文章编号:1000-5404(2011)04-0331-04

论著

利用全基因组SNP芯片筛查2型糖尿病患者大血管病变易感基因

朱 沂[1],张晓莉[1],张忠辉[2],张 峰[3],张永彪[3],府伟灵[1] (400038 重庆,第三军医大学西南医院:检验科[1],内分泌科[2];101300 北京,中国科学院北京基因组所[3])

【摘要】 **目的** 分析筛查2型糖尿病(T2DM)患者早期大血管病变的易感基因和单核苷酸多态性(single nucleotide polymorphism,SNP)位点。**方法** 利用ILLUMINA人类全基因组SNP芯片(HumanCytoSNP-12 v1.0 DNA Analysis BeadChip Kit),对在经多因素干预下仍出现大血管病变的34例T2DM病例及同样条件下未发大血管病变的52例对照进行SNP扫描分型;通过全基因组关联分析,筛选T2DM早期大血管病变的易感基因和遗传标记。**结果** 通过PLINK软件对芯片结果进行全基因组关联分析(genome-wide association study, GWAS),筛选出在病例组和对照组间有显著差异($P<0.01$)的SNP位点总计452个,其中处于T2DM大血管病变主要代谢通路相关基因内或者附近的SNP位点37个,通过这些SNP位点锁定T2DM患者大血管病变易感基因30个。**结论** T2DM大血管病变受遗传多态性影响,可能与多个基因以及位点相关。

【关键词】 2型糖尿病;大血管病变;单核苷酸多态性;易感基因;全基因组SNP芯片
【中图法分类号】 R394-33;R446.9;R587.104 【文献标志码】 A

Screening macroangiopathy susceptibility genes in type 2 diabetes mellitus patients with whole genome-wide SNP array

Zhu Yi[1], Zhang Xiaoli[1], Zhang Zhonghui[2], Zhang Feng[3], Zhang Yongbiao[3], Fu Weiling[1] ([1]Department of Laboratory Medicine, [2]Department of Endocrinology, Southwest Hospital, Third Military Medical University, Chongqing, 400038; [3]Genomics Institute of Chinese Academy of Sciences, Beijing, 101300, China)

【**Abstract**】 **Objective** To analyze and screen macroangiopathy susceptibility genes and single nucleotide polymorphism (SNP) loci in type 2 diabetes mellitus (T2DM) patients. **Methods** The DNA sample of 34 patients with type 2 diabetes who had been interfered with multiple risk factors of macroangiopathy but still developed to atherosclerosis and 54 other T2DM patients under the same conditions but having no arteriosclerosis were scanned by human whole genome-wide SNP array of ILLUMINA (HumanCytoSNP-12 v1.0 DNA Analysis Bead Chip Kit). The susceptibility genes and genetic markers of macroangiopathy in type 2 diabetes were screened through genome-wide association analysis (GWA). **Results** The results of chip test were studied for genome-wide association by PLINK software. There were 452 SNPs found to be significantly different between case group and control ($P<0.01$). Among them, 37 SNPs were in or near to the genes related the major metabolic pathway of T2DM macroangiopathy. According these 37 SNP loci, 30 genes were identified as macroangiopathy susceptibility genes for T2DM. **Conclusion** Our study shows that genetic polymorphism is associated with macroangiopathy of T2DM, and some multiple genes and SNPs might be involved.

【**Key words**】 type 2 diabetes mellitus; macroangiopathy; single nucleotide polymorphism; susceptibility gene; whole genome-wide SNP array

Supported by the General Program of National Natural Science Foundation of China (30972816). Corresponding author: Fu Weiling, E-mail: weilingfu@yahoo.com

大血管并发症是我国糖尿病患者主要的致残、致死原因。研究表明,传统危险因素的干预措施对某些易感者益处不大,血糖、血压、血脂等控制良好的患者也不能完全阻止大血管病变的发生,提示一些未知的非传统因素如遗传多态在糖尿病大血管病变的发生中可能起着重要的作用[1-2]。本研究通过全基因组SNP芯片对样本DNA进行SNP的扫描分型,对芯片结果进行全基因组关联分析,以期筛选出T2DM大血管病变的易感基因和遗传标记。

