DOI:10.16016/j.1000-5404.201908110

Gilbert 综合征 UGT1A1 致病变异的连锁关系研究

李露锋,邓国宏,毛 青 400038 重庆陆军军医大学(第三军医大学)第一附属医院感染科

[摘要] **目的** 明确 Gilbert 综合征患者(Gilbert syndrome,GS)UDP-葡萄糖醛酸基转移酶 1A1 基因(UDP-glucuronyl transferase 1A1 gene,UGT1A1)致病变异的连锁关系,进一步明确 GS 的遗传机制。**方法** 对 46 例 GS 患者与 80 名健康对照者 UGT1A1 基因的增强子 PBREM、启动子 TATA 盒、PE、DE 及外显子进行直接测序,把所测得的 SNP 进行连锁不平衡分析、并构建单体型和单体基因型进行分析。**结果** 共测到 c.-3279T>G、c.-1352C>A、c.-40_-39insTA、c.211G>A、c.686C>A、c.1456T>G 等 6 个 SNP 位点,对该 6 个 SNP 位点进行连锁不平衡分析表明 c.-3279T>G 和 c.-1352C>A 的 D'值和 r^2 均等于 1,呈完全的连锁不平衡;进一步构建单体型表明 c.-40_-39insTA 和 c.-3279T>G 呈连锁关系、c.211G>A 和 c.-1352C>A 呈连锁关系,他们所构成的单体基因型在病例组中的频率高达 0.717,而在对照组中的频率仅为 0.005,两者具有显著的统计学差异,$P<0.000$。**结论** Gilbert 综合征的发生与来源于不同同源染色体的 c.-40_-39insTA 和 c.211G>A 的两两变异相关,而 c.-3279T>G 和 c.-1352C>A 与它们的各自连锁发挥了剂量效应作用。

[关键词] Gilbert 综合征;UGT1A1;连锁;单体型;单体基因型

[中图法分类号] R363.21;R394.3;R596.1 [文献标志码] A

Linkage analysis of pathogenic UGT1A1 mutations in Gilbert syndrome: a case-control study of 46 cases

LI Lufeng, DENG Guohong, MAO Qing Department of Infectious Diseases, First Affiliated Hospital, Army Medical University (Third Military Medical university), Chongqing, 400038, China

[**Abstract**] **Objective** To clarify the linkage of pathogenic mutations in UDP-glucuronyl transferase 1A1 (UGT1A1) gene with Gilbert syndrome (GS) and explore the genetic mechanism of GS. **Methods** The genomic DNA was extracted from the peripheral blood samples of 46 unrelated GS patients and 80 healthy control subjects. The enhancer PBREM, the promoter TATA box, PE, DE and the coding regions of UGT1A1 gene were amplified with PCR, and the amplified DNA fragments were sequenced and analyzed. **Results** Six single nucleotide polymorphisms (SNPs) were identified in UGT1A1 gene of GS patients, namely c.-3279T>G, c.-1352C>A, c.-40_-39insTA, c.211G>A, c.686C>A, and c.1456T>G. The D' and r^2 were both 1 for c.-40_-39insTA and c.-3279T>G, demonstrating a complete linkage disequilibrium. Haplotype construction showed that c.-40_-39insTA was linked to c.-3279T>G, and c.211G>A was linked to c.-1352C>A; the frequency of the diplotypes generated by these haplotypes reached 0.717 in GS, as comparted to 0.005 in the control subjects ($P<0.000$). **Conclusion** The occurrence of GS is associated with the variations of c.-40_-39insTA and c.211G>A from different homologous chromosomes, and c.-3279T>G and c.-1352C>A both have a dose effect in their respective linkage.

