

GATA6 基因启动区多态性与急性心肌梗死患者易感性和临床特征的相关性分析

孙兆庆¹, 逢淑超², 崔英华^{3,4}, 闫波^{2,4,5}

(1. 山东大学齐鲁医学院, 山东省济南市 250012; 济宁医学院附属医院 2. 山东省心脏疾病诊疗重点实验室, 3. 心内科, 4. 心血管疾病分子遗传学中心, 5 山东省中美转化医学合作研究中心, 山东省济宁市 272029)

[关键词] 急性心肌梗死; 基因启动区; 单核苷酸多态性; 危险因素; 连锁不平衡

[摘要] 目的 探讨 GATA6 基因启动区-493(rs144923558)G/A、-172(rs146748749)G/A 2 个位点单核苷酸多态性(SNP)与急性心肌梗死(AMI)之间的关系及其相关的危险因素。方法 采用病例-对照研究方法,收集 328 例 AMI 患者和 344 例正常对照;应用聚合酶链反应-限制性内切酶片段长度多态性技术结合 DNA 测序后的序列比对进行数据统计;运用 Hardy-Weinberg 平衡检验后,应用 χ^2 检验进行相关性分析;利用 Logistic 回归对多种危险因素以及 2 个 SNP 位点与 AMI 发病进行关联性分析;利用 Haploview4.2 软件和 SHEsis 网站进行连锁不平衡及单倍体分析。结果 2 个 SNP 位点共检测出 3 种基因型为 GG、GA 和 AA,其基因型分布均符合 Hardy-Weinberg 平衡($P>0.05$),同时在 AMI 组与对照组间差异无显著性($P>0.05$)。多因素 Logistic 回归分析:增龄、高血压病、吸烟、低密度脂蛋白胆固醇(LDL-C)和甘油三酯(TG)升高是 AMI 发病的独立危险因素($P<0.05$),HDL-C 为保护因素($P<0.05$)。在显性、隐性和加性 3 种不同遗传模式下进行 Logistic 回归分析发现,2 个 SNP 位点与 AMI 发病无关联性。进行连锁不平衡和单倍体分析提示,该 2 个 SNP 位点处于同一个连锁不平衡区域($D'=1.000, r^2=1.000$),单倍型 GG 和 AA 均未增加 AMI 易感性($P>0.05$)。结论 GATA6 基因启动区-493(rs144923558)G/A 与-172(rs146748749)G/A 两个 SNP 位点为完全连锁不平衡,其中 GG 为主要单倍型。该 2 个 SNP 位点及其单倍体型与 AMI 发病无相关性,但提供了 GATA6 基因启动区多态性的群体遗传学资料。

[中图分类号] R5

[文献标识码] A

Correlation between GATA6 gene promoter region polymorphism and susceptibility and clinical features in patients with acute myocardial infarction

SUN Zhaoqing¹, PANG Shuchao², CUI Yinghua^{3,4}, YAN Bo^{2,4,5}

(1. Cheeoloo College of Medicine, Shandong University, Jinan, Shandong 250012, China; 2. Shandong Provincial Key Laboratory of Cardiac Disease Diagnosis and Treatment, 3. Division of Cardiology, 4. The Center for Molecular Genetics of Cardiovascular Diseases, 5. Shandong Provincial Sino-US Cooperation Research Center for Translational Medicine, Affiliated Hospital of Jining Medical University, Jining, Shandong 272029, China)

[KEY WORDS] acute myocardial infarction; gene promoter region; single nucleotide polymorphism; risk factors; linkage disequilibrium

[ABSTRACT] **Aim** To investigate the relationship between single nucleotide polymorphism (SNP) of GATA6 gene promoter region-493(rs144923558)G/A and -172(rs146748749)G/A 2 sites and acute myocardial infarction (AMI) and their associated risk factors. **Methods** A case-control study was performed to collect 328 AMI patients and 344 normal controls (NC). Data were analyzed by polymerase chain reaction-restriction fragment length polymorphism technique combined with sequence alignment after DNA sequencing. After the Hardy-Weinberg balance test, the χ^2 test was used for correlation analysis; Logistic regression was used to analyze the relationship between multiple risk factors and two SNP loci and AMI incidence; linkage disequilibrium and haploid analysis was performed using Haploview 4.2 software and