1 资料与方法

1.1 研究对象

86例重庆地区汉族、彼此无血缘关系、经过前期干预前瞻性研究同期入组未发生大血管病变的新诊2型糖尿病患者,样本来源于"十一五"国家攻关课题"2型糖尿病及其并发症预

【基金项目】 国家自然科学基金面上项目(30972816)
【通信作者】 府伟灵,E-mail:weilingfu@yahoo.com

警、干预的研究”，在8年前瞻性研究中对患者进行了大血管病变的传统危险因素如体质量、血压、血脂、血糖等的控制以及进行生活行为综合干预，具体的干预模式如下：分为4个不同的干预组，A组：强化血糖、血压控制；B组：强化血糖、血压、血脂控制；C组：强化血糖、血压、血脂控制＋维生素E干预；D组：强化血糖、血压、血脂控制＋复方丹参滴丸干预。对无禁忌证者，4组均同时给予阿司匹林治疗和生活行为综合干预。其中男性44例，女性42例，年龄(58.93±9.01)岁。经过8年随访观察，在实验的终点，采用惠普HD I5000型彩色超声仪，体表10 MHz宽频探头，于舒张期测量颈总动脉、股动脉、髂总动脉内膜中层的厚度(intima media thickness，IMT)，根据IMT把研究对象分为出现大血管病变组(病例组，34例)和未出现大血管病变组(对照组，52例)。分组的标准为：①3处动脉血管IMT≤1.0 mm且无斑块可判定为无大血管病变；②若其中3处动脉血管的任何1支IMT≥1.3 mm或发现斑块可判定为出现早期大血管病变；③对于3处动脉血管任何1支IMT＞1.0 mm但＜1.3 mm的个体，由于目前还无法断定其是否会发展为大血管病变，为避免错分病例的可能，所以暂不纳入研究。

1.2 研究方法

1.2.1 基因组DNA抽提 每个被试抽取静脉血5 ml，EDTA抗凝，天根全血基因组DNA提取试剂盒(离心柱型)提取基因组DNA，－70 ℃保存备用。

1.2.2 全基因组SNP分型 采用ILLUMINA人类全基因组SNP芯片(HumanCytoSNP-12 v1.0 DNA Analysis BeadChip Kit)，对样本DNA进行高通量SNP扫描分型。在进行芯片检测前对DNA样本进行了纯度和浓度检测，不符合芯片检测要求的样本进行了纯化，最终86例样本DNA都达到了检测要求。该芯片每张上有37万个从Hapmap一期和二期数据的基础上挑选的标签SNP(tagSNP)检测位点，最大限度地涵盖人全基因组的全部信息，包含了通过连锁不平衡分析选择在基因上下游10K区域或保守性区域内r2＞0.8的标签SNP位点，其他区域r2＞0.7的标签SNP位点，并包含了约7 300个非同义SNP及约1 200个MHC区域(主要组织相容性复合体)内的标签SNP。Illumina全基因组SNP芯片是光纤微珠芯片，在玻璃基片上包埋直径为3.1 μm无孔无荧光硅珠，并在每颗硅珠上连接100万左右条寡核苷酸探针，是一种高通量低成本的SNP分型检测芯片。

1.2.3 临床指标的测定 测定患者身高、体质量、血压、空腹血糖、糖化血红蛋白(HbA1c)、血脂[血清甘油三酯(TG)、总胆固醇(TC)、低密度脂蛋白胆固醇(LDL-C)、高密度脂蛋白胆固醇(HDL-C)]，计算体质量指数(body mass index，BMI)，并采用惠普HD I5000型彩色超声仪，体表10 MHz宽频探头，于舒张期测量颈总动脉IMT。