[**Key words**] Gilbert syndrome; UDP-glucuronyl transferase 1A1; linkage; haplotype; diplotypes

Supported by the Major Innovation Research Project of Southwest Hospital (SWH2016ZDCX3002). Corresponding author: MAO Qing, E-mail: qingmao@tmmu.edu.cn

[基金项目] 西南医院重大创新课题(SWH2016ZDCX3002)

[通信作者] 毛 青,E-mail:qingmao@tmmu.edu.cn

Gilbert 综合征(Gilbert syndrome,GS)是遗传性非结合型胆红素血症中的最常见类型,其发病率达 10%左右,以间断或持续的轻中度非溶血性非结合型胆红素升高为主要临床表现(OMIM:#143500)。GS 多数呈良性结果,但严重者可导致核黄疸甚至与精神疾患相关[1];且与其变异相关的胆红素代谢相关基因 UDP-葡糖醛酸基转移酶 1A1 基因(UDP-glucuronyl transferase 1A1 gene,UGT1A1)同时也是最重要的编码药物 II 相代谢酶的基因之一,因而 GS 与多种药物代谢及毒副作用相关,可引起化疗相关腹泻、黄疸、肝衰竭[2-5]。因此,对 Gilbert 综合征患者及其相关基因变异的精准识别及遗传机制研究尤为重要。

UGT1A1 基因编码的 UGT1A1 酶是肝脏催化脂溶性葡萄糖醛酸化为水溶性的胆红素的唯一酶,在既往研究中:一方面,通过对普通人群的全基因组研究表明该基因的 c.-40_-39insTA 和 c.211G>A 等变异与胆红素水平升高相关[6];而病例-对照研究也充分证实了在病例组中 c.-40_-39insTA 以及 c.211G>A 等致病变异与疾病的相关性,功能研究也充分验证了这些变异可以导致 UGT1A1 酶的转录及催化活性的显著下降[7,8],但并未对 c.-40_-39insTA 和 c.211G>A 的连锁关系得出明确结论,由此也未能充分解释包括 c.-3279T>G 和 c.-1352C>A 等高频变异在 GS 中的作用,并对它们是否为致病变异存在争议。

本研究对 46 例 GS 患者与 80 名健康对照者 UGT1A1 基因的增强子 PBREM、启动子 TATA 盒、PE、DE 及外显子进行直接测序,并通过对所测得的 SNP 进行连锁不平衡分析、构建单体型和单体基因型等研究阐明了 c.-40_-39insTA 和 c.211G>A 分别与完全不连锁的 c.-3279T>G 和 c.-1352C>A 连锁形成不同的单体型,阐明了来自不同同源染色体的 c.-40_-39insTA 和 c.211G>A 的两两变异是 GS 的主要遗传机制和变异 Marker,而 c.-3279T>G 和 c.-1352C>A 与它们的各自连锁则进一步发挥了剂量效应作用。

1 资料与方法

1.1 病例

所纳入的 46 例无血缘关系的 GS 患者,均为汉族;其中男性 34 例,女性 12 例,平均年龄为 32.6±11.57 岁,总胆红素水平的平均值为 43.01±11.65 μmol/L,以非结合型胆红素水平升高为主(详见表 1)。

1.1.1 诊断标准及排除标准

诊断标准:①总胆红素>1.5 mg/dL(25.65 μmol/L),<6 mg/dL(102.6μmol/L),且以 UCB 为主;②ALT、AST 正常,以排除肝脏炎症;③红细胞计数(red blood cell,RBC)、血红蛋白(hemoglobin,Hgb)及网织红细胞计数(reticulocyte count,RT)正常,以排除潜在的溶血可能。

表 1 GS 患者基本情况及胆红素水平等指标

	指标	值
例数(男:女)		46(34:12)
年龄(岁)		32.6±11.57
胆红素	TB/μmol·L^{-1}	43.01±11.65
	UCB/μmol·L^{-1}	33.80±10.18
	CB/μmol·L^{-1}	9.27±3.19
	TB/DB	0.22±0.058
转氨酶	ALT/U·L^{-1}	23.1±20.5
	AST/U·L^{-1}	20.1±18.6
	ALP/U·L^{-1}	70.4±20.7
	GGT/U·L^{-1}	23.0±31.2
蛋白	TP/g·L^{-1}	73.1±6.9
	ALB/g·L^{-1}	46.1±5.2
	GLO/g·L^{-1}	25.9±4.9
血脂	TG/mmol·L^{-1}	2.2±1.1
	Tch/mmol·L^{-1}	4.5±1.9
肾功	Cr/μmol·L^{-1}	40.3±20.9(mg/dl)
	BUN/mmol·L^{-1}	6.9±2.8(mg/dl)
空腹血糖	Glu	5.3±1.8(mg/dl)
肝脏超声	腹部超声提示肝大、脾大或胆道系统疾病等异常者	
病史及药物史	按纳入及排除标准进行,首次检查发现胆红素升高前 90 d 之内任何的用药史	