[收稿日期] 2019-01-26

[修回日期] 2019-04-25

[基金项目] 国家自然科学基金(81370271;81400291;81670341;81870279)

[作者简介] 孙兆庆, 硕士研究生, 研究方向为临床心血管病, E-mail 为 sunzqh01@163.com。通信作者闫波, 博士, 教授, 泰山学者, 博士研究生导师, 研究方向为心脏分子遗传学, E-mail 为 yanbo@mail.jnmc.edu.cn。

SHEsis website. **Results** Three genotypes were detected in the two SNP loci, including GG, GA and AA. The genotype distribution was consistent with Hardy-Weinberg equilibrium ($P>0.05$), and there was no significant difference between the AMI group and the NC group ($P>0.05$). Multivariate Logistic regression analysis: age, hypertension, smoking, LDLC and TG were independent risk factors for AMI ($P<0.05$), and HDLC was protective factors ($P<0.05$). Logistic regression analysis in three different genetic models of dominant, recessive and additive suggests that the two SNP loci are not associated with the onset of AMI. Linkage disequilibrium and haplotype analysis indicated that the two SNP loci were in the same linkage disequilibrium region ($D'=1.000$, $r^2=1.000$), and haplotypes GG and AA did not increase AMI susceptibility ($P>0.05$). **Conclusions** The GATA6 gene promoter-493 (rs144923558) G/A and -172 (rs146748749) G/A two SNPs are completely unbalanced, of which GG is the main haplotype. The two SNP loci and their haplotypes were not associated with the onset of AMI, but provided population genetics of the GATA6 gene promoter region polymorphism.

随着社会的老齡化,冠状动脉粥样硬化性心脏病简称冠心病(CHD)已经成为中国老年人最为常见的复杂的心血管疾病,其发病与遗传和后天环境因素紧密相关。急性心肌梗死(acute myocardial infarction, AMI)则是CHD当中最为严重的一种类型,具有发病急、病情危重和预后差等特点,也是世界上致死、致残的重要原因之一^[1]。目前,国内外研究多集中于后天环境因素方面,在AMI发病的遗传学机制方面的研究却相对较少。GATA结合蛋白6(GATA binding protein 6, GATA6)是位于人类18号染色体q11.1-q11.2的一小部分锌指转录因子的成员,是蛋白质编码基因,由7个外显子和6个内含子共同组成。先前有研究表明,GATA6基因参与人类早期胚胎发生过程,且在心脏前期发育中起重要作用,其基因突变可导致多种先天性心脏畸形,如二叶式主动脉瓣(bicuspid aortic valve, BAV)、室间隔缺损(ventricular septal defect, VSD)和圆锥动脉心脏缺陷等^[2-4]。另外,GATA6基因也在成人心肌细胞、动脉平滑肌细胞和实验动物的心肌细胞中均有检测到表达^[5-7]。然而,截至目前GATA6基因启动子区的SNPs与人类心肌梗死是否有关联,国内外仍未见有相关报告。因此本研究采用病例-对照的方法,选取中国北方汉族人群为研究对象,探讨GATA6基因启动子区-493(rs144923558)G/A、-172(rs146748749)G/A两个SNP位点和其它危险因素与AMI发病的相关性,并对该2个SNP位点进行连锁不平衡和单倍体型分析。

1 资料和方法

1.1 一般资料

选取2016年1月至2017年12月在济宁医学院附属医院心内科住院的AMI患者328例为AMI组,其中男236例,女92例,平均(63.70±12.21)岁

(28~94岁)。同期选取来本院健康体检者344例为对照组,其中男220例,女124例,平均(45.56±12.85)岁(19~81岁),两组均为汉族且没有血缘关系。AMI患者须结合病史、临床表现、心电图、生化指标、三维超声心动图等确诊,具体诊断标准依据2012年在欧洲心脏病学会(ESC)年会上制订的“心肌梗死通用定义”第三版^[8]。纳入对照组中的体检者为年龄相仿的、无CHD或AMI病史,且经心电图、超声心动图和心肌酶学等检查后无异常的健康人群。具体排除标准:(1)不符合纳入标准者;(2)合并有风湿性心脏病、心肌病、主动脉夹层等疾病者;(3)合并有急性或严重慢性感染性疾病、严重肝肾功能不全、肿瘤、自身免疫性疾病和家族性高胆固醇血症者;(4)不能配合者。研究协议经济宁医学院附属医院人文伦理委员会批准,并获得所有参与者知情同意。