1.2.4 数据处理 对样本的遗传背景采用EIG 3.0软件中的smartpca对数据进行主成分分析(PCA分析)，以验证所有的样本是否处于相同的遗传背景。用样本人群数据与HapMap数据库的北京汉族(CHB)、日本(JPT)、欧洲(CEU)及非洲(YRI)人群数据比较，同时在病例组和对照组之间进行比较。具体分组如下：①Asia组：包括JPT、CHB、病例组、对照组。②World组：包括YRI、CEU、JPT、CHB、病例组、对照组。③Case-control组：包括病例组、对照组。

1.3 统计学分析

临床资料各项指标以 $\bar{x} \pm s$ 表示，采用SPSS 12.0软件，两组均值间比较用 t 检验。利用PLINK 1.06软件进行全基因组关联分析和Logistic回归分析。

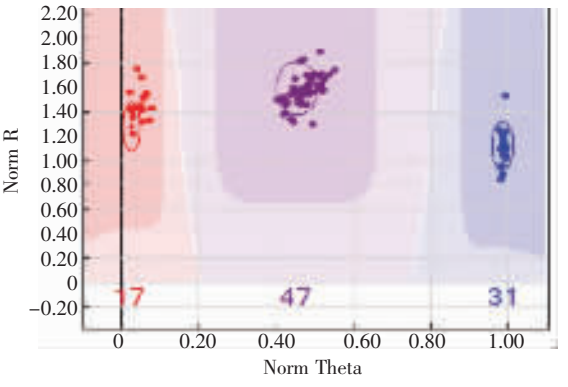
2 结果

2.1 病例组和对照组间相关临床指标比较

病例组和对照组相关临床指标比较结果见表1。除年龄之外，各项指标两组间比较差异无统计学意义(P＞0.05)。

2.2 芯片检测结果

86例样本在所有SNP检测位点的Call rate＞0.99，芯片检测质量良好。芯片SNP分型结果的基因型被定义为“AA”、“AB”、“BB”。其中“AA”和“BB”表示此样本在该位点上为纯合基因型，“AB”则为杂合基因型。以位点rs338556为例，其分型结果见图1。图中每个点代表1个样本的芯片检测信号(总计检测样本为95例，纳入研究86例)。



红色与蓝色的点代表纯合基因型AA和BB；紫色的点代表杂合基因型AB；横轴norm Theta(normlized Theta-value)＝2/Pi×arctan(B/A)，Pi代表圆周率；纵轴norm R(normlized R-value)＝A＋B，A、B表示两种碱基标准化后的荧光信号强度；17、47、31：样本数量

图1 位点rs338556芯片分型结果

表1 2组相关临床指标比较($\bar{x} \pm s$)

组别	n	年龄(岁)	BMI(kg/m²)	腰围(cm)	腰臀比	SBP (mmHg)	DBP (mmHg)
对照组	52	56.43±8.19	24.01±3.10	89.97±7.88	0.93±0.06	120.30±17.04	74.35±5.33
病例组	34	62.67±8.49[a]	23.64±2.64	89.75±10.78	0.94±0.07	126.06±10.27	74.21±4.11
组别	n	空腹血糖(mmol/L)	HbA1c(%)	TG(mmol/L)	TC(mmol/L)	HDL-C(mmol/L)	LDL-C(mmol/L)
对照组	52	6.93±1.24	6.57±0.89	1.80±1.74	4.37±0.87	1.30±0.29	2.48±0.68
病例组	34	6.67±1.27	6.55±1.08	1.53±0.79	4.52±0.70	1.38±0.34	2.50±0.52