GS:Gilbert 综合征;TB:总胆红素;UCB:非结合型胆红素;CB:结合型胆红素;ALT:谷丙转氨酶;AST:谷草转氨酶;ALP:碱性磷酸酶;GGT:谷氨酰转肽酶;TB:总蛋白;ALB:白蛋白;GLO:球蛋白;TG:甘油三酯;Tch:胆固醇;Cr:肌酐;BUN:尿素氮;Glu:血糖

排除标准:①合并其他疾病;②有饮酒史,且超过 5 年,酒精量男性≥40 g/d,女性≥20 g/d,或 2 周内有大量饮酒史,折合酒精量>80 g/d;③首次检查发现胆红素升高前 90 d 之内任何的用药史;④腹部超声提示肝大、脾大或胆道系统疾病等异常者。

1.1.2 病例筛选流程

根据诊断标准及排除标准,病例筛选按以下流程进行:1)对于转氨酶正常的以 UCB 为主的胆红素血症患者,详细询问病史并行体格检查,排除有慢性肝病史及药物服用史及合并高血压、糖尿病等其他疾病的患者,并再次复查肝功,仍提示以 UCB 为主的胆红素血症,ALT、AST 等转氨酶仍正常;2)进一步行血常规及网织红细胞计数等检查,排除血红蛋白及网织红细胞计数异常的患者;3)再进一步行血清学及腹部彩超等检查:排除病毒表面标志物阳性、自身抗体阳性及肝脏、胆道形态异常的患者;4)对部分患者还行溶血试验、血清铜、铜蓝蛋白、异常血红蛋

白电泳等检查进一步排除相关疾病。

在本研究中,共383例患者进入初筛,所有患者均为在第三军医大学西南医院感染科门诊或住院部就诊患者。通过询问病史、查体等进行初筛,所有患者都存在持续或间断的非结合型胆红素血症至少3年以上,且无ALT及AST升高等肝脏炎症表现。除尿黄、眼黄、肤黄等黄疸表现外,无阳性体征及不适主诉。再进一步行血常规、网织红细胞计数、甲肝抗体(HAV-Ab)、乙肝表面抗原(HBsAg)、丙肝抗体(HCV-Ab)、戊肝抗体(HEV-Ab)、自身抗体以及腹部超声等检查后,有289名患者被排除:273例患者检测到乙肝表面抗原阳性、5例患者结合型胆红素大于非结合型胆红素、3例患者HCV-IgM阳性、1例患者HAV-IgM阳性、1例患者为急性乙肝、5例患者病史不明、1例患者正在接受干扰素治疗。在余下的79例胆红素水平符合GS诊断标准的患者中,有33例患者被排除:18例患者未测网织红细胞计数、8例患者网织红细胞计数超过1.5%、2例患者在随访过程中出现转氨酶升高,1例患者存在贫血、4例患者在随访过程中胆红素呈下降趋势。

1.1.3 对照 80例健康对照组成员从我科工作人员及研究生中招募,包括44名男性和36名女性,均为汉族,平均年龄36.7±11.4岁。入组前,至少每两年进行一次健康体检,体检项目包括血常规、肝肾功、乙型肝炎表明抗原等血清学检测,胸部X光检测,腹部超声等检查,这些检测检查结果均提示正常。入组时,再次检测包括ALT、AST、TB、UCB、CB等指标的肝功能检测,均提示正常。

1.2 UGT1A1基因突变分析

1.2.1 引物设计合成 UGT1A1基因位于2号染色体3区7带,包含6个外显子,除启动子TATA盒外[-38到-23 bp,A(TA)6TAA],尚存在启动子DE(-1 362 bp到-1 220 bp,distal element)和PE(-113到-70 bp,proximal element),另外,在其上游还存在一个与苯巴比妥相关的增强子(-3 483 bp到-3 194 bp)[9-12]。本次扩增片段为UGT1A1基因的6个外显子及所有功能元件,但并不包括羧基端的非编码区(引物序列详见表4)。扩增引物用software Primer 3.0(Whitehead Institute, Cambridge, MA)辅助设计。UGT1A1基因的序列号为:GenBank accession number NT_005120.16。引物位置避开SNP位点,并选取发夹结构形成少的引物,以避免引物二聚体的产生。对所选取的引物用UCSC浏览器进一步检测扩增产物是否为目的片段(详见表2)。本次一共扩增9个片段,包括了UGT1A1基因的增强子,启动子区TATA盒、DE、PE,外显子1、2、3、4、5a及5b的编码区。在上述引物中,引物Cgt及CDE引用自既往文献。引物设计完成后,在北京华大基因合成。

