的作用关系,有利于揭示中药成分复杂作用机制^[1-2]。网络药理学概念自2007年被提出以来,国内外已有不少相关的研究报道,如范晓辉等^[3]采用网络药理学方法研究复方丹参方,网络分析表明9个活性成分均作用于多个靶点,呈现出方剂多成分、多靶点、整合调节作用特点;乔延江等^[4]采用网络药理学方法研究活血化瘀中药,通过网络对比计算,筛选得到中药与西药治疗冠心

[收稿日期] 20131122(007)

[基金项目] 广东省自然科学基金项目(S2012010009166);国家自然科学基金项目(81274060)

[第一作者] 苑婕,在读硕士,从事中药信息技术研究, E-mail: yuanjie2364@163.com

[通讯作者] *陈超,副教授,从事中药信息技术研究, Tel:020-39352181, E-mail: cep02cc@gmail.com

*王淑美,教授,从事中药质量控制研究, Tel:020-39352559, E-mail: shmwang@sina.com

病 138 个相同作用节点和 397 个中药独特作用节点,表明中药治疗冠心病确实存在有别于西药的独特作用机制;王毅等^[5]以中药附子的 22 个化学成分预测出多个作用靶点,所建网络模型中每个化合物的平均靶点数为 16.3,平均每个靶点与 4.77 个化合物相关联,反映出中药“多成分、多靶点”的特点。

银杏叶为银杏科银杏属植物银杏(*Ginkgo biloba* L.)的树叶。银杏叶提取物(GBE)是从银杏叶中分离纯化的提取物,具有清除氧自由基、拮抗血小板活化因子、降血脂、增强中枢神经系统功能、调节神经递质和激素水平、改善血液流变状态等作用^[6],因而在临床上被广泛用于预防和治疗心脑血管类疾病。但是,其药效成分及关键作用靶点尚待系统研究和揭示。本文以银杏叶为研究对象,利用随机森林算法建立药物分子-靶点相互作用预测模型,对银杏叶化学成分进行靶点预测,然后构建银杏叶“成分-靶点-心脑血管相关疾病”复杂网络,从网络药理学角度预测银杏叶治疗心脑血管类疾病的有效成分及其潜在靶点,为开发更好的心脑血管类候选药物提供参考依据。

1 方法

1.1 数据集 从 KEGG^[7-8]数据库下载小分子化学药的分子结构及其相应靶蛋白受体,剔除缺乏实验数据的药物-靶点,最终获得 4 782 条药物-靶点配对,其中包括 2 711 条药物-酶蛋白受体、1 365 条药物-离子通道蛋白受体、620 条药物-G 蛋白受体、86 条药物-核蛋白受体。以上数据作为本研究的阳性样本集。

阴性样本集:将阳性样本集中的药物-靶点配对拆开,各类分别进行重新组合配对,剔除已包含在阳性样本集中的配对结果,最后随机挑选出两倍于阳性样本集的配对结果,即生成阴性样本集。

训练集:分别合并上述 4 类药靶配对阴阳样本,即得训练集数据。

1.2 分子描述符 应用 PowerMV (www.niss.org/PowerMV) 和 ProFeat (<http://jing.cz3.nus.edu.sg/cgi-bin/prof/prof.cgi>) 分别计算药物和靶蛋白的分子描述符,获得配对数据合计 7 202 维。如此高维数据若直接用于建模,一是存在信息冗余,影响模型精度;二是增大计算量,影响模型速度。因此,本研究采用主成分分析法(Principle Component Analysis, PCA)对上述数据进行降维处理,按保留原始信息 99.7% 的信息量进行降维,获得数据矩阵分别为:8 133 × 17, 4 095 × 4, 1 860 × 24, 258 × 17。然后,再按下式进行归一化处理。

$$x(i) = [x_0(i) - \text{Min}(x_0)] / [\text{Max}(x_0) - \text{Min}(x_0)] \quad (1)$$

其中 $x_0(i)$ 是原始数据; $x(i)$ 是归一化后数据,其值在 -1 到 1 之间。

1.3 模型 随机森林(Random Forest, RF)是 Leo Breiman 于 2001 年提出的一个组合分类器算法^[9],它是利用 boot strap 重抽样方法从原始样本中抽取多个样本,对每个 boot strap 样本进行决策树建模,然后组合多棵决策树的预测,通过投票得出最终预测结果^[10]。在树的每个节点处,从 M 个特征中

