

用主成分分析探索不同中药组分配伍干预脑缺血的基因表达模式

荆志伟¹ 周才秀¹ 王 忠¹ 张占军² 王丽颖¹ 胡木林¹ 武红莉¹ 张 亮³ 李澎涛⁴ 王永炎¹

(1. 中国中医科学院中医临床基础医学研究所, 北京市东直门内南小街 16 号, 100700; 2. 北京师范大学认知神经科学与学习国家重点实验室; 3. 生物芯片北京国家工程技术研究中心; 4. 北京中医药大学)

[摘 要] 目的 比较清开灵主要组分黄芩苷、栀子苷及其配伍干预脑缺血损伤的基因表达模式。方法 12 只 4 周龄小鼠随机分为黄芩苷组(BA 组)、栀子苷组(JA 组)、栀子苷+黄芩苷组(BJ 组)、尼莫地平组(NM 组), 建立全脑缺血再灌注小鼠模型后, 各组分别给予尾静脉注射黄芩苷、栀子苷、栀子苷+黄芩苷以及尼莫地平。抽提小鼠脑组织的总 RNA, 制备小鼠全基因组表达谱, 对其进行基因本体论功能分类, 构建相关系数矩阵, 对差异表达的细胞黏附相关基因进行主成分分析。结果 获取 16 463 个小鼠脑测全基因组表达谱, 筛选出 32 个差异表达的细胞黏附相关基因。主成分分析结果显示, 4 组的前 2 个主成分的累积贡献度为 100%, BJ 组总得分大于单一组分, NM 组对细胞黏附相关基因干预的总得分高于 BA 和 JA 组。各次检测的主成分分布显示, NM 组和 3 个中药组均距离较远, BJ 组和 BA 组的距离均近于 JA 组。结论 黄芩苷、栀子苷及其配伍均能有效干预细胞黏附相关基因, 配伍后的干预对黏附相关功能基因的影响大于单一组分, 其基因表达模式和黄芩苷相近。

[关键词] 黄芩苷; 栀子苷; 组分配伍; 脑缺血; 基因芯片; 主成分分析; 细胞黏附

Principal Component Analysis of Gene Expression of Baicalin, Jasminoidin and Their Combination in Experimental Stroke

JING Zhiwei¹, ZHOU Caixiu¹, WANG Zhong¹, ZHANG Zhanjun², WANG Liying¹, HU Mulin¹, WU Hongli¹, ZHANG Liang³, LI Pengtao⁴, WANG Yongyan¹

(1. Institute of Clinical Fundamental Chinese Medicine, China Academy of Chinese Medicine Sciences, Beijing, 100700; 2. State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University; 3. Biochip Beijing National Engineering Research Center; 4. Beijing University of Chinese Medicine)

ABSTRACT Objective To explore the gene expression pattern of Baicalin and Jasminoidin, which are main ingredients of Qingkailing (Chinese medicine for clearing heat to restore consciousness), and their combination in cerebral ischemia-reperfusion injury.

Methods Twelve 4-week-old mice were randomized into Baicalin group (Group BA), Jasminoidin group (Group JA), Baicalin and Jasminoidin combination group (Group BJ), Nimodipine group (Group NM). After making models of whole brain ischemia and reperfusion, they were administered Baicalin, Jasminoidin, their combination, and Nimodipine respectively. Total RNA was obtained from brain tissue to prepare the whole gene profile for gene ontology classification by gene function to construct the related coefficient matrix to make a comparative analysis with the principal component analysis (PCA) on the cell adhesion genes of different expressions. **Results** The 16 463 whole gene profiles were obtained, from which 32 cell adhesion-related genes of different expressions were selected. PCA showed that the accumulation contribution rate of top two principal components of the four groups was 100%. The total score of Group BJ was more than that of Group BA and Group JA respectively. The total score of Group NM was higher than that of group BA and JA. The distribution of principal components showed that there was a long distance between Group NM and the three Chinese medicine groups, and the distance from Groups BJ to BA was nearer than to group JA. **Conclusion** Baicalin and Jasminoidin can effectively intervene the adhesion-related genes. The effect on adhesion-related genes of the combination of the two is greater than the single component, and their expression pattern is similar to that of Baicalin.