a：P＜0.05，与对照组比较

表 2 与 T2DM 大血管病变相关主要代谢途径内相关的差异 SNP 位点列表

dbSNP ID	Gene	Gene nearby(kbaway)	Region	Alleles	*P* 值	所属 T2DM 大血管病变代谢通路
rs650616	DDOST		5′ UTR	A/G	0.000 093 45	氧化应激反应
rs12431702	PRKD1		intron	T/G	0.000 114 67	炎症反应
rs10799655		PINK1(2) DDOST(21)		T/C	0.000 129 08	氧化应激
rs781443		AGPAT9 (-32)		A/C	0.000 300 92	脂质代谢
rs4969425	RPTOR		intron	T/C	0.000 378 67	胰岛素信号转导
rs3169121	AKAP13		3′ UTR	T/G	0.000 480 82	血管内皮功能完整性
rs11884290		THSD7B(29)		A/G	0.000 544 93	血凝调节
rs4969230	RPTOR		intron	A/G	0.000 557 52	胰岛素信号转导
rs4790646		ALOX15(62)		A/G	0.000 558 21	炎症反应
rs968726		RPH3AL(63)		T/C	0.000 597 32	胰岛素信号转导
rs4896219	MAP3K5		intron	T/G	0.000 604 65	胰岛素信号转导
rs2448026	AKAP13		intron	A/G	0.000 610 73	血管内皮功能完整性
rs3742076	FOXM1		exon	T/C	0.000 702 17	血管内皮功能完整性
rs6060959		MYLK2(17)		T/C	0.000 702 64	血凝调节
rs4243449	MAP3K5		intron	A/G	0.000 761 05	胰岛素信号转导
rs2542606		AKAP13(-10)		A/G	0.000 761 02	血管内皮功能完整性
rs527912		PINK1(25) DDOST(44)		A/G	0.000 793 47	氧化应激反应
rs338556	AKAP13			T/C	0.000 809 13	血管内皮功能完整性
rs3775195	VEGFC		intron	T/G	0.000 809 11	血管内皮功能完整性以及炎症反应
rs11062386		FOXM1(3)		A/G	0.000 867 52	血管内皮功能完整性
rs3733328	AGPAT9		intron	A/C	0.000 927 37	脂质代谢
rs12213573		PECI(-76)		T/C	0.000 928 64	脂质代谢
rs2772395		RUNX2(-6)		T/C	0.000 946 82	血管内皮功能完整性
rs3094513	HNF1B		intron	T/C	0.000 953 34	和多条代谢通路相关
rs2641256	AGPAT9		exon	T/C	0.000 953 32	脂质代谢
rs312077		GNA11(-90)		T/C	0.000 336 77	血凝调节
rs26770	ARL15		intron	T/C	0.000 460 75	胰岛素信号转导
rs262426	LDLRAD3		intron	T/G	0.000 472 12	脂质代谢
rs4854610		SLCO2A1(-56)		T/C	0.000 493 13	血管内皮功能完整性
rs728115	KCND2		intron	T/C	0.000 676 87	胰岛素信号转导
rs2238455	ADCY9		intron	T/C	0.000 737 21	胰岛素信号转导
rs7407025	HNF1B		intron	T/C	0.000 746 73	和多条代谢通路相关
rs7596463		NMS(-50)		A/G	0.000 805 82	血管内皮功能完整性
rs1957516		CYP46A1(-15)		T/C	0.000 836 68	脂质代谢
rs1362355		CHN2(-52)		T/C	0.000 879 23	动脉基质平衡
rs2529412	GRB10		intron	A/G	0.000 916 21	胰岛素信号转导
rs9294970		THBS2(46)		T/C	0.000 938 67	血凝调节

2.3 病例组和对照组遗传背景 PCA 分析结果

World 组、Asia 组的 PCA 结果显示，病例组和对照组间以及样本内部均不存在显著的遗传背景差异。Case-control 组的 PCA 结果显示，没有明显的主成分可以区分样本，说明样本不存在人群分层，也不存在样本逸出。

2.4 全基因组关联分析以及 Logistic 回归分析筛选出的差异位点

用 PLINK 1.06 软件对芯片结果进行 Allelic 和 Genotype 两种遗传模型下的分析，并以年龄为协变量作 Logistic 回归分析，得出在病例组和对照组间差异显著（$P < 0.01$）的位点总计 452 个。

