

新疆维吾尔族、哈萨克族、汉族冠心病差异基因筛选及功能路径分析

李霞¹, 王枚¹, 马依彤², 谢翔², 杨毅宁², 马翔², 陈铀², 罗俊一², 刘芬³

(1.新疆医科大学第五附属医院老年病科, 新疆乌鲁木齐市 830011; 2.新疆医科大学第一附属医院心脏中心, 新疆乌鲁木齐市 830054; 3.新疆心血管病重点实验室, 新疆乌鲁木齐市 830011)

[关键词] 基因芯片; 基因表达谱; 冠心病

[摘要] **目的** 筛选新疆维吾尔族、哈萨克族、汉族冠心病相关的差异基因。**方法** 选用病例-对照研究按照 1:1 匹配原则, 选择经冠状动脉造影检查明确诊断为冠心病患者 9 例, 其中包括维吾尔族 3 例, 哈萨克族 3 例, 汉族 3 例; 对照组选择同期住院经冠状动脉造影检查冠状动脉血管完全正常者 9 例, 两组民族、性别、年龄均匹配; 利用人类全基因组表达谱芯片筛查与冠心病相关的差异基因。**结果** 筛查出与冠心病相关的差异基因 33 条, 与汉族冠心病相关的差异基因 26 条, 与维吾尔族冠心病相关的差异基因 2 条, 而哈萨克族未筛选出差异基因表达; 经 GO 和 Pathway 分析后, 差异基因多与免疫反应有关。**结论** 基因芯片筛查的结果表明, 对生物学进程有影响的差异基因主要与免疫反应有关。

[中图分类号] Q343

[文献标识码] A

The Screening and the Functional Pathway Analysis of Differential Genes in Patients with Coronary Heart Disease in Uygur, Kazakh, and Han Population of Xinjiang

LI Xia¹, WANG Mei¹, MA Yi-Tong², XIE Xiang², YANG Yi-Ning², MA Xiang², CHEN You², LUO Jun-Yi², and LIU Fen³

(1. Department of Geriatrics, the Fifth Affiliated Hospital of Xinjiang Medical University, Urumqi, Xinjiang 830011, China; 2. Heart Center of the First Affiliated Hospital of Xinjiang Medical University, Urumqi, Xinjiang 830054, China; 3. Xinjiang Key Laboratory of Cardiovascular Disease Research, Urumqi, Xinjiang 830011, China)

[KEY WORDS] Gene Microarray; Gene Expression Profile; Coronary Heart Disease

[ABSTRACT] **Aim** To screen differential genes associated with coronary heart disease in Uygur, Kazakh, and Han population of Xinjiang using gene microarray expression profiles. **Methods** A case-control study including 9 cases (3 cases in Uygurs, 3 cases in Kazakhs, and 3 cases in Hans) with coronary heart disease confirmed by coronary angiography and 9 gender-, age-, and ethnic-matched normal controls was designed for the present study. Human whole genome expression microarray was used to screen differential genes associated with coronary heart disease. **Results** 33 genes associated with coronary heart disease were differentially expressed, 26 genes associated with coronary heart disease in Han population were differentially expressed, 2 genes associated with coronary heart disease in Uygur population were differentially expressed, while the differentially expressed genes were not screened in Kazakh population. Through GO and Pathway analysis, the differentially expressed genes were related to immune response. **Conclusion** The results of gene microarray screening showed that the differentially expressed genes were mainly related to immune response.

