

• 论 著 •

葡萄糖激酶基因 4 个标签单核苷酸多态性位点与 2 型糖尿病的相关性研究^{*}

孙各琴, 张秀明, 罗楚君, 李晶晶, 韩 慧, 陈 琼
(中山大学附属中山医院医学检验中心, 广东中山 528403)

摘要:目的 探讨葡萄糖激酶(GCK)基因 4 个标签单核苷酸多态性(tagSNPs)位点 rs12702070、rs2268569、rs2268573、rs1476891 与 2 型糖尿病(T2D)的关系。**方法** 选取中山大学附属中山医院住院的中国南方汉族 T2D 患者 499 例(T2D 组),同时选择汉族健康人 499 例作为对照组,对 GCK 基因的 4 个 tagSNPs 位点进行基因分型并检测样本代表性。比较 T2D 组和对照组基因型和等位基因频率的差异,并在 3 种遗传模型下对各 SNP 位点进行相关性分析。构建 GCK 基因 4 个 tagSNPs 位点的单体型,分析是否存在连锁不平衡(LD)及不同的 GCK 单体型与 T2D 易感性的关系。**结果** rs2268573、rs2268569、rs1476891 的基因型和等位基因频率在 T2D 组和对照组之间差异均无统计学意义。rs12702070 的基因型和等位基因分布在 D2M 组和对照组之间差异有统计学意义,且在显性遗传模式下以及在加性遗传模式下其基因型分布在两组之间差异有统计学意义。GCK 基因 4 个位点中 rs2268569、rs12702070 和 rs1476891 有 1 个 LD 域,其中共存四种主要单体型,均与个体患 T2D 的风险无相关性。**结论** 在汉族人群中,GCK 基因区域的 rs12702070 位点与糖尿病遗传易感性密切相关,而 rs2268573、rs2268569、rs1476891 位点与糖尿病遗传易感性无明确相关性。rs2268569、rs12702070、rs1476891 LD 域四种主要单体型均与个体患 T2D 的风险无相关性。

关键词: 2 型糖尿病; 基因; 葡萄糖激酶; 多态性,单核苷酸

DOI:10.3969/j.issn.1673-4130.2016.18.002

文献标识码:A

文章编号:1673-4130(2016)18-2507-04

The relationship between glucokinase gene 4 tag single nucleotide polymorphism sites and type 2 diabetes mellitus^{*}

SUN Geqin, ZHANG Xiuming, LUO Chujun, LI Jingjing, HAN Hui, CHEN Qiong

(Department of Examination Medical center, Zhongshan Affiliated Hospital of Sun Yat-sen University, Zhongshan, Guangdong 528403, China)

Abstract: Objective To investigate the relationships between Glucokinase(GCK) gene 4 tag single-nucleotide polymorphisms, tagSNPs) sites which named rs12702070, rs2268569, rs2268573 and rs1476891 polymorphisms and type 2 diabetes in Chinese Southern Han Population. **Methods** This study was designed as a case-control. 499 type 2 diabetes patients and 499 healthy controls were chosen. 4 GCK tagSNPs sites were analyzed by improved multiple ligase detection reaction(iMLDR), and genotype and allele frequency between T2D group and healthy controls could be determined by chi-square test, logistic regression analysis, and tagSNPs were further analyzed under three genetic modes(dominant, recessive and additive). What's more, Haploview software was used to construct the haplotype of 4 GCK tagSNPs and the linkage disequilibrium(LD) and relationship between various GCK haplotype and T2D susceptibility could be analyzed. **Results** Genotype distribution of rs2268573, rs2268569 and rs1476891 and allele frequency in T2D group were no significant differences with health controls. Significant differences in genotype distribution of rs12702070 and allele frequency were observed between T2D group and health controls. Under dominant model and additive model, genotype distribution of rs12702070 in T2D was significantly different from health controls. One LD domain was observed in 3 tagSNPs among those 4 sites of GCK gene. There are 4 main haplotypes(TAG, TGG, TAT, CGG) in rs2268569, rs12702070 and rs1476891, but all these haplotypes have no relevance to the individual risk of T2D($P > 0.05$). **Conclusion** The results indicated that the GCK gene tagSNPs site rs12702070 imparts susceptibility to T2D in Han Chinese, but not rs2268573, rs2268569 and rs1476891. Four main haplotypes in rs2268569, rs12702070 and rs1476891 have no relevance to T2D.