1.2 提取样本基因组DNA

使用含有乙二胺四乙酸二钾(EDTA-K2)的抗凝试管,抽取研究对象清晨空腹时肘部静脉血3 mL。利用天津灏洋公司的人全血单个核细胞分离液,采用分层梯度离心法提取外周血单个核细胞。使用QIAGEN牌的DNA提取试剂盒提取出基因组DNA,暂存于-80℃冰箱中备用。

1.3 设计引物,扩增目的片段

从Genebank数据库中获取GATA6基因启动区生物基因组序列,选取GATA6基因转录起始点上游约2000个碱基进行研究分析,并结合dbSNP数据库中筛选出rs144923558 G/A和rs146748749 G/A 2个SNP位点。针对其位点所在的基因启动序列中的位置,使用软件Primer 5进行引物设计,经NCBI中Blast与UCSC网站对比后,选取特异性高引物序列(见表1),由上海生工生物有限公司合成。PCR反应条件:94℃预变性3 min,经(94℃ 30 s, 62℃ 30 s, 72℃ 50 s)×35个循环,72℃延伸5 min,4℃保温。

表 1. PCR 的引物序列表

Table 1. Primer sequence list of PCR

SNP 位点	引物位置	引物序列	产物 (bp)
-493 (rs144923558) G/A	上游	5'-CCGAAACCACCACGACCTGAG-3'	663
-172 (rs146748749) G/A	下游	5'-TGGGCTCCTGATTGGACTCACC-3'	

1.4 检测 GATA6 基因启动子区

将体外扩增得到的目的基因片段发往上海生工生物有限公司进行 Sanger 法测序。使用 DNAMAN 软件与野生型进行序列比对,并通过测序图谱核对,将样本中的两个 SNP 位点及对应基因型数据进行统计。

1.5 统计学分析

所有定量资料包括年龄、BMI (kg/m^2)、高密度脂蛋白胆固醇 (high density lipoprotein cholesterol, HDLC)、低密度脂蛋白胆固醇 (low density lipoprotein cholesterol, LDLC)、总胆固醇 (total cholesterol, TC)、甘油三酯 (triacylglycerol, TG) 等以 $\bar{x} \pm s$ 表示,其分析采用 t 检验。二分类变量包括性别、高血压史、糖尿病史和吸烟史等,Hardy-Weinberg 平衡检测、基因型和等位基因频率的分布采用 χ^2 检验进行分析,同时基因易感性风险以比值比 (OR) 和 95% 置信区间 (95% CI) 表示。利用多因素 Logistic 回归对多种危险因素进行分析,采用显性、隐性和加性 3 种遗传模型对 2 个 SNP 位点与 AMI 发病的关联性进行分析,同时利用多因素 Logistic 回归分析校正年龄、性别、BMI、高血压、糖尿病和吸烟等混杂因素。以上分析由 SPSS 22.0 统计软件完成, $P < 0.05$ 被视为有统计学意义。连锁不平衡分析采用 Haploview 4.2 软件,连锁不平衡系数 D' 及 r^2 界限定为 0.8,在线使用 SHEsis 网站进行单倍体构建和分析。

2 结果

2.1 两组一般临床资料的比较

AMI 组年龄、血压、血糖、吸烟和血脂水平 (TG

除外)均显著高于对照组 ($P < 0.05$,表 2),具有统计学意义,这与传统认知的冠心病高危因素相符合。其中 TG 两组间差异无显著性 ($P > 0.05$),在本研究中不排除 AMI 组中患者因服用调脂药物而受影响。