Key Words Baicalin; Jasminoidin; Combination of ingredients; Cerebral ischemia; Genechip; Principal component analysis (PCA); Cell adhesion

中药复方作用机制研究逐渐得到世界医学界的研究兴趣^[1]。运用快速、并行分析的基因芯片技术及生物信息学方法可以一定程度上揭示中药干预脑缺血的机制^[2-3], 其高维的基因并行表达数据采用降维分析^[4], 有助于比较中药配伍的复杂作用机制, 而神经元间的

黏附、前炎症过程与认知功能关系密切^[5]。故本研究以清开灵组分配伍为研究对象, 定制 Oligo 高通量基因芯片获取差异基因表达谱, 运用主成分分析(PCA)比较清开灵主要组分黄芩苷、栀子苷及其配伍干预脑缺血损伤的基因表达模式。

基金项目: 国家“十一五”科技支撑计划资助项目(2006BAI08B04—06); 中国中医科学院自主选题项目(Z02154)

通讯作者: 王忠, zhonw@vip.sina.com, (010)64014411 转 3308

1 材料与方法

1.1 动物

12 只 4 周龄昆明种清洁级小鼠购自北京大学实

验动物中心,动物许可证号:SCXK(京)2006-0002,饲养至 6 周。

1.2 药品、试剂与仪器

黄芩苷、栀子苷标准品购自中国生物制品检定所(批号:0824-200102,0740-200109);尼莫地平片由石家庄市华龙药业股份有限公司生产(批号:20041017)。总 RNA 提取试剂盒(美国 Promega 公司,批号:900146),焦碳酸二乙脂(DEPC,BBI 公司,批号:DB0154),琼脂糖、无水乙醇、DNA Mark (DL2000,TaKaRa 公司)。Axon Genepix 4000B 基因芯片扫描仪(美国 Axon 公司生产,批号:Cx03946)。

1.3 动物造模及分组

按随机数字表法分为黄芩苷组(BA 组)、栀子苷组(JA 组)、栀子苷+黄芩苷组(BJ 组)、尼莫地平组(NM 组)各 3 只。结合 Zealanga 等^[6]和 Himori^[7]描述的方法,建立全脑缺血再灌注模型。术后观察左侧上肢屈曲、行走左转或左侧跌倒者为手术成功。2h 后,BA 组、JA 组、BJ 组分别尾静脉注射(40mg/kg)黄芩苷(5mg/ml)、栀子苷(5mg/ml)、黄芩苷+栀子苷(1:1 配比,5mg/ml),NM 组按照 10mg/kg 给予尼莫地平溶液(5mg/ml)。

1.4 RNA 抽提与芯片制备

按一步法^[8]分别抽提脑组织的总 RNA,紫外分析及电泳检测显示获得了高纯度的总 RNA。纯化后,将每组鼠的脑组织总 RNA 等量混合,用于芯片实验。芯片(博奥生物有限公司)包含 16463 个小鼠脑 Oligo 片段,每组重复实验 3 次。将制备的靶基因溶解于 3×SSC 溶液中,用 Cartesian Pixsys 7500 点样仪点样于硅烷化玻片。玻片经水合(2h)、室温干燥(0.5h)、UV 交连(能量值设定 65mj/cm),再分别用 0.2%SDS、水及 0.2%硼氢化钠溶液处理晾干备用。

1.5 探针制备

参照 Schena 的方法^[9]逆转录标记 RNA,于 50μl 逆转录反应体系中加入 60μg RNA,用 Cy3 dUTP 标记 NM 组 RNA,Cy5 dUTP 标记鼠脑组织 RNA,利用 Superscript II 逆转录酶合成探针,乙醇沉淀后将 Cy3d UTP 和 Cy5d UTP 标记的探针混合溶解在 20μl 5×SSC+0.2%SDS 杂交液中,分别进行实验。

1.6 杂交及洗涤

将芯片和杂交探针分别于 95℃变性后置于杂交舱内,加入混合探针。恒温杂交箱内 42℃杂交 18h。按顺序用洗涤液 92×SSC+0.12%Tween-20、0.12×SSC+0.1%×SSC 洗涤 10min,室温晾干备扫描。