2.5 对差异位点进行基因定位和功能学筛选

通过 UCSC genome browser 对得出的位点进行基因定位，找出这些 SNP 位点所处的基因以及其附近的基因（±100 kb 范围以内）。通过 NCBI 数据库对这些基因进行了功能注释。以与 T2DM 大血管病变相关的主要代谢途径或者机制为标准进行功能学筛选，总计筛选出 SNP 位点 37 个。这 37 个位点分处于 30 条基因内或附近。具体结果见表 2。

3 讨论

糖尿病患者大血管病变（包括脑血管、心血管以及下肢血管病变）的发病率高于普通人群[3]，已成为我国糖尿病患者主要的致残致死原因。目前认为，T2DM 大血管病变的发生与遗传因素密切相关，但其属多基因复杂疾病，由于涉及的代谢通路众多，且各通路中又包含大量的相关基因及蛋白产物，因而对其遗传背景的研究极具挑战。本研究中的病例组和对照组患者分别是经多因素干预治疗后发生早期动脉硬化和未发生动脉硬化的 2 型糖尿病患者，由于统一了干预标准，两组对象除年龄之外相关检查指标如体质量、血压、血糖、血脂等代谢指标干预达标情况均无显著性差

异,因此尽可能减少了环境因素的影响,这样观察遗传因素的作用会更加准确。同时本研究采用全基因组SNP分型芯片进行检测和全基因组关联分析,全基因组关联分析是寻找复杂疾病遗传风险因子最有效的方法之一[4-5],这些都使我们能够从全基因组水平考察遗传因素对疾病的影响。

初步筛选出的30个基因与T2DM大血管病变发生发展的6条主要代谢通路相关,如rs26770所处的ARL15(ADP核糖基化样因子15)基因,其基因内存在的某些遗传多态性可导致外周血中脂联素(adiponectin,ADPN)水平降低[6],而脂联素具有降糖,提高胰岛素敏感性和抗炎、抗动脉粥样硬化的作用,与T2DM大血管病变发生密切相关。rs3742076处于的FOXM1(forkhead box M1)基因,通过对血管内皮细胞周期的影响进而影响其在受到损伤时的修复,而损伤修复功能的下降是大血管病变发生的一个重要原因[7]。THSD7B和THSD2基因参与编码血小板凝集素1。这个蛋白可以使凝血机制增强和促进附壁血栓形成,同时释放炎症介质及细胞增殖因子,引起血管壁的重塑,加速动脉粥样斑块的形成。rs12431702所处的PRKD1(蛋白激酶1)基因以及rs3775195所处的VEGFC基因共同参与了伴随粥样斑块形成全程的炎症反应,它们所构筑的VEGF- PRKD1炎症反应轴介导了诸多血管内皮内炎症因子(如IL-6, IL-8, GRO-alpha)的产生[8]。而血管内皮因子家族单独也可影响血管内皮细胞的通透性和血管内皮细胞的损伤修复,刺激血管平滑肌增生,参与了大血管病变的多个病理过程[9]。rs650616所处的DDOST基因编码的多萜基长醇二磷酸寡糖蛋白环糊精糖基转移酶影响晚期糖基化终末产物(AGEs)的形成,AGEs对血管内皮细胞内的活性氧簇(reactive oxygen species, ROS)有强烈的抑制作用,从而通过抑制细胞内的氧化作用来抵抗血管内皮细胞的损伤[10]。血管平滑肌的增殖与迁移也是大血管病变的一个重要病理过程,rs1362355下游52kb的嵌合素2(chimerin 2, CHN2)基因编码蛋白可以刺激血管平滑肌的增生和迁移,是大血管病变的一个重要因素[11]。rs262426所处的LDLRAD3(low density lipoprotein receptor class A domain containing 3),该基因参与编码低密度脂蛋白受体,虽然至今无文献报道其与大血管病变相关,但该受体可表达于血管内皮细胞,与动脉粥样斑块形成早期的脂质条纹的形成密切相关,故也将其筛选入内,其具体的生物学意义还有待于下一步实验验证。此外,HNF1B(肝细胞核因子1 HOMEBOX B)是T2DM发生的一个重要易感基因,其多态性可以影响包括脂联素水平在内的多种T2DM病理过程,虽目前无文献报道其与大血管病变有直接联系,但它对T2DM诸多已知易感基因的调控提示其在T2DM大血管病变中可能起着上游调节的作用[12]。