表2 引物序列及其对应的靶序列片段长度

引物名称	引物方向	序列(5′→3′)	片段大小/bp	目的片段
Cgt	正向	CACCTCCTCCTTATTCTCTT	400	增强子
	反向	CTCATTCCTCCTCTCTAGCC		
CDE	正向	CTCTAAGCACATCCCCAAGTA	525	DE
	反向	TAAGCAAGTTTCCATCCTTCA		
CPE	正向	AAGTGAACTCCCTGCTACCTT	244	PE
	反向	TCAACAGTATCTTCCCAGCAT		
TATA-E1-1	正向	AGCTCCACCTTCTTTATCTCTGAA	578	TATA盒和外显子1上半段
	反向	TGAGCTCCTTGTTGTGCAGT		
E1-2	正向	GATTCTTTCCTGCAGCGTGT	649	外显子1下半段
	反向	GCATGTAAAAGTCCCACTCCA		
E2	正向	CAAACACGCATGCCTTTAATC	500	外显子2
	反向	CAGGGAAAAGCCAAATCTAAGG		
E3	正向	AAGTTGCCAGTCCTCAGAAGC	546	外显子3
	反向	GTTACTCACATGCCCTTGCAG		
E4	正向	CTGCAAGGGCATGTGAGTAAC	632	外显子4
	反向	GAGGCAGAACTGCTTGAACG		
E5b	正向	TCAAGCTCACTGGTAATAGGC	483	外显子5b编码区
	反向	CTCCAACTTCTGGGCTCAAG		
E5a	正向	CAGGTTTCCTTTCCCAAGTTT	555	外显子5a编码区
	反向	AAGCACTCTGGGGCTGATTA		

1.2.2 基因组 DNA 制备及 PCR 扩增 采用 Wizard Genomic DNA Purification kit(Promega,USA)从 5 mL 新鲜全血中抽提 DNA。按鼎国 DNA *Taq* 聚合酶(Beijing Dingguo Changsheng Biotech CO. LTD,China)说明书配制成 50 μL PCR 反应体系,扩增条件如下:95 ℃ 2 min 后,以 94 ℃ 30 s、58 ℃ 30 s、72 ℃ 40 s 进行 35 个循环,再以 72 ℃延长 7 min。

PCR 产物经 2%琼脂糖凝胶电泳鉴定后送北京华大基因纯化特异性片段并使用原 PCR 扩增引物测序。

1.2.3 SNP 鉴定及基因分型 用 PolyPhred,program(http://www.genome.washington.edu)进行 SNP 的鉴定,先通过 Consed 程序编写拼接序列,然后使用 FTP 网络资源,在 Windows 环境下,读取拼接峰图,进行比对。除根据程序提示外,对每一个峰再次进行核查。对于噪音大,背景杂的样品,重新进行 PCR 扩增,并采用切胶纯化后再进行测序。

1.2.4 H-W 检验、LD 分析、单体型构建 用 haploview.lnk(Cambridge,MA 02141,USA)进行 LD 分析,先制备输入文件,输入文件包括测序结果文件及 SNP 位点物理距离文件,均为 txt 格式。该软件可直接输出每个 SNP 的 H-W 检验结果,并以点阵图的形式展示 LD 的结果,可根据需要展示配对 r2 值或 D'值等,并使用 Four Famete 原则和 Solid Spine of LD 法定义单体型块。但 haploview 只能展示单体型频率及 tagSNP 的结果。对于每个个体的单体型及双体型,则使用基于 Bayesian 统计方法的 PHASE 软件来构建,其中模糊不清的单体型可根据家系结果选择,PHASE 软件在 DOS 环境下运行,构建单体型时,选取在一个单体型块内的 SNP 位点来构建。