Key words: type 2 diabetes; gene; glucokinase; polymorphism, single nucleotide

糖尿病是一组由胰岛素分泌不足或新陈代谢紊乱所导致的高血糖病。据估计,2013 年在中国就有 9 800 万糖尿病患者,居世界首位^[1]。遗传因素在糖尿病的发生和发展中起重要作用。全基因组关联研究(GWAS)和候选基因方法已被广泛用于解释糖尿病的遗传基础,并发现了多个与糖尿病相关的基因^[2-3]。但是,糖尿病相关的基因数目和特性、潜在的遗传模型以及与环境因素的相互作用还不清楚^[4]。葡萄糖激酶(GCK)

是糖酵解的关键酶,也是通过 β 细胞维持葡萄糖自身稳定的传感器^[5]。GCK 基因位于 7 号染色体上(7p15.3~p15.1),由 12 个外显子组成,编码由 465 个氨基酸组成的蛋白质^[6]。通过对 GCK 基因的标签单核苷酸多态性(tagSNPs)网上系统分析和文献查阅,与糖尿病相关的 GCK 基因 4 个 tagSNPs 位点 rs12702070、rs2268569、rs2268573、rs1476891 引起了笔者的关注。因此,本研究对 GCK 基因上的这 4 个位点与中国汉族人

^{*} 基金项目:广东省科技计划基金资助项目(2013B02180012)。

作者简介:孙各琴,女,副主任技师,主要从事分子生物学方面的研究。

2 型糖尿病 (T2D) 易感性的关系进行探讨。

1 资料与方法

1.1 一般资料 选取 2013 年 8 月至 2014 年 12 月在本院诊治的 T2D 患者 499 例 (T2D 组), 均符合 1999 年世界卫生组织糖尿病诊断标准, 其中男 257 例, 女 242 例, 年龄 28~93 岁, 平均 58.6 岁^[7]。另取同期体检健康者 499 例作为对照组, 其中男 290 例, 女 209 例, 年龄 27~86 岁, 平均 55.7 岁。本研究已通过医院伦理委员会审核 (项目名称: 糖尿病 HbA1c 相关的遗传易感基因研究), 患者均签署知情同意书。

1.2 仪器与试剂 DU800 核酸检测仪、5415D 高速离心机 (德国 Eppendorf 公司), DYY-6B 型稳压稳流电泳仪 (北京六一仪器厂), UVP 凝胶成像系统 (美国 UVP 公司), 血液基因组提取试剂盒 (Gentra Puregene Blood Kit, 美国 Qiagen 公司)。

1.3 标本采集 入选者知情同意后签字, 分别抽取静脉血 5 mL, EDTA 抗凝, 室温放置 15 min, 2 000 r/min (400 g) 离心 10 min, 吸取上层血浆, 有核细胞层经红细胞裂解液处理后, 置-70℃低温冰箱保存备用。

1.4 基因组 DNA 提取 采用 Gentra Puregene Blood Kit 基因组提取试剂盒, 提取样本外周血白细胞基因组 DNA, DNA 样本经 1% 琼脂糖凝胶电泳对其进行完整度检查, 核酸检测仪测定浓度, 并将样本 DNA 稀释到工作浓度 (5~10 ng/L)。