表 2. AMI 组与对照组临床资料的比较

Table 2. Comparison of clinical data between AMI group and NC group

临床资料	AMI 组 ($n=328$)	对照组 ($n=344$)	P
男/女 (例)	236/92	220/124	0.032
年龄 (岁)	63.70 \pm 12.21	45.56 \pm 12.85	0.000
BMI (kg/m^2)	24.80 \pm 3.74	25.47 \pm 3.46	0.024
高血压病史 [例 (%)]	147 (44.8)	84 (24.4)	0.000
糖尿病史 [例 (%)]	72 (22.0)	26 (7.6)	0.000
吸烟史 [例 (%)]	168 (51.2)	56 (16.3)	0.000
HDLC (mmol/L)	1.05 \pm 0.37	1.31 \pm 0.30	0.000
LDLC (mmol/L)	2.50 \pm 0.81	2.79 \pm 0.72	0.000
TG (mmol/L)	1.49 \pm 0.94	1.44 \pm 1.09	0.583
TC (mmol/L)	4.27 \pm 1.08	4.92 \pm 1.44	0.000

2.2 Hardy-Weinberg 平衡检测

将 AMI 组与对照组人群中 rs144923558 G/A 和 rs146748749 G/A 的基因型分布用 χ^2 检验后均符合 Hardy-Weinberg (HWE) 平衡 ($P > 0.05$,见表 3),表明两个 SNP 位点基因型分布在本研究的人群中达到了遗传平衡,具有较好的群体代表性。

2.3 基因型及等位基因频率分布

GATA6 基因启动区 rs144923558 G/A 和 rs146748749 G/A 的基因型均为 GG、GA 和 AA 基因型。其中,以上 2 个 SNP 位点在 AMI 和对照组中的

表 3. GATA6 基因启动区 2 个 SNP 在 AMI 组和对照组中基因型及等位基因频率分布

Table 3. Genotype and allele frequency distribution of two SNPs in GATA6 promoter region in AMI and NC groups

分组	基因型 [例 (%)]			HWE	P 值	等位基因 [例 (%)]		P 值	OR	95% CI
	AA	GA	GG			A	G			
AMI 组	1 (0.3)	16 (4.9)	311 (94.8)	$P=0.119 > 0.05$	0.580	18 (2.7)	638 (97.3)	0.885	1.050	0.542 ~ 2.036
对照组	0	18 (5.2)	326 (94.8)	$P=0.618 > 0.05$		18 (2.6)	670 (97.4)			

SNP 位点 rs144923558 和 rs146748749 的数据是一致的。HWE: Hardy-Weinberg 平衡, P 值均 > 0.05 , 群体代表性较好。

主要基因型均为 GG。2 个 SNP 位点基因型分布及等位基因频率分布在 AMI 组和对照组之间差异均无统计学意义 ($P>0.05$, 表 3)。

2.4 多种危险因素及 GATA6 基因启动区 2 个 SNP 位点与 AMI 发病的关联分析

以研究对象是否患有 AMI 作为因变量, 把性别、糖尿病、高血压和吸烟作为因子变量, 年龄、BMI、HDLc、LDLc、TG 和 TC 作为协变量进行多因素 Logistic 回归分析。结果显示: 增龄、高血压病、吸烟、LDLc 和 TG 升高是人群中 AMI 发病的独立危险因素 ($P<0.05$), HDLc 为保护因素 ($P<0.05$) (见表 4)。该研究结果提示健康的生活方式如: 防控高血压、戒烟和调血脂等是预防 AMI 发病的有效举措。

在加性、显性和隐性 3 种不同遗传模式下, 运用多项 Logistic 回归分析方法对两个 SNP 位点与 AMI 的发病进行单因素分析, 在校正性别、年龄、BMI、高血压、糖尿病和吸烟史后再次分析, 基因的易感性风险以比值比 (OR) 和 95% 置信区间 (95% CI) 表示。单因素分析结果为 2 个 SNP 位点的显性模型与 AMI 的发病无相关性 ($P=0.977>0.05$, OR=0.990, 95% CI=0.501~1.956)。在校正性别、年龄、BMI、高血压、糖尿病和吸烟史等因素后显性模型分析结果无统计学意义 ($P=0.533>0.05$, OR=1.366, 95% CI=0.512~3.643)。同时, 对 2 个 SNP 位点进行隐性和加性模型分析均未提示与 AMI 发病有相关性 (表 5)。