1.2.5 统计学分析 使用 SPSS PASW Statistics(version 18.0)进行等位基因频率、单体型频率、双体型频率的统计分析、比较,采用 χ^2 检验,根据期望值,样本量大小选取不同的双侧检验结果(样本量>40、期望值>5:Pearson Chi-Squear;样本量>40、期望值<5:Continuity Correction;样本量<40 或期望值<1:Fisher's Exat Test)。

2 结果

2.1 UGT1A1 基因分型结果

在本研究中,在 46 例无血缘关系的 GS 患者和 80 个健康对照者中,除同义突变外,共检测到 6 个变异类型(详见表 3),它们分别位于增强子,启动子 DE、TATA 盒,外显子 1、4 及 5a,而启动子 PE、外显子 2、3 及 5b 未测到变异。这些变异引起的蛋白质改变、各变异等位的等位基因频率及与对照组的比较结果详见表 3。

表 3 UGT1A1 测序结果、变异等位频率及比较结果

变异名称	位置氨基酸改变蛋白改变	GS(N=46)纯合子杂合子等位基因频率	对照(N=80)纯合子杂合子等位基因频率	PGS 与对照	dbSNP
UGT1A1*60	增强子	14	13	0.048	rs4124874
	c.-3279T>G	23	42		
		0.554	0.425		
UGT1A1*112	启动子 DE	9	25	0.048	rs3755319
	c.-1352C>A	23	42		
		0.446	0.575		
UGT1A1*28	TATA 盒	9	1	0.000	rs3064744
	c.-40_-39insTA	19	21		
		0.402	0.144		
UGT1A1*6	外显子 1	6	2	0.007	rs4148323
	c.211G>A	17	23		
	p.G71R	0.315	0.169		
UGT1A1*27	外显子 1	1	0	0.261	rs35350960
	c.686C>A	2	2		
	p.P229Q	0.043	0.013		
UGT1A1*7	外显子 5a	1	0	0.033	rs34993780
	c.1456T>G	2	0		
	P.Y486D	0.043	0.000		

位于 TATA 盒的插入变异 c.-40_-39insTA 变异,使得野生型的 TATA 盒从 6 个 TA 重复[A(TA)6TAA]变异为 7 个 TA 重复[A(TA)7TAA],该变异在 GS 组中的等位基因频率明显高于对照组(0.402 和 0.144;P=6.237E-6)。位于外显子 1 的 c.211G>A 变异使 UGT1A1 基因的第 71 个密码子从 GGA 变异为 AGA,导致 UGT1A1 酶氨基酸序列的第 71 个氨基酸从甘氨酸变异为精氨酸,该变异在 GS 组中的等位基因频率也明显高于对照组(0.315 和 0.169;P=0.007)。位于外显子 5a 的 c.1456T>G 变异使 UGT1A1 基因第 486 个密码子从 TAC 变异为 GAC,导致 UGT1A1 酶第 486 个氨基酸从酪氨酸变异为天冬氨酸,该变异的等位基因频率,GS 组高于对照组(0.043 和 0.000;P=0.033)。位于外显子 1 的 c.686C>A 变异使 UGT1A1 基因的的 229 个密码子从 CCG 变异为 CAG,导致 UGT1A1 酶氨基酸序列的第 229 个氨基酸从脯氨酸变异为谷氨酰胺,该变异在 GS 组和对照组中均可测到,其等位基因频率 GS 组与对照组无统计学差异(0.043 和 0.013;P=0.261)。而位于增强子的 c.-3279T>G 变异和位于启动子 DE 的 c.-1352C>A 变异则在 GS 组、对照组中均可测到,且其等位基因频率在各组中均高于 0.300,该两个变异的等位基因频率在 GS 组与对照组中比较的结果,P 值都为 0.048,略小于 0.05。

2.2 Hardy-Weinberg 平衡检验和 LD 分析结果

所有 SNP 均符合 Hardy-Weinberg 平衡,P 值均大于 0.05,相应 P 值如下:c.-3279T>G (P=1.000),c.-1352C>A (P=1.000),c.-40_-39insTA (P=0.104),c.211G>A (P=0.714),c.686C>A (P=0.110),c.1456T>G (P=0.074)。LD 分析显示各个 SNP 之间存在强 LD 关系,所有 SNP 位于一个单体型块内。c.-3279T>G 变异和 c.-1352C>A 变异的配对 D'值及 r2 值均等于 1,该两个 SNP 呈完全 LD。除 c.-3279T>G 和 c.-40_-39insTA、c.-1352C>A 和 c.-40_-39insTA 间的配对 D'值分别为 0.96、0.96 外,其余各位点间的配对 D'值均为 1。