1.5 基因检测 SNP 位点的检测使用改良多重高温连接酶检测反应技术 (iMLDR), 所需 PCR 扩增引物和多重单碱基延伸反应延伸引物 (表 1) 均由上海天昊生物科技有限公司设计, 上海生工生物科技有限公司合成, 按要求稀释到指定浓度。由于待测的 SNP 附近有其他高频的 SNP, 设计连接引物时必须紧挨目的位点, 所以必定要覆盖到这些附近的 SNP。因此, 在设计连接引物时两种碱基都合成, 这样才不会由于只设计其中一种, 而另一种多态的染色体因为和引物存在错配而效率低, 甚至不出现碱基峰, 从而导致基因型判读错误。分型试验委托上海天昊生物科技有限公司进行。反应体系 (10 μL) 包含: 1×GCI buffer, 3.0 mmol/L Mg²⁺, 0.3 mmol/L dNTP, 1 U HotStar Taq polymerase, 1 μL 样本 DNA 和 1 μL 多重 PCR 引物。PCR 循环程序: 95℃ 2 min; 11 个循环 (94℃ 20 s, 65℃ 和 0.5℃/cycle 40 s, 72℃ 90 s); 24 个循环 (94℃ 20 s, 59℃ 30 s, 72℃ 90 s); 4℃ 60 min; 在 PCR 产物中加入试剂盒中的 Exo I /SAP 酶 1 μL, 37℃ 温浴 1 h, 然后 75℃ 灭活 15 min; 取 10×连接缓冲液 2 μL、高温连接酶 0.2 μL、连接引物混合液 1 μL 与纯化后多重 PCR 产物 3 μL、ddH₂O 3.8 μL, 混匀。连接程序: 38 个循环 (94℃ 1 min, 56℃ 4 min); 4℃ 60 min; 取 0.5 μL 稀释后的连接产物, 与 0.5 μL Liz500 SIZE STANDARD 9 μL HiDi 混匀, 95℃ 变性 5 min 后上 ABI 3130 XL 测序仪; 收集的原始数据用 GeneMapper 4.1 (美国应用生物系统公司) 分析并输出结果。

1.6 相关定义 如果双亲的性状同时在 F1 个体上表现出

来, 这种显性表现称为共显性, 或叫并显性。显性就是表现出来的特征, 隐性是没有表现出来但是自身携带。比如 rs2235321 (G>A), G 为野生型、A 为突变型。共显性模型假设等位基因 G 和 A 均显性表达, 致病基因型为 GA、AA 型; 显性模型假定突变型 A 为显性致病基因, 致病基因型为 GA、AA 型; 隐性模型假定突变型 A 为隐性致病基因, 致病基因型为 AA 型; 超显性模型假定纯合子不发病而杂合子发病, 致病基因型为 GA 型。

1.7 统计学处理 采用病例对照研究。运用 Hardy-Weinberg 平衡规律检测样本代表性; 采用 SPSS17.0 统计软件处理基因型和等位基因频率结果, 计数资料各百分率比较用 χ^2 检验, $P<0.05$ 表示差异有统计学意义; 运用 Haploview 软件进行 GCK 的 4 个位点 tagSNPs 单体型的构建; <http://bioinfo.iconcologia.net/snpstats/start.htm> 在线分析 LD, 分析是否存在 LD 及不同的 GCK 单体型与 T2D 风险的关系。

表 1 扩增 SNPs 的引物序列

位点	引物名称	序列 (5'~3')
rs2268573	rs2268573F	GGTGACATGGCCACAATTTGAA
	rs2268573R	TGCATCTTCCAGCTCTTCGACTACA
rs2268569	rs2268569F	GGTTCCCCCAAGACAAGGACTG
	rs2268569R	AATGCCTCGGGCACAAACTTTA
rs12702070	rs12702070F	TCTCCTGCCAGGGCTTACTGTG
	rs12702070R	GCCCCAGGATCTGAACAGGTG
rs1476891	rs1476891F	TCAAGTGTCACGCTGACTGGT
	rs1476891R	TCCAGGAGCCATCCAGAAGAAA

2 结 果

2.1 候选 SNPs 基因型与等位基因频率 采用 iMLDR 多重 SNP 分型试剂盒对 998 例样本的 4 个 SNP 进行分型。经统计发现, 所有 SNP 位点均符合 Hardy-Weinberg 检验平衡定律, 提示所选人群具有较好代表性。通过多变量回归分析, rs2268573、rs2268569、rs1476891 的基因型 (χ^2 值分别为 3.361、0.582、0.918, $P>0.05$) 和等位基因频率 (χ^2 值分别为 0.222、0.590、0.851, $P>0.05$) 在 T2D 组和对照组之间差异均无统计学意义 (见表 2)。rs12702070 的基因型 (χ^2 值为 7.990, $P<0.05$) 和等位基因分布 (χ^2 值为 5.979, $P<0.05$) 在 T2D 组和对照组之间差异有统计学意义 (表 2), 提示此位点 SNP 多态性与 T2D 相关。在显性遗传模式下 ($OR=1.54$, 95% $CI=1.17\sim2.04$, $P<0.05$) 以及在加性遗传模式下 ($OR=1.26$, 95% $CI=1.04\sim1.52$, $P<0.05$) rs12702070 的基因型分布在 D2M 组和对照组之间差异有统计学意义 (表 3), 提示此位点在显性和加性遗传模式下与 T2D 的易感性相关。