新疆是一个多民族聚居区, 随着本地区经济水平的提高, 冠心病发病率和死亡率不断增加, 严重

威胁着该地区各族人民群众的身体健康。冠心病是由多种基因、多种环境因素以及基因与环境因素

[收稿日期] 2016-03-09

[修回日期] 2016-09-02

[基金项目] 新疆维吾尔自治区自然科学基金医学联合基金资助项目(2015211C166)

[作者简介] 李霞, 博士, 副主任医师, 研究方向为冠心病的遗传学, E-mail 为 250456230@qq.com。王枚, 硕士, 主治医师, 研究方向为老年心血管病诊治。通讯作者马依彤, 博士, 主任医师, 教授, 博士研究生导师, 研究方向心血管疾病的介入治疗, E-mail 为 myt-xj@163.com。

相互作用所致的复杂性疾病,且在不同人群之间存在一定的差异性。本研究利用人类全基因组表达谱芯片筛选新疆地区维吾尔族、哈萨克族、汉族冠心病差异基因,从分子水平进一步认识冠心病的发生、发展机制。

1 对象和方法

1.1 研究对象

按照病例对照研究中 1:1 个体匹配原则,选取冠心病患者共 9 例,其中维吾尔族 3 例(男 2 例,女 1 例,年龄 57.0 ± 1.53 岁),哈萨克族 3 例(男 2 例,女 1 例,年龄 56.67 ± 1.20 岁),汉族 3 例(男 2 例,女 1 例,年龄 56.8 ± 1.86 岁);对照组 9 例,其中维吾尔族 3 例(男 2 例,女 1 例,年龄 55.67 ± 1.45 岁),哈萨克族 3 例(男 2 例,女 1 例,年龄 57.02 ± 0.58 岁),汉族 3 例(男 2 例,女 1 例,年龄 57.0 ± 1.15 岁),所有研究对象均选自 2013 年 2 月至 4 月在新疆医科大学第一附属医院心脏中心的住院患者,签署知情同意后纳入研究。冠心病组按照世界卫生组织制定的缺血性心脏病的诊断标准^[1],选择经冠状动脉造影证实有一支或多支血管狭窄程度 $\geq 50\%$ 的冠心病患者。对照组入选标准:住院行冠状动脉造影检查排除冠心病且年龄、性别、民族与冠心病组相匹配,冠心病组与对照组中体质指数、吸烟、饮酒、冠心病相关家族史、高血压病史、高脂血症病史、糖尿病病史等方面基本相似,具有可比性。剔除标准^[1]:冠心病组合并有下列疾病之一者均剔除:先天性心脏病、主动脉夹层、多脏器功能衰竭、风湿性心脏病、感染性疾病、慢性阻塞性肺病、肺动脉栓塞、免疫系统疾病、肿瘤以及具有精神障碍不能配合者;对照组:先天性心脏病、主动脉夹层、多脏器功能衰竭、风湿性心脏病、感染性疾病、慢性阻塞性肺病、肺动脉栓塞、免疫系统疾病、肿瘤及有精神障碍者或经多普勒检查证实有颈动脉斑块或狭窄者。两组入选对象均无血缘关系。

1.2 血样采集

本次研究经申请并且通过了新疆医科大学第一附属医院伦理委员会批准,入选前签署知情同意书。在采血前,空腹 12 h 以上,用 EDTA-K3 抗凝真空管和普通生化管于肘静脉处采集外周静脉血各 5 mL 备用。

1.3 外周血单个核细胞提取

利用 Ficoll 密度梯度离心法分离外周血单个核细胞(peripheral blood mononuclear cell, PBMC)。

EDTA-K3 抗凝真空管采用无菌操作技术抽取新鲜外周静脉血 5 mL,将外周血与生理盐水按 1:1 稀释;取 15 mL 无菌离心管,加入 5 mL 淋巴细胞分离液(血:淋巴细胞分离液为 1:1),距分离液上 1 cm 处,沿管壁加入稀释后的血液,注意血液与淋巴细胞分离液分层要清楚;将装有淋巴细胞分离液与血液的无菌离心管放入水平离心机中,400 g,20℃ 水平离心 20 min,此时离心管中形成血浆层、PBMC 层、淋巴细胞分离液层、粒细胞层、红细胞层 5 层;用一次性直头滴管吸取 PBMC 层,移至另一 15 mL 无菌离心管中;在放置 PBMC 层的离心管中加入生理盐水 4~5 mL,充分震荡混匀后离心 20 min(400 g, 20℃),将离心后的液体倾去,留取沉淀,再次倒入生理盐水 2 mL,充分震荡混匀后用移液器移至 2.0 mL EP 管中,离心 2 min(13000 r/min, 20℃),所得沉淀即为 PBMC;在 PBMC 中加入 1 mL Trizol 液并用移液器反复吹打混匀,备用。