在筛选出的37个SNP位点中,有17个处于基因区域之外,20个处于基因区域内的SNP绝大部分也分布于内含子区域,这与我们所采用的Illumina芯片主要是检测标签SNP(tagSNP)为主的情况符合。此外,由于对有些基因以及SNP的功能目前还知之甚少,除了我们筛选出来的37个位点,剩余的413个位点也不能排除其与T2DM大血管病变之间存在联系,可以作为后续研究的参考。

本研究主要着眼于对T2DM大血管病变易感基因以及遗传标记的筛查,为体外验证这些遗传标记的生物学功能奠定基础,为尽早识别易感者,尽早提供有效干预措施和针对性的个体化,从而为预防或延缓大血管病变的发生提供理论依据。

参考文献:

[1] Ordovas J M. Genetic links between diabetes mellitus and coronary atherosclerosis[J]. Curr Atheroscler Rep, 2007, 9(3):204-210.

[2] Fumeron F, Reis A F, Velho G. Genetics of macrovascular complications in diabetes[J]. Curr Diab Rep, 2006, 6(2):162-168.

[3] Ziegler D. Type 2 diabetes as an inflammatory cardiovascular disorder[J]. Curr Mol Med, 2005, 5(3):309-322.

[4] Kronenberg F. Genome-wide association studies in aging-related processes such as diabetes mellitus, atherosclerosis and cancer[J]. Exp Gerontol, 2008, 43(1):39-43.

[5] Hattersley A T, McCarthy M I. What makes a good genetic association study?[J]. Lancet, 2005, 366(9493):1315-1323.

[6] Richards J B, Waterworth D, O'Rahilly S, *et al*. A genome-wide association study reveals variants in ARL15 that influence adiponectin levels[J]. PLoS Genet, 2009, 5(12):e1000768.

[7] Zhao Y Y, Gao X P, Zhao Y D, *et al*. Endothelial cell-restricted disruption of FoxM1 impairs endothelial repair following LPS-induced vascular injury[J]. J Clin Invest, 2006, 116(9):2333-2343.

[8] Hao Q, Wang L, Tang H. Vascular endothelial growth factor induces protein kinase D-dependent production of proinflammatory cytokines in endothelial cells[J]. Am J Physiol Cell Physiol, 2009, 296(4):C821-C827.

[9] 王海珍. 血管内皮生长因子与2型糖尿病大血管病变的关系[J]. 中国煤炭工业医学杂志, 2007, 10(12):1381-1382.

[10] Cai W, Torreggiani M, Zhu L, *et al*. AGER1 regulates endothelial cell NADPH oxidase-dependent oxidant stress *via* PKC-delta:implications for vascular disease[J]. Am J Physiol Cell Physiol, 2010, 298(3):C624-C634.

[11] Maeda M, Kato S, Fukushima S, *et al*. Regulation of vascular smooth muscle proliferation and migration by beta2-chimaerin, a non-protein kinase C phorbol ester receptor[J]. Int J Mol Med, 2006, 17(4):559-566.

[12] Ruchat S M, Elks C E, Loos R J, *et al*. Evidence of interaction between type 2 diabetes susceptibility genes and dietary fat intake for adiposity and glucose homeostasis-related phenotypes[J]. J Nutrigenet Nutrigenomics, 2009, 2(4/5):225-234.

(收稿:2010-10-11 修回:2010-11-30)

(编辑 汪勤俭)