2.3 单个个体单体型构建结果及分析

用所测到的 6 个 SNP 位点构建单体型,共得到 7 个单体型(详见表 4)。其中,存在完全连锁不平衡关系的两个 SNP c.-3279T>G 和 c.-1352C>A 的两个变异等位 c.-1379G 和 c.-1352A 位于不同的同源染色体上,单体型 H1 和 H2 除分别含这两个变异等位中的一个变异等位外,不含其他变异等位,其他单体型分别由其他的 1 个或 2 个变异等位与这两个存在完全不连锁关系并位于不同的同源染色体上的变异等位(c.-3279G 或 c.-1352A)相连形成:c.211A、c.1456G 变异等位与 c.-1352A 变异等位分别相连形成单体型 H5、H7;c.-40_-39insTA、c.1091T 变异等位与 c.-3279G 变异等位分别相连形成单体型 H3 和 H6;c.686A 变异等位和 c.-40_-39ins 变异等位与 c.-3279G 变异等位相连形成单体型 H4。

除 c.-3279G 变异等位和 c.-1352A 变异等位外,c.-40_-39insTA 变异等位与 c.211A 变异等位和 c.1456G 变异等位也总是存在于不同的同源染色体上,但它们的相应 SNP 之间并不像 SNP c.-3279T>G 和 c.-1352C>A 那样呈完全连锁不平衡的关系。

在这些单体型中,只含有 c.-3279G 变异等位的单体型 H1 和只含有 c.-1352A 变异等位的单体型 H2 在对照组中占有明显的优势,均高于 GS 组(0.282 和 0.098,P=3.160×10^{-4};0.381 和 0.087,P=4.557×10^{-7});而含有 c.-40_-39insTA 变异等位的单体型 H3 和含有 c.211A 变异等位的单体型 H5 则在 GS 组中占有明显的优势,均高于对照组(0.359 和 0.131,P=2.271×10^{-5};0.315 和 0.169,P=0.007)。

表 4 单体型构建结果及其在 GS 组和对照组中的频率与比较结果

单体型	位点										GS 个数(频率) n=46	对照个数(频率) n=80	PGS 与对照
	-3 345	-3 279	-1 352	-40_-39	211	686	715	1 091	1 253	1 456			
H1	C	G	C		G	C	C	C	T	T	9(0.098)	45(0.282)	0.000
H2	C	T	A		G	C	C	C	T	T	8(0.087)	61(0.381)	0.000
H3	C	G	C	insTA	G	C	C	C	T	T	33(0.359)	21(0.131)	0.000
H4	C	G	C	insTA	G	A	C	C	T	T	4(0.043)	2(0.013)	0.261
H5	C	T	A		A	C	C	C	T	T	29(0.315)	27(0.169)	0.007
H6	C	G	C		G	C	C	T	T	T	5(0.054)	1(0.006)	0.047
H7	C	T	A		G	C	C	C	T	G	4(0.043)	0(0.000)	0.033