1.4 总 RNA 提取

总 RNA 的提取采用 Trizol 一步提取法:将加有 Trizol 液的 PBMC 匀浆后转入无 RNA 酶的离心管中,加入 0.5 mL 异丙醇,在室温下放置 10 min,然后在 4℃、以 12000 g 离心 10 min,弃上清液,加入 1 mL 75% 乙醇,震荡后在 4℃、以 7500 g 离心 5 min,弃上清,真空干燥 5~10 min,加入 5% SDS 溶液 50 μ L,吹打混匀后放入 60℃ 恒温箱中孵育 15 min, -70℃ 冰箱保存,备用。

1.5 芯片选择

芯片选用美国昂飞公司人类全基因组表达谱芯片(Affymetrix GeneChip Human Genome U133 Plus 2.0 Array),该芯片包含 54676 个探针组和 47000 个转录本,包括 38500 个人类已知基因,芯片的标记、杂交及扫描由依托博奥生物有限公司实验平台完成。

1.6 数据处理与分析

用 RMA 法^[2]预处理数据后进行差异表达基因分析:有 3 个及以上生物学重复的实验设计,利用 SAM(significance analysis of microarray) R 程序包分析差异基因,差异基因的筛选标准是 q 值 ≤ 0.05 且差异倍数 ≥ 2 或 ≤ 0.5 ;Cluster 软件进行分层聚类分析和结果展示;博奥生物开发的分子功能注释系统(MAS3.0)(<http://bioinfo.capitalbio.com/mas3/>)进行基因本体论(gene ontology, GO)和 Pathway 分析。通过 GO 和 KEGG、GenMAPP、Biocarta Pathway 数据库计算 P 值和 q 值,当 P 值和 q 值均 < 0.05 说明基

因在功能和通路上具有富集性,对生物学功能影响明显。

2 结 果

2.1 RNA 质检

用紫外分光光度计检测 RNA 样品纯度, $A_{260/280} \geq 1.70$, RNA 总量 $\geq 1 \mu\text{g}$; 经甲醛变性胶电泳检测, RNA 样品电泳条带清晰, 28 S : 18 S rRNA 条带亮度大于或接近 2 : 1。说明 RNA 的样品纯度、总量及完整性均符合表达谱芯片实验要求(图 1)。

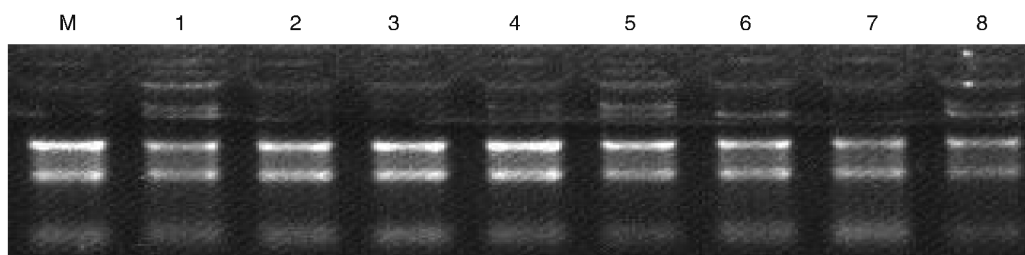


图 1. RNA 样品质检凝胶电泳结果 M 泳道为 Hela 细胞质控 RNA, 1~8 为样本序号。

Figure 1. Gel electrophoresis results of RNA samples



图 2. 冠心病组全基因组芯片扫描结果截图(阵列名称)

Figure 2. Whole genome scan results screenshot (array name) in coronary heart disease group