随机挑选 $mtry$ 个特征 ($mtry < M$),通常假设 $mtry = \sqrt{M}$ 。按照节点不纯度最小的原则,从这 $mtry$ 个特征中选出一个特征进行分支生长。单棵分类树进行充分生长,使每个节点的不纯度达到最小,不进行通常的剪枝操作。将生成的 $ntree$ 棵分类树组成随机森林,用随机森林分类器对新的数据进行判别与分类,分类结果按树分类器的投票多少而定,即:

$$c = \text{argmax}_c \left(\frac{1}{ntree} \sum_{k=1}^{ntree} I(h(x, \theta_k) = c) \right) \quad (2)$$

在上述随机森林构建过程中,生成每棵分类树时所应用的自助样本集从原始数据中随机选取,每个节点处所采用的变量也是从所有变量中随机选取。此 2 次随机过程使得随机森林具有较大的稳定性和较优的性能,能在较好地容忍噪声的同时提高分类精确度。RF 作为一种优秀的数据挖掘算法,目前已被广泛运用于生物、医药、物理、化学等领域。

2 结果与讨论

2.1 模型优化与评价 应用随机森林法建立模型,需要优化其主要参数,即树数($ntree$)和每棵树的节点数($mtry$)。本文采用网格搜索法进行参数优化,以获得最高的 K 折交叉验证(K -fold cross-validation)正确率。所谓 K 折交叉验证,即将初始采样分割成 K 个子样本,一个单独的子样本被保留作为验证模型的数据,其他 $K-1$ 个样本用来训练。交叉验证重复 K 次,每个子样本验证一次,平均 K 次的结果或者使用其他结合方式,最终得到一个单一估测。本研究中 $K=10$,正确率($accuracy$)见公式(3)定义,优化结果列于表 1。

$$\text{正确率}(\text{accuracy}) = (1 - NY/Y) \times 100\% \quad (3)$$

其中 NY 代表分类错误的样本个数, Y 代表所有样本个数。

表 1 本文模型的参数优化

模型	树数	节点数	正确率/%
1 酶蛋白	500	1	71.34
2 离子通道	400	2	67.08
3 G 蛋白	200	1	73.17
4 核蛋白	100	1	67.83

2.2 银杏叶作用靶点预测 从化学专业数据库(<http://www.organchem.csdb.cn/scdb/>)和 PubChem 数据库(<http://www.ncbi.nlm.nih.gov/pccompound>)下载已报道的银杏叶所含化学成分,共计 69 个。为去除类药性不高的非药效成分,本研究计算夹角余弦(Cosine)比较银杏叶成分与训练集中小分子药物的化学空间相似性。夹角余弦越大表示两个向量的夹角越小,即相似性越高。本文中夹角余弦的阈值设定为 0.96,从而在 69 个银杏叶成分中挑选出了 34 个与训练集中小分子药物化学空间相近的化合物。

将上述 34 个化合物与四种已知的靶点进行组合,构建预测集。预测集按训练集相应的方法计算分子描述符、降维和归一化等处理,然后代入 RF 模型进行靶点预测。表 2 列出了与银杏叶化学成分作用频次较高的前 15 个靶点,提示

它们可能是银杏叶治疗心脑血管类疾病的关键靶标;进一步地,应用 Cytoscape 软件^[11]构建银杏叶成分-靶点-疾病网络模型,见图 1。网络中黄色三角形节点表示银杏叶化学成分;

红色圆形节点表示靶点;绿色正方形节点表示心脑血管相关疾病;若某一靶点为某化合物的潜在作用靶点,则以边相连。

表 2 本文预测的银杏叶 15 个潜在作用靶点

No.	靶点名称	名称缩写	类别	心脑血管相关疾病
1	α_{2A} 肾上腺素受体	ADRA2A	G 蛋白	心衰,高血压,缺血性心脏病
2	细胞色素 P450 2C19 酶	CYP2C19	酶	室性心动过速,炎症
3	组胺 H1 受体	HRH1	G 蛋白	高血压,局部贫血,全身动脉血管舒张
4	核受体	ER	核蛋白	脑损伤,心血管疾病,冠状动脉粥样硬化
5	电压依赖型钙通道 α -2B 亚基	CACN A1B	离子通道蛋白	原发性高血压,心搏徐缓,卒中
6	γ -氨基丁酸 γ -2 亚基受体	GABRG2	离子通道蛋白	脑损伤
7	α_{2C} 肾上腺素受体	ADRA2C	G 蛋白	心衰
8	磷酸二酯酶 3B	PDE3B	酶	心衰,坏死
9	β_1 肾上腺素受体	ADRB1	G 蛋白	心律失常,心血管疾病,冠状动脉心脏疾病,扩张型心肌病,高血压
10	钠通道蛋白 5- α 亚基	SCN5A	离子通道蛋白	心律不齐,持续性室性心动过速
11	钙激活钾通道蛋白 β -1 亚基	KCN MB1	离子通道蛋白	高血压,缺氧,原发性高血压,心肌炎,胆固醇
12	电压依赖型钙通道 α -1A 亚基	CACNL1A4	离子通道蛋白	脑缺血
13	血管紧张素转化酶	ACE	酶	心衰,高血压,血管疾病
14	环氧合酶 2	COX-2	酶	心肌梗死,卒中,病理性血管生成
15	5-羟色胺 2A 受体	5-HT-2A	G 蛋白	动脉栓塞和血栓症,原发性高血压