表 2 SNPs 位点基因型及等位基因分布在 2 组的分析

SNP 位点	基因型分析					等位基因分析				
	基因型	T2D 组[n(%)]	对照组[n(%)]	χ^2 值	P 值	基因型	T2D 组[n(%)]	对照组[n(%)]	χ^2 值	P 值
rs2268573	G/G	208(0.4)	228(0.5)	3.361	0.120	G	651(65.2)	661(66.2)	0.222	0.670
	G/T	235(0.5)	205(0.4)			T	347(34.8)	337(33.8)		

续表 2 SNPs 位点基因型及等位基因分布在 2 组的分析

SNP 位点	基因型分析					等位基因分析				
	基因型	T2D 组[n(%)]	对照组[n(%)]	χ^2 值	P 值	基因型	T2D 组[n(%)]	对照组[n(%)]	χ^2 值	P 值
T/T	56(0. 1)	66(0. 1)								
rs12702070	T/T	446(89. 4)	420(84. 2)	6. 896	0. 008	C	59(5. 9)	84(8. 4)	4. 708	0. 018
	C/T	47(9. 4)	74(14. 8)			T	939(94. 1)	914(91. 6)		
	C/C	6(1. 2)	5(1. 0)							
rs2268569	G/G	338(67. 7)	349(69. 9)	0. 582	0. 820	G	818(82. 0)	831(83. 3)	0. 590	0. 443
	G/T	142(28. 5)	133(26. 6)			T	180(18. 0)	167(16. 7)		
	T/T	19(3. 8)	17(3. 4)							
rs1476891	A/A	281(56. 3)	269(53. 9)	0. 918	0. 800	A	751(75. 3)	733(73. 4)	0. 851	0. 192
	G/A	189(37. 9)	195(39. 1)			G	247(24. 7)	265(26. 6)		
	G/G	29(5. 8)	35(7. 0)							

表 3 不同遗传背景下各 SNP 位点分析

SNP 位点	显性模型		隐性模型		加性模型	
	OR(95 %CI)	P 值	OR(95 %CI)	P 值	OR(95 %CI)	P 值
rs2268573	0. 85(0. 66~1. 09)	0. 200	1. 21(0. 82~1. 76)	0. 330	0. 96(0. 80~1. 15)	0. 640
rs12702070	1. 74(1. 17~2. 57)	0. 005	0. 68(0. 20~2. 38)	0. 550	1. 51(1. 06~2. 14)	0. 020
rs2268569	0. 90(0. 69~1. 18)	0. 450	0. 89(0. 46~1. 74)	0. 730	0. 92(0. 73~1. 15)	0. 450
rs1476891	1. 11(0. 86~1. 41)	0. 440	1. 22(0. 74~2. 03)	0. 440	1. 10(0. 90~1. 35)	0. 350

表 4 rs12702070C/T、rs1476891G/A 和 rs2268569T/G 位点配对单体型频率比较

rs12702070	rs1476891	rs2268569	T2D 组频率(%)	对照组频率(%)	OR 值	95 %CI	P 值
T	A	G	0. 572 1	0. 567 1	1. 00	—	—
T	G	G	0. 188 4	0. 181 4	0. 97	0. 76~1. 23	0. 790
T	A	T	0. 180 4	0. 167 3	0. 94	0. 74~1. 19	0. 610
C	G	G	0. 059 1	0. 084 2	1. 39	0. 99~1. 96	0. 056

注:—表示无数据。

2.2 GCK 单体型的构建及与 T2D 风险的关系 运用 Haplo-view 软件进行单体型构建。GCK 基因 4 个位点的 3 个位点组成 LD 域,即为 rs2268569、rs12702070 和 rs1476891。GCK 的 rs2268569、rs12702070 和 rs1476891 3 个位点共存在 TAG、TGG、TAT、CGG 四种主要单体型(表 4),均与个体患 T2D 的风险无相关性($\chi^2=2. 718,P>0. 05$)。