1.7 荧光扫描和数据处理

用 Axon Genepix 4000B 扫描仪扫描芯片,链接

Genebank 输出 Unigene 号和功能描述。所得芯片数据均进行整体校正^[10],整体校正后的数据采用容许区间法计算各信号 Ratio 值,确定差异表达阈值。所有数据经过筛选,筛选由 Genespring 分析软件自动完成。Ratio(Cy3/Cy5)的自然对数的绝对值大于 2 或者小于 0.5 的基因分别代表上调或下调的差异表达基因^[11]。

1.8 细胞黏附相关基因的选取

依据基因本体论(GO)分类原则,使用美国斯坦福大学在线资源(<http://genome-www5.stanford.edu/cgi-bin/SMD>),对差异表达基因进行功能分类和定义^[12],提取细胞黏附相关基因。

1.9 统计学方法

主成分分析采用 SAS 9.0 统计软件分析。原始数据为 n×m 阶矩阵。采用负对数化处理,将上调和下调的基因表达比值标准化,未差异表达的为 0^[13]。分组采用样本相关系数矩阵来计算其特征值和特征向量、主成分的贡献率。具体算法见参考文献^[14]。

2 结果

2.1 基因功能分类与细胞黏附相关基因的选取

各组差异表达基因的功能分类结果显示细胞黏附相关基因共 32 个,其中整合素家族基因为 13 个(占 40.62%)。

2.2 主成分个数的选取

经过构建相关系数矩阵,具有相关系数为 1 的基因,属于黏附功能的差异表达基因的主成分计算结果见表 1。4 组的前 2 个主成分的累积贡献度为 100%,表明其所携带的信息概括了 32 个基因的全部信息。

表 1 各组细胞黏附调节差异表达基因主成分特征值表

组别	主成分	特征值	贡献度(%)	累计贡献度(%)
BA 组	Z ₁	16.73	59.76	59.76
	Z ₂	11.26	40.24	100.00
JA 组	Z ₁	15.30	63.76	63.76
	Z ₂	8.69	36.24	100.00
BJ 组	Z ₁	14.83	67.43	67.43
	Z ₂	7.16	32.57	100.00
NM 组	Z ₁	17.24	61.57	61.57
	Z ₂	10.75	38.43	100.00

注:Z_i表示第 i 个主成分

2.3 各组主成分主要贡献基因的比较

表 2 显示了各组主成分的贡献度,其中第一个主成分中 BA 组的主要贡献基因是 Celsr2、整合素 α3(Itga3)、层粘连蛋白 α1(Lama1)、信号转导因子和转录激活因子 5(Stat5A)、转酮酶(Tke)、肿瘤坏死因子超家族成员 5 引导蛋白 1(Tnfsf5)等(其载荷系数大于 0.2,下同),JA 组的主要贡献基因是 Lama1、血小板内皮细胞附着分子(Pecam)、整合素 α4(Itga4)、整合素 α2b(It