表 4. AMI 危险因素的多项 Logistic 回归分析

Table 4. Multiple Logistic regression analysis of AMI risk factors

因素	B	SE	Wald	P	OR	95% CI	
						下限	上限
年龄	0.118	0.011	109.204	0.000	1.125	1.101	1.151
BMI	-0.050	0.037	1.774	0.183	0.951	0.884	1.024
[性别=男]	-0.212	0.295	0.516	0.472	0.809	0.454	1.443
[糖尿病=否]	-0.298	0.335	0.791	0.374	0.743	0.385	1.431
[高血压=否]	-0.757	0.250	9.145	0.002	0.469	0.287	0.766
[吸烟=否]	-2.249	0.301	55.778	0.000	0.105	0.058	0.190
HDLc	-1.699	0.401	17.911	0.000	0.183	0.083	0.402
LDLc	0.980	0.383	6.539	0.011	2.664	1.257	5.645
TG	0.350	0.174	4.034	0.045	1.419	1.008	1.996
TC	-1.172	0.323	13.159	0.000	0.310	0.165	0.584

[性别=男]: 与女性比较; [糖尿病=否]: 与患糖尿病比较; [高血压=否]: 与患高血压比较; [吸烟=否]: 与吸烟比较。

表 5. GATA6 基因启动区 2 个 SNP 与 AMI 发病的 Logistic 回归分析

Table 5. Logistic regression analysis of the incidence of two SNPs and AMI in the promoter region of GATA6 gene

SNP	遗传模式	单因素			多因素校正后		
		Wald χ^2 值	P 值	OR(95% CI)	Wald χ^2 值	P 值	OR(95% CI)
rs144923558G/A	显性	0.001	0.977	0.990(0.501, 1.956)	0.388	0.533	1.366(0.512, 3.643)
或	隐性	-	-	-	-	-	-
rs146748749G/A	加性	-	-	-	-	-	-

显性: (GA+AA) VS GG; 隐性: AA VS (GA+GG); 加性: AA VS GG。多因素 Logistic 回归校正了性别、年龄、BMI、高血压、糖尿病和吸烟史。

2.5 遗传连锁不平衡和单倍体型分析

运用 Haploview 4.2 软件对 GATA6 基因启动区 2 个 SNP 位点(rs144923558 和 rs146748749) 进行连锁不平衡分析, 结果显示: $D' = 1 > 0.8$ 和 $r^2 = 1 > 0.8$, 证明这 2 个 SNP 位点处于同一完全连锁不平衡区域。根据连锁不平衡系数 D' 和 r^2 作图 (图 1),

图中灰色表示强连锁不平衡, 以 $D' > 0.8$ 和 $r^2 > 0.8$ 作为参考。在线使用 SHEsis 网站进行单倍体构建和分析, 发现以上 2 个 SNP 位点可产生 2 个单倍体型 GG 和 AA (其中频率 < 0.01 的在分析中被忽略)。经分析未提示 2 个单倍体型与 AMI 发病有相关性 ($P > 0.05$, 表 6)。

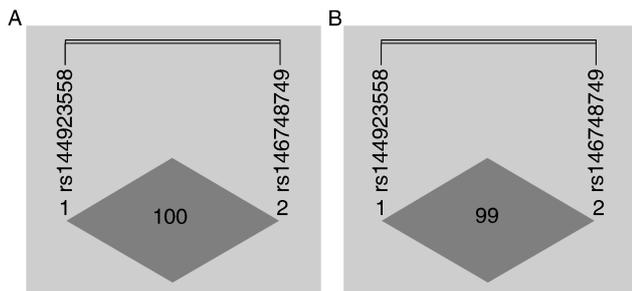


图 1. GATA6 基因启动区 2 个 SNP 位点 (rs144923558 和 rs146748749) 连锁不平衡图 A: 显示 D' 值 (D' = 1.000); B: 显示 r² 值 (r² = 1.000)。

Figure 1. Linkage disequilibrium map of two SNP loci in GATA6 promoter region (rs144923558 and rs146748749)

表 6. GATA6 基因启动区的单倍体型与 AMI 发病的关联性

Table 6. Association between haplotypes of GATA6 promoter region and the onset of AMI

编号	单倍体型		频率 (%)		χ^2	Fisher's P	OR(95% CI)
	rs144923558	rs146748749	AMI 组	对照组			
H1	G	G	97.3	97.4	0.021	0.885	1.010(0.516,1.977)
H2	A	A	2.7	2.6	0.021	0.885	0.952(0.491,1.847)

表 7. 不同区域人群 rs144923558 和 rs146748749 的基因型及等位基因频数分布比较

Table 7. Comparison of genotype and allele frequency distribution of rs144923558 and rs146748749 in different regions