3 讨 论

近年来,探讨遗传因素在糖尿病发生和发展中的作用倍受关注。已有研究表明,钙蛋白酶 10(Calpain-10)基因、KIR6. 2 基因、HNF-4 α 基因、TCF7L2 基因等与糖尿病的发生和发展相关^[8]。本研究通过对 GCK 基因的 tagSNP 网上系统分析和文献查阅,选取 GCK 基因 4 个 tagSNP 位点 rs12702070、rs2268569、rs2268573、rs1476891,探讨其与中国汉族人 T2D 风险的关系。本研究在挑选位点时,基于以下两点:(1)在疾病相关基因中,基因编码区、5'非编码区(5' UTR)、3'非编码区(3' UTR)、5'附近的基因(5' near gene)、3'附近的基因(3' near gene)的 SNP 位点,最小等位基因频率(MAF)一般大于 5%,进入美国国家生物技术信息中心(NCBI)的 SNP 数据库网站([http://www.ncbi.nlm.](http://www.ncbi.nlm.nih.gov/snp/)

[nih.gov/snp/](http://www.ncbi.nlm.nih.gov/snp/)),输入目的基因,找出 MAF 频率大于 0. 05 的多态性位点,这些位点在基因的编码区、5' UTR、3'UTR、5' near gene、3' near gene。同时,注意选择人种,因为不同的人种各 SNP 的频率是不同的。(2)1 个基因上有很多个 SNP 位点,并且基因上位点间往往是有关联的,可以用 1 个位点或多个位点的组合代表另外的位点或单体型,所以,可以挑选有代表性的位点来代表基因上其他的位点或人群中存在的常见单体型。在选取 tagSNP 时主要考虑序列范围(一般常用基因前 5k 和后 2k)、SNP 入选频率标准(MAF $\geq 0. 05$)和 r^2 标准($r^2 \geq 0. 8$)。

在 dbSNP 数据库中查询得到:rs12702070 位点的 MAF 值 C 在欧洲和中国南方汉族人群分别占 0. 08 和 0. 076。有文献报道,南斯拉夫高加索白单基因糖尿病患者出现 rs12702070 位点突变,而对照组未见该位点突变^[9]。本研究结果显示,GCK 位点 rs12702070 与 T2D 有关,在对背景、年龄和性别做出调整后,仍与 T2D 相关。在显性遗传模式和加性遗传模式下,rs12702070 tagSNPs 位点的基因型分布在 D2M 组和对照组之间差异有统计学意义(见表 3),提示此位点在显性和加性

遗传模式下与 T2D 的易感性密切相关,可能增加个体患 T2D 的风险,支持上述文献的研究结论。此外,有文献还报道,印第安人群 rs2268573 突变与 T2D 有相关性,而 rs2268569 突变在西班牙人、欧洲裔美国人、非裔美国人、中国人群与空腹血糖的水平有关,rs1476891 则与空腹血糖的水平无关^[10-12]。本研究发现,GCK 位点 rs2268569、rs2268573、rs1476891 等位基因及基因型分布频率与健康人群相比较差异无统计学意义(表 2),与上述研究不完全一致,提示这些基因与糖尿病及糖代谢的关系尚需深入研究。单体型研究结果显示,rs12702070、rs1476891 和 rs2268569 形成了 LD 域,表达显著关联($P < 0.05$),表明在这 3 个 SNP 位点中有基因-基因相互作用。

总之,本研究初步验证了在汉族人群中,GCK 基因区域的 rs12702070 位点与糖尿病遗传易感性密切相关,而 rs2268573、rs2268569、rs1476891 位点与糖尿病遗传易感性无明确相关性。rs12702070、rs1476891 和 rs2268569 的 LD 域有基因-基因相互作用。但这些突变基因的具体作用机制仍不明确,值得深入研究,可在今后结合更多临床资料及数据进一步对比分析。

参考文献

[1] 刘子杰,晋臻,段勇. HbA1c 常规工作中引入测量不确定度的意义和存在的困难[J]. 中华检验医学杂志,2014,37(12):884-886.

[2] Sladek R,Rochelleau G,Rung J, et al. A genome-wide association study identifies novel risk loci for type 2 diabetes[J]. Nature,2007,445(7130):881-885.