表 2 各组细胞黏附相关基因的主成分特征向量表

基因	中文名称	BA 组		JA 组		BJ 组		NM 组	
		Z ₁	Z ₂	Z ₁	Z ₂	Z ₁	Z ₂	Z ₁	Z ₂
Arhe	rnd3 鸟苷酸结合蛋白超家族	-0.12	0.26	0.00	0.00	0.00	0.00	0.00	0.00
Cd2	动力蛋白激活蛋白 2(p50)	-0.12	0.26	0.00	0.00	-0.26	0.02	0.17	0.21
Celsr2	钙粘着蛋白	0.24	0.05	0.00	0.00	0.00	0.00	0.00	0.00
Ctgf	结缔组织生长因子	-0.12	0.26	0.00	0.00	0.12	-0.33	-0.03	0.30
Ddr2	盘蛋白域受体家族成员 2	0.00	0.00	0.00	0.00	0.15	0.30	0.06	-0.30
Dll1	δ 样 1(果蝇)	-0.12	0.26	-0.22	-0.18	0.00	0.00	-0.23	0.10
Emr1	含表皮生长因子样粘蛋白	0.00	0.00	-0.25	0.07	0.00	0.00	0.00	0.00
Itga2	整合素 α2	0.23	0.10	0.09	-0.31	0.25	0.08	-0.23	-0.10
Itga2b	整合素 α2b	0.23	-0.11	0.25	-0.07	0.14	0.31	-0.17	-0.21
Itga3	整合素 α3	0.24	0.05	0.17	0.26	0.20	-0.24	-0.12	0.26
Itga4	整合素 α4	0.00	0.00	0.25	-0.07	0.26	-0.02	0.17	0.21
Itga5	整合素 α5	-0.24	0.01	0.25	-0.07	0.00	0.00	0.17	0.21
Itga7	整合素 α7	-0.08	0.28	-0.02	-0.34	0.16	0.29	-0.17	-0.21
Itgae	整合素 αE	-0.12	0.26	0.25	-0.04	0.11	0.34	-0.17	-0.21
Itgal	整合素 αL(p180)	0.00	0.00	0.00	0.00	0.00	0.00	-0.17	-0.21
Itgam	整合素 αM	0.23	0.10	-0.14	0.28	0.26	-0.02	0.15	-0.24
Itgb3	整合素 β3	-0.22	-0.13	-0.25	0.07	0.11	-0.34	0.12	-0.26
Itgb5	整合素 β5	0.13	0.25	0.00	0.00	0.00	0.00	0.00	0.00
Itgb6	整合素 β6	0.13	0.25	-0.25	0.07	0.22	0.19	-0.06	0.30
Itgb7	整合素 β7	-0.12	0.26	-0.25	0.07	-0.24	0.13	0.24	0.02
Lama1	层粘连蛋白 α1	0.24	-0.01	0.26	0.02	0.00	0.00	0.24	0.01
Mtss1	转移抑制因子 1	0.17	0.22	0.20	0.20	0.00	0.00	-0.20	0.17
Nmb	神经调节肽 B	-0.13	-0.25	0.17	0.26	0.00	0.00	-0.23	0.08
Nphp4	肾结核 4	-0.22	0.14	-0.17	-0.26	-0.26	0.02	-0.23	0.08
Nphs1	肾变性蛋白 1	-0.10	0.27	-0.17	-0.26	-0.26	0.02	0.24	0.01
Pecam	血小板内皮细胞附着分子	-0.24	0.01	0.26	0.00	-0.25	0.10	-0.23	0.08
Pkp1	血小板亲和蛋白 1	-0.24	0.04	0.08	-0.32	0.26	-0.02	0.23	-0.08
Pxn	桩蛋白	0.16	0.23	0.17	0.26	-0.12	0.33	0.24	0.03
Stat5a	信号转导因子和转录激活因子 5A	0.24	-0.01	-0.17	0.26	-0.26	0.02	0.17	0.21
Stat5b	信号转导因子和转录激活因子 5B	0.00	0.00	0.00	0.00	0.00	0.00	0.23	-0.08
Tkt	转酮酶	0.24	0.00	0.00	0.00	0.00	0.00	0.23	-0.08
Tnfsf5	肿瘤坏死因子超家族成员 5 引导蛋白 1	0.24	0.01	-0.08	0.32	0.26	0.05	0.00	0.00

ga2b)、整合素 αE(Itgae)、整合素 α5(Itga5)、转移抑制因子 1(Mtss1),其基因表达中除了整合素家族外,还较强地抑制了血小板内皮细胞黏附分子(Pecam),该分子参与调节单核细胞与内皮细胞的黏附以及随后的经内皮迁移;BJ 组中的主要贡献基因为 Tnfsf5、Itgam、Itga4、Pkp1、Itga2、Itgb6、Itga3,包含 5 个整合素家族基因,其中 BJ 组和 BA 组均共同调控了 Tnfsf5、Itga3、Itgam、Itga2 等 5 个整合素家族基因。