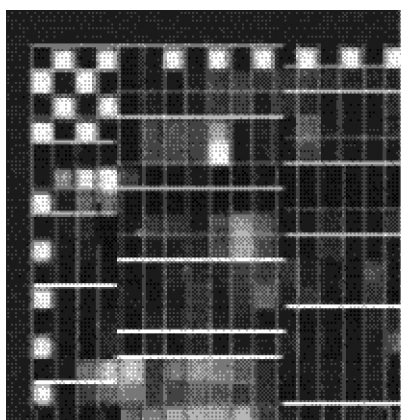


图 3. B2 对照组探针与检测芯片杂交扫描结果(突出阵列边缘与四角)

Figure 3. Probe and detection chip hybridization scan results in control group (B2 Outstanding array edge and the four corner)

2.2 芯片质控及扫描情况

B2 Oligo 做为阳性杂交对照, 与 Affymetrix GeneChip Human Genome U133 Plus 2.0 Array 芯片杂交后可以界定探针的区域, 在扫描图上 B2 Oligo 的杂交通过以下方式得到突出: ①边缘呈现信号强度的强弱交替出现; ②每一角呈现棋盘状的信号, 说明 RNA 样品质量可靠; ③U133 Plus2.0 芯片与探针杂交扫描结果显示, 芯片左上部或者中上部呈现阵列的名称“GeneChip HG-U133 Plus 2”, 说明看家基因信号值及 3’/5’ 比值正常, 平均背景值与噪音值较低, 符合要求范围, 芯片质控良好(图 2 和 3)。

2.3 冠心病组与对照组相关差异基因比较

根据筛选差异基因的原则, 冠心病组与对照组共有 33 条差异基因表达, 均为上调基因。为了进一步排除不同民族间本身遗传基因差异对筛选结果的影响, 进一步进行各民族组内筛选, 结果发现汉族冠心病组与汉族对照组相关差异基因共有 26 条, 为上调基因; 维吾尔族冠心病组与维吾尔族对照组相关差异基因只有 2 条, 为下调基因; 而哈萨克族冠心病组与哈萨克族对照组未发现差异基因。

2.4 汉族、维吾尔族、哈萨克族间相关差异基因比较

维吾尔族冠心病组与汉族冠心病组有 43 条差异基因, 其中上调基因 41 条, 下调基因 2 条; 维吾尔族冠心病组与哈萨克族冠心病组有 1 条下调差异基

因;哈萨克族冠心病组与汉族冠心病组有 1 条下调差异基因;维吾尔族对照组与汉族对照组有 3 条上调差异基因,维吾尔族对照组与哈萨克族对照组有 1 条下调差异基因,哈萨克族对照组与汉族对照组也有 1 条下调差异基因。

2.5 差异基因 GO 和 Pathway 分析

冠心病组与对照组间相关的 33 条差异基因进行 GO 分析,明显影响生物学进程的差异基因中有 14 条与免疫反应有关 ($P = 5.3\text{e-}34, q = 5.3\text{e-}34$),其余涉及细胞凋亡、细胞增殖及 RNA 聚合酶 II 启动子转录负调控等(表 1)。通过 KEGG、GenMAPP、Bio-carta 数据库 Pathway 分析,发现差异基因中肿瘤坏死因子受体超家族成员 17(tumor necrosis factor receptor superfamily member 17, TNFRSF17)基因在细胞因子和细胞因子受体相互作用的通路中有富集性,参与体液免疫和细胞增殖过程。汉族冠心病组与汉族对照组相关的 26 条差异基因进行 GO 分析,有 8 条与免疫反应有关的差异基因影响生物学进程,其余涉及细胞凋亡、细胞增殖、内皮细胞迁移等(表 2)。经上述数据库 Pathway 分析,发现差异基因中 TNFRSF17 基因在细胞因子和细胞因子受体相互作用的通路中有富集性,IGHA1、IGHG1 基因参与炎症反应通路,对生物学功能影响明显 ($P < 0.05$ 和 $q < 0.05$)。维吾尔族冠心病组与维吾尔族对照组间 2 条差异基因,进行 GO 和 Pathway 分析后未发现对生物学功能产生明显影响。维吾尔族冠心病组与汉族冠心病组间的 43 条差异基因进行 GO 分析,同样与免疫反应有关的差异基因最多,共 16 条。维吾尔族冠心病组与哈萨克族冠心病组间差异基因及哈萨克族冠心病组与汉族冠心病组间差异基因进行 GO 和 Pathway 分析后未发现对生物学功能产生明显影响。同样,维吾尔族对照组与汉族对照组间、维吾尔族对照组与哈萨克族对照组间、哈萨克族对照组与汉族对照组间的差异基因进行 GO 和 Pathway 分析后均未发现对生物学进程产生明显影响。