[3] Billings LK,Florez JC. The genetics of type 2 diabetes: what have we learned from GWAS[J]. Ann N Y Acad Sci,2010,1212:59-77.

[4] Bao XY,Peng B,Yang MS. Replication study of novel risk variants in six genes with type 2 diabetes and related quantitative traits in the Han Chinese lean individuals[J].

Mol Biol Rep,2012,39(3):2447-2454.

[5] Matschinsky F,Liang Y,Kesavan P, et al. Glucokinase as pancreatic beta-cell glucose sensor and diabetes gene[J]. Journal of Clinical Investigation,1993,92(5):2092-2098.

[6] Iynedjian PB. Mammalian glucokinase and its gene[J]. Biochem J,1993,293(Pt 1):1-13.

[7] WHO. Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1: diagnosis and classification of diabetes mellitus[S]. WHO, Geneva, 1999: 1-59.

[8] 张秀明,黄宪章,曾方银,等. 临床生化检验诊断学[M]. 北京:人民卫生出版社,2012:866.

[9] Borowiec M,Antosik K,Fendler W, et al. Novel glucokinase mutations in patients with monogenic diabetes-clinical outline of GCK-MD and potential for founder effect in Slavic population[J]. Clin Genet,2012,81(3):278-283.

[10] Echouffo-Tcheugui JB,Dieffenbach SD,Kengne AP. Added value of novel circulating and genetic biomarkers in type 2 diabetes prediction:a systematic review[J]. Diabetes Res Clin Pract,2013,101(3):255-269.

[11] Rasmussen-Torvik LJ,Guo XQ,Bowden DW, et al. Fast-ing glucose GWAS candidate region analysis across ethnic groups in the multiethnic study of atherosclerosis(Mesa) [J]. Genet Epidemiol,2012,36(4):384-391.

[12] Weedon MN,Clark VJ,Qian YA, et al. A common haplo-type of the glucokinase gene alters fasting glucose and birth weight: Association in six studies and population-genetics analyses[J]. Am J Hum Genet,2006,79(6):991-1001.

(收稿日期:2016-03-12 修回日期:2016-05-20)

(上接第 2506 页)

is a Ubiquitin-Protein ligase that promotes p53 degradation[J]. Cellular Physiology and Biochemistry, 2015, 35(1):237-245.

[5] Lee JY,Kong G. MEL-18, a tumor suppressor for aggressive breast cancer[J]. Oncotarget, 2015, 6(18):15710-15711.

[6] 黄湘,谭家余,王穗海,等. 过表达 HSPC238 对 Bel7402 细胞周期的影响[J]. 中国免疫学杂志,2011,27(2):120-125.

[7] Sun XX,Devine T,Challagundla KB, et al. Interplay between ribosomal protein S27a and MDM2 protein in p53 activation in response to ribosomal stress[J]. J Biol Chem,2011,286(26):22730-22741.

[8] Tan JY,Chen JL,Huang X, et al. Screening and verification of proteins that interact with HSPC238[J]. Oncol Rep,2015,34(6):3097-3103.

[9] 谭家余,黄湘,陈敬林,等. HSPC238 对 HMOX1、RPS27a、

MT2A、UBB 调节的初步研究[J]. 中国免疫学杂志,2016, 32(4):509-512.

[10] Yin YX,Solomon G,Deng CX, et al. Differential regulation of p21 by p53 and Rb in cellular response to oxidative stress[J]. Mol Carcinog,1999,24(1):15-24.

[11] Schwartzman JM,Duijf PH,Sotillo RA, et al. Mad2 is a critical mediator of the chromosome instability observed upon Rb and p53 pathway inhibition[J]. Cancer Cell, 2011,19(6):701-714.

[12] Iaquina PJ,Aslanian A,Lees JA, et al. Regulation of the Arf/p53 tumor surveillance network by E2F[J]. Cold Spring HaRb Symp Quant Biol,2005,70:309-316.

[13] Aksoy O,Chicas A,Zeng TY, et al. The atypical E2F family member E2F7 couples the p53 and RB pathways during cellular senescence[J]. Genes Dev,2012,26(14): 1546-1557.

(收稿日期:2016-04-21 修回日期:2016-06-28)