区域	基因型比例 [例 (%)]			等位基因 [频数 (%)]	
	GG	GA	AA	G	A
EAS	463(91.9)	40(7.9)	1(0.2)	966(95.8)	42(4.2)
CHB	96(93.2)	7(6.8)	-	199(96.6)	7(3.4)
CHS	99(94.3)	5(4.8)	1(1.0)	203(96.7)	7(3.3)
JPT	90(86.5)	14(13.5)	-	194(93.3)	14(6.7)
AFR	576(87.1)	81(12.3)	4(0.6)	1233(93.3)	89(6.7)
AMR	327(94.2)	19(5.5)	1(0.3)	673(97.0)	21(3.0)
EUR	503(100.0)	-	-	1006(100.0)	-
SAS	489(100.0)	-	-	978(100.0)	-

SNP 位点 rs144923558 和 rs146748749 的数据是一致的。数据来源于 1000 Genomes Project Phase 3 于 2018 年 10 月发布的 94 版。EAS: 东亚, CHB: 中国北方汉族, CHS: 中国南方汉族, JPT: 日本东京, AFR: 非洲, AMR: 美国, EUR: 欧洲, SAS: 南亚。

3 讨论

众所周知,基因表达程序的失调已经影响到人类许多疾病,如心血管疾病、糖尿病、炎症和癌症等^[9-10]。AMI 属于 CHD 的一种严重类型且为多基因遗传疾病,在中国老年人中的发病率越来越高。此前有报道 CHD 传统危险因素达 300 多种,而与 CHD 相关的易感基因有 500 多个,从未来个体化治疗的层面出发,将 AMI 可能的潜在关联基因进行发

2.6 世界不同国家、地区间的 2 个 SNP 位点情况与本研究结果的比较

本研究虽以中国北方汉族人群为研究对象,而由于地域环境因素的影响,中国南方与北方的研究结果存在一定差异,如在中国南方人群中可检测到 AA 纯合子基因型,而在本研究的 AMI 组中发现 1 例,不排除有外来流动人员的可能。此外,由于不同国家和地区存在种族和地理位置的差异,也给研究数据带来较大不同,如东亚(EAS)、非洲(AFR)、美国(AMR)均可检测到 GG、GA 和 AA 基因型,而欧洲(EUR)、南亚(SAS)却只检测到 GG 基因型。这可能与样本量的大小和遗传异质性有关系(表 7)。

病风险预测的作用不亚于传统危险因素的预防。基因多态性是人类基因组在受到外界环境影响后的一种适应性表现,在疾病易感性和机体对疾病的抵御等方面有着重要作用。其中,SNP 是基因多态性中最常见的一种。目前有大量研究证实,AMI 的发生与遗传因素和血管内皮损伤有很大关系,许多全基因组关联研究已经确定了大量的遗传基因座,而单个或多个基因的多态性改变会对 AMI 的发病产生直接影响,这表明常见的遗传变异有助于 AMI

的发展^[11-13]。但由于不同种族、地域或同一种族不同区域的差异存在,导致人群中的 AMI 易感性有所不同^[14]。尽管在 AMI 发病过程中,低频率和罕见的遗传变异可能是遗传力缺失的部分原因,但这些遗传变异仅能解释病例总数的 10%^[15]。因此,有关 AMI 发生的遗传因素和潜在的分子遗传机制仍有很多不清楚。同时,随着现代分子生物学、遗传学知识及相关技术的不断成熟,人们对 AMI 的发病遗传学机理也有了更深的理解,所以研究基因的多态性和 AMI 发病相关性具有重要的临床价值。