3 讨 论

目前在全球范围内,冠心病是引起死亡和伤残调整生命年损失的最主要原因,每年大概造成 7 百万人口死亡和 1.29 亿伤残调整生命年损失^[3-4]。尤其在低中等收入地区,自上世纪 90 年代始,随着城市化、机械化发展,静坐式工种增加,加速了冠心病

表 1. 冠心病组与对照组明显改变生物学进程的差异基因
Table 1. Differentially expressed genes that significantly changed biological processes between coronary heart disease group and control group

GO 生物学过程	数量	P 值	q 值	基因名称
免疫反应	14	5.68E-34	5.68E-34	TNFRSF17
				IGHG1
				IGHD
				IGHA1
				KV123_HUMAN
				IGLC3
				IGJ
				IGKV1-5
				IGKV4-1
				IGLC3
				KV309_HUMAN
				KV308_HUMAN
				KV303_HUMAN
				KV310_HUMAN
细胞凋亡	1	0.043883	0.004708072	PACAP
细胞增殖	1	0.044524	0.004708072	TNFRSF17
信号转导	1	0.171049	0.013683883	TNFRSF17
RNA 聚合酶 II 启动子转录负调控	1	0.009618	0.002404594	FOSB

表 2. 汉族冠心病组与汉族对照组明显改变生物学进程的差异基因
Table 2. Differentially expressed genes that significantly changed biological processes between coronary heart disease group and control group in Han population

GO 生物学过程	数量	P 值	q 值	基因名称
免疫反应	8	3.10E-16	3.10E-16	TNFRSF17
				IGHA1
				IGHG1
				HV303_HUMAN
				IGLC3
				IGJ
				KV123_HUMAN
				IGHM
细胞凋亡	1	0.029	0.003	PACAP
细胞增殖	1	0.029	0.003	TNFRSF17
B 淋巴细胞酪氨酸激酶	1	0.015	0.0023	BLK

死亡率和伤残调整生命年的增加^[5-6]。正如前所述,冠心病是一种多因素疾病,其发病一方面受环境因素的影响,另一方面与遗传因素也有很大关系。不同地区、不同种族及同一地区、不同种族间冠心病的患病率存在很大差异。本研究利用人类全基因组表达谱芯片技术筛选新疆地区维吾尔族、哈萨克族、汉族冠心病相关的差异基因。从筛选的

结果来看,与冠心病有关的差异基因有 33 条,通过数据库 Pathway 分析后发现与生物学进程有关的基因主要涉及参与了免疫反应;在三个不同民族中进行冠心病组与对照组差异基因比较后发现,汉族冠心病组与汉族对照组相关差异基因共有 26 条,均为上调基因,同样涉及免疫反应的最多,有 8 条基因,GO 和 Pathway 分析后发现对生物学进程有影响的基因仍然是 TNFRSF17、IGHA1、IGHG1。维吾尔族冠心病组和对照组的差异基因经分析后对生物学进程无意义。由此看出,本研究筛选出的冠心病差异基因,主要是汉族冠心病组与汉族对照组间的差异基因,而且对生物学进程有影响的差异基因主要与免疫反应有关。