表 2 各变量对应于两个主成分的载荷值,载荷值反映的是主成分与变量的相关系数,可以据此写出主成分载荷表达式。如 BA 组: $Z_{1A} = 0.24Celsr2 + 0.24Itga3 + 0.24Tnfsf5 + 0.24Tkt + 0.24Lama1 + 0.24Stat5a + 0.23Itga2 + 0.23Itgam + 0.23Itga2b - 0.22Nphp4 - 0.22Itgb3 - 0.24Pkp1 - 0.24Itga5 - 0.24Pecam$ 。

2.4 各组对细胞黏附相关基因表达影响总得分比较

根据各主成分的贡献率,可以分别各组的基因的得分模型: $Z_{iBA} = 0.5976Z_1 + 0.4024Z_2$, $Z_{iJA} = 0.6376Z_1 + 0.3624Z_2$, $Z_{iBJ} = 0.6743Z_1 + 0.3257Z_2$, $Z_{iNM} = 0.6157Z_1 + 0.3843Z_2$ (检测次数 $i=1,2,3$)将各

组的原始数据代入上述公式,可以得到各组 3 次检测的综合得分(见表 3),4 组对细胞黏附相关基因的改变显示,BJ 组总得分大于单一组分组,NM 组对细胞黏附相关基因干预的总得分高于 BA 组和 JA 组。提示 3 个中药组均能有效干预细胞黏附相关基因,BJ 组分配伍的基因表达模式和黄芩苷相近。

表 3 各组黏附相关差异表达基因主成分综合得分

组别	检测编号	Z ₁	Z ₂	Z	总分
BA 组	1	-6.60	-4.97	-6.01	4.38
	2	-5.69	20.37	3.75	
	3	27.09	-5.19	15.39	
JA 组	1	30.91	-4.83	19.27	0.77
	2	5.07	5.07	5.07	
	3	-15.68	-35.15	-22.02	
BJ 组	1	76.00	2.66	46.49	12.62
	2	-10.18	-10.22	-10.19	
	3	-3.89	9.66	1.56	
NM 组	1	-9.01	-9.88	-9.34	6.64
	2	63.73	-32.56	26.73	
	3	-12.07	25.97	2.55	

3 讨论

用于本研究的清开灵组分(黄芩苷、栀子苷)是在 973 课题组建立的中药标准组分提取分离平台、指纹图谱质量评价体系和多组分配伍优化筛选模式下筛选出来的^[15]。探索组分配伍的加和效应,有助于揭示方剂配伍的复杂机制。故本研究本着“复杂问题由简单入手”的思路,从基因表达谱层面探索两个组分的加和关系。

基因芯片数据为高维的海量数据,对其生物信息的挖掘分析发展很快^[4,16]。基于多元分析思想的主成分分析旨在原始数据信息丢失最小的情况下,对高维变量进行降维处理。选取的主成分一般为贡献率最大(方差最大)的若干主成分,这些主成分被认为是原始数据中的真实信号,需要对其进行进一步分析才能解释其深刻内涵,而其余主成分则被认为是噪声信号而丢弃。因此 PCA 成为解决高通量生物信息分析的有效手段^[4,16]。

对两个组分及其配伍多基因的降维分析结果显示, BJ 组总得分大于单一组分。这一结果提示,黄芩苷和栀子苷配伍后对细胞黏附相关基因干预作用优于单一组分,起到加和效应的主要贡献基因主要集中在整合素家族基因,如 Itgam、Itga4、Pkp1、Itga2、Itgb6、Itga3 等。这些基因对于调控细胞的黏附与趋化及炎症细胞的聚集等过程有关。3 个中药组均能有效干预细胞黏附相关基因,组分配伍的基因表达模式和黄芩苷相近。BJ 组和 BA 组均共同调控了 Tnfsf5、Itga3、Itgam、Itga2 等在细胞与细胞外基质(ECM)间的黏附及信息传递中发挥重要作用的基因,反映了配伍和其某一组分具有共同的基因表达基础。经配伍后 Pkp1(血小板亲和蛋白 1)得以上调,起到稳定纤维分子骨架的形成,改善细胞黏附后炎症状态的作用,体现了“增效”效应。单一组分黄芩苷下调了重要的 Celsr2,抑制了神经元分化^[17],而经过和配伍栀子苷后,这种抑制消失了,提示了两组分配伍的“减毒”效应。这两个单一组分配伍组整体影响大于单一组分与本课题组前期的药效学(行为学、形态学)研究结果^[18]是一致的, NM 组与单一组分栀子苷的比较结果却与既往不一致,这可能与本研究限于数据挖掘方法学,仅对细胞黏附相关基因进行分析有关,更全面的证据仍需进一步研究。