从图 1 可见,银杏叶所含 34 个成分与靶点存在复杂的网络关系,网络模型中平均相邻节点数目为 10.517,反映出中药多成分-多靶点-多疾病相关的特点。同时可见 5 号化合物(山奈酚)、20 号化合物(白果苦内酯 A)、31 号化合物(异鼠李素)、40 号化合物(表儿茶素)、50 号化合物(银杏内酯 A)、52 号化合物(银杏酮)、62 号化合物(白果苦内酯 B)、63 号化合物(白果苦内酯 C)、64 号化合物(白果苦内酯 J)、65 号化合物(白果苦内酯 M)等成分在网络中连接度较高(见表 3),提示其可能是银杏叶发挥药效的重要成分。

表 3 图 1 中连接度较高的前 10 位化合物

化合物编号	连接度	化合物编号	连接度
5	48	52	33
20	29	62	26
31	39	63	33
40	40	64	28
50	44	65	27

2.3 文献验证 对计算得到的银杏叶与心脑血管疾病相关的潜在作用靶点进行了文献验证,发现预测出的四类受体相应靶点均有相关文献报道。

2.3.1 酶蛋白受体 潜在作用靶点主要是 CYP 酶、PDE 酶、ACE 酶、COX 酶等。目前国内外关于 GBE 对细胞色素 CYP 酶的影响没有得到一致的结果^[12]。一项以人 P450 微粒体

为研究对象的试验中,荧光法测定 GBE 及其黄酮类成分和萜类内酯成分对 CYP450 的活性影响,发现 GBE 能抑制 CYP2C9、CYP1A2、CYP2E1 和 CYP3A4 的表达,黄酮类成分能抑制 CYP2C9、CYP1A2、CYP2E1 和 CYP3A4 的表达,萜类内酯成分只显著抑制 CYP2C9 的表达^[13]。Sugiyama T 等^[14]采用小鼠研究 GBE 对甲苯磺丁脲降血糖作用的影响,实验提示长期服用 GBE,诱导了小鼠的 CYP2C9;而单独一次服用 GBE,可能抑制 CYP2C9 的活性。人体临床试验关于 GBE 对 CYP2C9 的影响并没有得到与上述一致的结果。Yin O Q 等^[15]研究提示,健康志愿者长期服用 GBE 可能对 CYP2C19 活性具有诱导作用,并且 GBE 对 CYP2C19 的诱导作用具有基因型依赖性。

谭萍等^[16]研究发现,GBE 可使急性脑梗死患者血小板 PDE3 活性进一步下降,加强 PDE3 在脑梗死急性期的保护作用,有效抑制血小板聚集,改善血液循环及神经功能缺损。耿秀芳等^[17]研究表明,银杏叶单黄酮山奈酚、槲皮素对 ACE 的抑制率分别是 74%,53%。张爱林^[18]研究表明,复方银杏滴丸可以显著降低模型大鼠脑皮质 COX-2mRNA 表达,影响脑缺血后 PAF-COX-2 信号通路在神经元损伤中发挥作用。

2.3.2 离子通道受体 潜在作用靶点主要是 GABA 受体、CACN 受体、KCN 受体、SCN 受体等。例如,王何等^[19]研究发现,脑缺血后,单纯亚低温干预及分别在常温或亚低温状态下给予 GBE 干预,均使 Glu、MDA 水平显著降低,GABA、SOD、GSH-Px 水平显著上调,亚低温状态下 GBE 干预的效果