据报道,人类 GATA6 基因在成人心脏、胰腺、卵巢、肺、肝脏、中枢神经系统、肾上腺和血管平滑肌细胞中均检测到表达,其中在心脏和胰腺发育中有着特别重要作用^[16]。此前国内外对 GATA6 基因的研究大多集中于心脏前期发育中因基因突变而引起的相关疾病,然而 GATA6 基因在成年人心肌及血管平滑肌细胞中的仍有表达,但一直未有基因表达程序失调后的相关疾病报道。本研究则以中国北方汉族人群为研究对象,对 GATA6 基因启动区中 2 个 SNP 位点(rs144923558G/A 和 rs146748749G/A)和其它危险因素与 AMI 的发病进行了探索性分析。对 2 个 SNP 位点的基因型分布、等位基因频率进行分析显示:GG 和 GA 基因型在 AMI 和对照组中均有发现,且 GG 为野生型,AA 纯和突变只在 AMI 组中找到 1 例;G 等位基因频数在 AMI 组中为 638 (97.3%),在对照组中为 670(97.4%),但无统计学差异($P>0.05$),未发现与 AMI 发病相关联的基因型和等位基因。把临床资料引入 Logistic 回归分析后提示:增龄、高血压病、吸烟、LDLC 和 TG 升高是人群中 AMI 发病的独立危险因素,而健康的生活方式则有助于减少或延缓 AMI 的发生。另外,在 3 种遗传模式下未发现与 AMI 发病有关的 SNP。在本研究中证实了此 2 个 SNP 位点处于同一完全连锁不平衡区域($D'=1>0.8$ 和 $r^2=1>0.8$),可产生 GG、AA 两种单倍型,其中 GG 为主要单倍型,与 AMI 发病进行相关分析得到 GG/AA 的 Fisher's P 值均 >0.05 ,结果提示该 2 个 SNP 位点均不增加 AMI 的易感性。

实际上,人类心血管疾病的发生通常是由内在多个基因相互作用并结合外在多种因素共同导致。GATA6 基因可参与调节体内许多关键转录因子和结构分子,并与它们相互协助而发挥作用。例如:GATA6 和 GATA5 等位基因的复合缺失导致双出口右心室和 VSD^[17];GATA6 和 TBX5 协同激活 ANF

基因在心脏发育中起作用^[18]等。然而像 CHD 及 AMI 这种多基因遗传疾病具有多种微效基因累加效应的特点,GATA6 基因是否与体内其它未知基因共同作用后才能通过干扰心脏基因调控网络,并与可控的传统危险因素(如高血压、糖尿病、吸烟和生活方式等)以及不可控制因素(年龄、性别和 CHD 家族史等)结合后共同促进 CHD 及 AMI 的发生发展仍未知。另外有研究证实,GATA6 基因突变是胰腺发育不全的主要原因^[19],也是部分新生儿及成人糖尿病的形成重要因素^[20-21]。从该角度出发,GATA6 基因启动区的 SNPs 虽在 AMI 人群中未显示出相关性,但不排除影响胰岛细胞水平后,以 DM 的形式促进了 CHD 及 AMI 的发生。此外,鉴于基因表达具有时间和空间的特异性,选择性表达也很复杂,其影响因素既有内因又有外因。因此,GATA6 基因是否促进成年人 CHD 或 AMI 的发展还有待于进一步研究验证。

本研究对 GATA6 基因启动子区 2 个 SNP 位点在中国北方汉族 AMI 人群中的易感性进行了初步探索,对其连锁不平衡及单倍型进行了相关分析,并证实了 2 个 SNP 位点为完全连锁不平衡,虽未证实该 2 个 SNP 位点及其组成的单倍型 GG/AA 与 AMI 发病的有相关性,但提供了 GATA6 基因启动区多态性的群体遗传学资料。当然,本研究也存在几个局限性。首先,在分析某个或某些基因 SNPs 对疾病的贡献时,最好在整体基因中随机选择 SNP。然而,为了探讨 GATA6 启动子区中聚集的 SNP 对 AMI 的影响,本研究只选择了其中 2 个 SNP 位点。因此,下一步需要做的是扩增 SNP 位点的研究数量以及增加 GATA6 基因其他区域的 SNPs 分析。其次,结合现实中不同地域存在基因型分布等差异情况,后续研究还待于扩大样本来源即进行其他族群的确认研究,来对 GATA6 基因启动子区的 SNPs 以及序列中的其它变异与 AMI 发病之间的关系作进一步功能验证。

[参考文献]

- [1] Murray CJ, Barber RM, Foreman KJ, et al. Global, regional, and national disability-adjusted life years (DALYs) for 306 diseases and injuries and healthy life expectancy (HALE) for 188 countries, 1990 ~ 2013: quantifying the epidemiological transition[J]. *Lancet*, 2015, 386(10009): 2145-2191.
- [2] Xu YJ, Di RM, Qiao Q, et al. GATA6 loss-of-function mutation contributes to congenital bicuspid aortic valve[J].