参考文献

[1] Li XM, Huang CK, Zhang TF, et al. The chinese herbal medicine formula MSSM-002 suppresses allergic airway hyperreactivity and modulates TH1/TH2 responses in a murine model of allergic asthma[J]. J Allergy Clin Immunol, 2000, 106(4):660-668.

[2] Sharp FR, Lu A, Tang Y, et al. Multiple molecular penumbras after focal cerebral ischemia[J]. J Cereb Blood Flow Metab, 2000, 20(7): 1011-1032.

[3] 荆志伟, 王忠, 杨俊红, 等. 基因芯片技术与缺血性脑损伤的基础研究[J]. 中风与神经疾病杂志, 2008, 25(3): 253-256.

[4] 黄德双. 基因表达谱数据挖掘方法研究[M]. 北京: 科学出版社, 2009: 65.

[5] Prabhakar S, Noonan JP, Paabo S, et al. Accelerated evolution of conserved noncoding sequences in humans[J]. Science, 2006, 314(5800): 786.

[6] Longa EZ, Weinstein PR, Carlson S, et al. Reversible middle cerebral artery occlusion without craniectomy in rats[J]. Stroke, 1989, 20(1): 84-91.

[7] Himori N, Watanabe H, Akaike N, et al. Cerebral ischemia model with conscious mice. Involvement of NMDA receptor activation and derangement of learning and memory ability[J]. J Pharmacol Methods, 1990, 23(4): 311-327.

[8] Chomczynski P, Sacchi N. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol chloroform extraction[J]. Anal Biochem, 1987, 162(1): 156-159.

[9] Kim HS, Lee SJ, Kim DS, et al. Effects of brain-derived neurotrophic factor and neurotrophin-3 on expression of mRNAs encoding c-Fos, neuropeptides and glutamic acid decarboxylase in cultured spinal neurons[J]. Neuroreport, 2000, 11(17): 3873-3876.

[10] Yang P, Otto CM, Sheehan FH. The effect of normalization in reducing variability in regional wall thickening[J]. J Am Soc Echocardiogr, 1997, 10(3): 197-204.

[11] Jiang CH, Tsien JZ, Schultz PG, et al. The effects of aging on gene expression in the hypothalamus and cortex of mice[J]. Proc Natl Acad Sci USA, 2001, 98(4): 1930-1934.

[12] Siegel GJ, Agranoff BW, Albers RW, et al. Basic neurochemistry. Molecular, cellular and medical aspects[M]. Philadelphia: Lippincott-Raven, 1999: 102-150.

[13] Yeung KY, Ruzzo WL. Principal component analysis for clustering gene expression data[J]. Bioinformatics, 2001, 17(9): 763-774.

[14] 孙振球, 徐勇勇. 医学统计学[M]. 北京: 人民卫生出版社, 2002: 321-330.

[15] 张伯礼, 王永炎. 方剂关键科学问题的基础研究——以组分配伍研制现代中药[J]. 中国天然药物, 2005, 3(5): 258-261.

[16] 荆志伟, 王忠. 基因芯片数据分析方法研究进展[J]. 生物技术通讯, 2007, 18(1): 144-148.

[17] 张占军, 李澎, 邱蕾, 等. 神经元钙传感蛋白的研究进展及其在脑缺血中的作用[J]. 中国药理学通报, 2009, 25(1): 12-15.

[18] 张占军, 汪丽娅, 王忠, 等. 黄芩苷、栀子苷及其配伍治疗局灶性脑缺血大鼠药效评价及作用机制研究[J]. 中国中药杂志, 2006, 31(11): 907-909.

(收稿日期: 2009-05-11; 修回日期: 2009-09-19)

[编辑: 崔京艳]