- [4] 黄明峰,张燕玲,任真真,等. 网络药理学方法探讨活血化瘀中药治疗冠心病作用机理[J]. 世界科学技术——中医药现代化,2012(5):1969.
- [5] 吴磊宏,高秀梅,王林丽,等. 附子多成分作用靶点预测及网络药理学研究[J]. 中国中药杂志,2011,36(21):2907.
- [6] 李红梅,刘顺良,姜静岩. 银杏叶提取物对心脑血管疾病的药理作用研究进展[J]. 时珍国医国药,2002,13(2):105.
- [7] Goto S, Nishioka T, Kanehisa M. LIGAND; chemical_database_for_enzyme_reactions1 [J]. Bioinformatics, 1998,14(7):591.
- [8] Kanehisa M, Goto S, Hattori M, et al. From genomics to chemical genomics; new developments in KEGG[J]. Nucleic Acids Res,2006,34(Database issue):D354.
- [9] L B. Random Forests [J]. Machine Learning,2001,45(1):5.
- [10] 袁敏,胡秀珍. 随机森林方法预测膜蛋白类型[J]. 生物物理学报,2009(5):349.
- [11] Christopher L, Andrew H. Navigating chemical space for biology and medicine [J]. Nature, 2004, 432(7019):855.
- [12] 陶共由,周宏灏. 银杏叶提取物对细胞色素 P450 影响的研究进展[J]. 中国临床药理学与治疗学,2008,13(9):1071.
- [13] Gaudineau C, Beckerman R, Welbourn S. Inhibition of human P450 enzymes by multiple constituents of the Ginkgo biloba extract [J]. Biochem Biophysical Res Communications,2004,318(4):1072.
- [14] Tomomi S, Yoko K, Kazumasa S, et al. Ginkgo biloba extract modifies hypoglycemic action of tolbutamide via hepatic cytochrome P450 mediated mechanism in aged rats[J]. Life Sciences,2004,75(9):1113.
- [15] P Y O Q, Brian T, Y W M M, et al. Pharmacogenetics and herb-drug interactions: experience with Ginkgo biloba and omeprazole [J]. Pharmacogenetics, 2004, 14(12):841.
- [16] 谭萍,刘雁,郝勇,等. 银杏叶提取物对急性脑梗死患者血小板磷酸二酯酶 3 及相关因素的影响[J]. 华国防医学杂志,2011,25(3):206.
- [17] 耿秀芳,孟祥国,栾美英,等. 银杏叶黄酮扩血管机理的探讨[J]. 中医药研究,1996(1):55.
- [18] 张爱林. 复方银杏滴丸防治血管性痴呆的作用及其机理分析[D]. 北京:北京中医药大学,2005.
- [19] 王何,陈红,杨明秀,等. 不同脑温状态下银杏叶提取物对大鼠缺血脑组织氨基酸及自由基的影响[J]. 现代中西医结合杂志,2011,20(28):3530.
- [20] 张绪国,王家宁,李建军,等. 银杏叶提取物对豚鼠心室肌细胞动作电位及 L 型钙离子通道的影响[J]. 中国临床药理学与治疗学,2004,9(3):345.
- [21] 孙柯,高尔,李宁. 银杏内酯 B 对大鼠海马神经元钾离子通道的作用[Z]. 济南:山东省药学会,20051.
- [22] 黄璞,刘力强,韩勇,等. 银杏叶提取物对中枢神经系统药理作用的研究进展[J]. 中国医院药学杂志,2005,25(6):553.
- [23] Caltagirone A, Weiss G, Pantopoulos K. Modulation of cellular iron metabolism by hydrogen peroxide. Effects of H₂O₂ on the expression and function of iron-responsive element-containing mRNAs in B6 fibroblasts [J]. J Biol Chem,2001,276(23):19738.
- [24] 黄九龙. 银杏叶总黄酮对豚鼠肠系膜下神经节细胞 ls-EPSP 的影响[D]. 合肥:安徽医科大学,2007.
- [25] Min O S, Hyuck C K. Antiestrogenic activities of Ginkgo biloba extracts [J]. J Steroid Biochem Mol Biol,2006,100(4/5):167.
- [26] 朱诗平,刘康丽,葛茂功,等. 银杏叶提取物对腺嘌呤致肾间质纤维化大鼠肾组织 ACE2 蛋白及 mRNA 的影响[J]. 中成药,2011,33(11):1854.
- [27] 周燕红,白育庭,高卉,等. 银杏叶提取物对大鼠实验性结肠炎 COX-2 表达的影响[J]. 咸宁医学院学报:医学版,2008,22(1):4.

[责任编辑 聂淑琴]

欢迎投稿

欢迎订阅