近年来随着分子生物医学的发展,人们逐渐认识到遗传因素在冠心病发病机制中的重要作用,大约 40%~60% 冠心病易感性和遗传有关,基因检测对认识冠心病危险因素和治疗起很大作用。基因芯片是一种平面的基质载体,它上面有规则地、特异地吸附着基因或基因的产物。在近十多年来,微阵列技术得到了快速的发展,科学家们能够明确认定特定生理或病理状态下的全基因组表达谱,而这些是传统生物学技术不能完成的^[7]。

基因芯片表达谱分析是一种假设性研究,具有高通量、高灵敏度检测特点,但是在芯片杂交过程中容易产生假阳性。为了提高芯片结果的可靠性,在实验过程中要严格掌握样本入选和排除标准。RNA 提取、探针标记、cDNA 逆转录、芯片杂交调节及图像分析等任何环节的微小差别都有可能造成实验结果的不同。与传统分子生物学技术相比,基因芯片技术能比较好地观察基因组的整体性变化,全面反映和预测单个或少数基因相应的生物学机制,为后续研究确定方向。

动脉粥样硬化是心血管疾病重要的病理生理过程,主要累及大中型动脉。脂质异常堆积并伴有炎症、免疫因素参与是动脉粥样硬化的特征性病理改变^[8]。本研究中筛选出与冠心病有关的差异基因如 TNFRSF17 基因,该基因主要在成熟 B 淋巴细胞中表达,与 B 细胞发育和自身免疫反应有关,与肿瘤坏死因子家族成员 13B 结合,激活 NF- κ B 和 MAPK8/JNK 通路活性,再与肿瘤坏死因子受体家族成员结合,参与细胞存活和增殖信号转导通路。

免疫球蛋白 IgA1 (IGA1) 基因、IgG1 (IGHG1) 基因、IgD (IGHD) 基因等一批与冠心病相关的基因,使我们有理由认为筛选出的差异基因多与免疫反应有关。为了确定单个基因与疾病的关系,必须对筛出的目标基因进行验证,通过验证后方可确定表达谱研究的可靠性,我们将在后续的研究中进一步对目标基因进行验证。

[参考文献]

- [1] Bahri R, Esteban E, Moral P, et al. Apolipoprotein gene polymorphisms and plasma levels in healthy Tunisians and patients with coronary artery disease [J]. *Lipids Health Dis*, 2008, 17 (7): 46.
- [2] Irizarry RA, Hobbs B, Collin F, et al. Exploration, normalization, and summaries of high density oligonucleotide array probe level data [J]. *Biostatistics*, 2003, 4 (2): 249-264.
- [3] Lozano R, Naghavi M, Foreman K, et al. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010 [J]. *Lancet*, 2012, 380: 2095-128.
- [4] Murray CJ, Vos T, Lozano R, et al. Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010 [J]. *Lancet*, 2012, 380: 2197-223.
- [5] Moran AE, Forouzanfar MH, Roth GA, et al. The global burden of ischemic heart disease in 1990 and 2010: the Global Burden of Disease 2010 study [J]. *Circulation*, 2014, 129: 1493-501.
- [6] Moran AE, Forouzanfar MH, Roth GA, et al. Temporal trends in ischemic heart disease mortality in 21 world regions, 1980 to 2010: the Global Burden of Disease 2010 study [J]. *Circulation*, 2014, 129: 1483-492.
- [7] Montagna P, Cortelli P, Avoni P, et al. Clinical features of fatal familial insomnia: phenotypic variability in relation to a polymorphism at codon 129 of the prion protein gene [J]. *Brain Pathol*, 1998, 8: 515-520.
- [8] 喻思扬,王燕,刘洋,等. 炎性体在动脉粥样硬化等心血管疾病中的作用及机制 [J]. *中国动脉硬化杂志*, 2014, 22 (12): 1281-286.

(此文编辑 文玉珊)