- Gene, 2018, 663: 115-120.
- [3] Li C, Li X, Pang S, et al. Novel and functional DNA sequence variants within the GATA6 gene promoter in ventricular septal defects[J]. *Int J Mol Sci*, 2014, 15(7): 12677-12687.
- [4] Zhang E, Hong N, Chen S, et al. Targeted sequencing identifies novel GATA6 variants in a large cohort of patients with conotruncal heart defects [J]. *Gene*, 2018, 641: 341-348.
- [5] Suzuki E, Evans T, Lowry J, et al. The human GATA-6 gene: structure, chromosomal location, and regulation of expression by tissue-specific and mitogen-responsive signals [J]. *Genomics*, 1996, 38(3): 283-290.
- [6] Morrissey EE, Ip HS, Lu MM, et al. GATA-6: a zinc finger transcription factor that is expressed in multiple cell lineages derived from lateral mesoderm [J]. *Dev Biol*, 1996, 177(1): 309-322.
- [7] Pikkariainen S, Tokola H, Kerkela R, et al. GATA transcription factors in the developing and adult heart[J]. *Cardiovasc Res*, 2004, 63(2): 196-207.
- [8] Thygesen K, Alpert JS, Jaffe AS, et al. Third universal definition of myocardial infarction[J]. *Circulation*, 2012, 126(16): 2020-2035.
- [9] Lee TI, Young RA. Transcriptional regulation and its misregulation in disease [J]. *Cell*, 2013, 152(6): 1237-1251.
- [10] Gong M, Simaite D, Kuhnen P, et al. Two novel GATA6 mutations cause childhood-onset diabetes mellitus, pancreas malformation and congenital heart disease[J]. *Horm Res Paediatr*, 2013, 79(4): 250-256.
- [11] Assimes TL, Roberts R. Genetics: implications for prevention and management of coronary artery disease[J]. *J Am Coll Cardiol*, 2016, 68(25): 2797-2818.
- [12] Björkegren JLM, Kovacic JC, Dudley JT, et al. Genome-wide significant loci: how important are they? Systems genetics to understand heritability of coronary artery disease and other common complex disorders[J]. *J Am Coll Cardiol*, 2015, 65(8): 830-845.
- [13] McPherson R, Tybjaerg H. A genetics of coronary artery disease[J]. *Circ Res*, 2016, 118(4): 564-578.
- [14] Dandona S, Roberts R. The role of genetic risk factors in coronary artery disease[J]. *Curr Cardiol Rep*, 2014, 16(5): 479.
- [15] Lettre G. Rare and low-frequency variants in human common diseases and other complex traits[J]. *J Med Genet*, 2014, 51(11): 705-714.
- [16] Tiyaboonchai A, Cardenas-Diaz FL, Ying L, et al. GATA6 plays an important role in the induction of human definitive endoderm, development of the pancreas, and functionality of pancreatic beta cells[J]. *Stem Cell Reports*, 2017, 8(3): 589-604.
- [17] Laforest B, Nemer M. GATA5 interacts with GATA4 and GATA6 in outflow tract development [J]. *Dev Biol*, 2011, 358(2): 368-378.
- [18] Maitra M, Schluterman MK, Nichols HA, et al. Interaction of Gata4 and Gata6 with Tbx5 is critical for normal cardiac development [J]. *Dev Biol*, 2009, 326(2): 368-377.
- [19] Shi ZD, Lee K, Yang D, et al. Genome editing in hPSCs reveals GATA6 haploinsufficiency and a genetic interaction with GATA4 in human pancreatic development [J]. *Cell Stem Cell*, 2017, 20(5): 675-688.
- [20] Yau D, De Franco E, Flanagan SE, et al. Case report: maternal mosaicism resulting in inheritance of a novel GATA6 mutation causing pancreatic agenesis and neonatal diabetes mellitus[J]. *Diagn Pathol*, 2017, 12(1): 1.
- [21] Villamayor L, Rodríguez-Seguel E, Araujo R, et al. GATA6 controls insulin biosynthesis and secretion in adult β -cells[J]. *Diabetes*, 2018, 67(3): 448-460.

(此文编辑 朱雯霞)