注:下划线等位基因所示为变异等位

表 5 双体型构建结果及在 GS 组和对照组中的频率与比较结果

双体型	变异等位	GS(个数/频率)	对照(个数/频率)	*P*
H1/H2	[c.-3279G]/[c.-1352A]	3/0.065	26/0.325	0.001
H2/H2	[c.-1352A]/[c.-1352A]	1/0.022	13/0.163	0.015
H1/H3	[c.-3279G]/[c.-3279G;c.-40_-39insTA]	2/0.043	6/0.075	0.750
H1/H4	[c.-3279G]/[c.-3279G;c.-40_-39insTA;c.686A]	1/0.021	1/0.013	1.000
H1/H5	[c.-3279G]/[c.-1352A;c.211A]	2/0.043	3/0.038	1.000
H1/H6	[c.-3279G]/[c.-3279G;c.-40_-39insTA;c.1456G]	1/0.022	1/0.013	1.000
H2/H3	[c.-1352A]/[c.-3279G;c.-40_-39insTA]	3/0.065	12/0.150	0.157
H3/H3	[c.-3279G;c.-40_-39insTA]/[c.-3279G;c.-40_-39insTA]	7/0.152	1/0.013	0.007
H3/H4	[c.-3279G;c.-40_-39insTA]/[c.-3279G;c.-40_-39insTA;c.686A]	1/0.022	0/0.000	0.778
H3/H5	[c.-3279G;c.-40_-39insTA]/[c.-1352A;c.211A]	11/0.239	1/0.013	0.000
H3/H6	[c.-3279G;c.-40_-39insTA]/[c.-3279G;c.1091T]	1/0.022	0/0.000	0.778
H3/H7	[c.-3279G;c.-40_-39insTA]/[c.-1352A;c.1456G]	1/0.022	0/0.000	0.778
H4/H4	[c.-3279G;c.-40_-39insTA;c.686A]/[c.-3279G;c.-40_-39insTA;c.686A]	1/0.022	0/0.000	0.778
H5/H5	[c.-1352A;c.211A]/[c.-1352A;c.211A]	6/0.130	2/0.025	0.050
H5/H6	[c.-1352A;c.211A]/[c.-3279G;c.1091T]	3/0.273	0/0.000	0.000
H5/H7	[c.-1352A;c.211A]/[c.-1352A;c.1456G]	1/0.022	0/0.000	0.778
H7/H7	[c.-1352A;c.1456G]/[c.-1352A;c.1456G]	1/0.022	0/0.000	0.778

2.4 单个个体单体基因型(双体型)结果分析

在病例组及对照组中,一共存在 17 种单体基因型(双体型)(详见表 5)。其中,由单体型 H1 和 H2 组成的双体型在对照组中占有明显优势:双体型 H1/H1 只存在于对照组中;双体型 H1/H2 和 H2/H2 的频率高于 GS 组(0.325 和 0.065,$P=0.001$;0.163 和 0.022,$P=0.015$)双体型 H3/H3、H3/H5 和 H5/H6 在 GS 组中占有明显优势,其频率远远高于对照组(0.152 和 0.013,$P=0.007$;0.239 和 0.013,$P=1.147\times10^{-4}$;0.273 和 0.000,$P=1.184\times10^{-4}$);双体型 H5/H5 的频率在 GS 组中也高于对照组(0.130 和 0.025,$P=0.05$)。

3 讨论

c.-3279T>G 变异位于 UGT1A1 基因的增强子上,UGT1A1 基因的增强子又称苯巴比妥反应元件(phenobarbital-responsive enhancer module, PBREM),位于 UGT1A1 基因的上游(-3 499 到-3 210)[11]。含 c.-3279G 等位的增强子较含 c.-3279T 等位的增强子而言,其转录活性下降到 62%[13],可导致转录活性下降 18%~30% 的被认为是白种人的 GS 致病变异的 c.-40_-39insTA 变异等位与 c.-3279G 变异等位连锁[14-15],构成单体型 H3。但 c.-3279G 等位基因频率是否在 GS 组中占优势及可否被鉴定为 GS 的致病变异,不同的研究存在争议:很多个案报道和病例研究在患者中都测到了 c.-3279G 变异等位,首都医科大学附属北京佑安医院梁晨、郑素军等人在 43 例包括 GS 及 Crigler-Najjar 综合征 II 型的患者中发现,c.3279G 等位普遍存在于 GS 组中高达 35%[16],但在陈志威等人[17]所做的 UGT1A1 基因与成人高胆红素血症的风险与定量研究中,胆红素水平在正常范围内波动的人群 c.3279G 等位的基因频率仍高达 29.5。在本研究中,c.-3279G 等位的等位基因频率在 GS 组和对照组并无明显的统计学差异(0.554 和 0.425,$P=0.048$),且只包含 c.-3279G 变异等位的单体型 H2 在对照组中占有明显的优势,其频率远远高于 GS 组(0.282 和 0.098,$P=3.16\times10^{-4}$);在病例组中占明显优势的是除 c.3279G 变异等位外,同时存在 c.-40_-39insTATA 变异的单体型 H4(0.359 和 0.131,$P=2.271\times10^{-5}$);因此,通过个基因位点的连锁关系分析,本研究阐明了造成 c.-3279G 变异等位在病例组中集中的原因是 c.-40_-39insTATA 变异和该变异的连锁关系,c.-40_-39insTATA 变异才是 GS 发病的致病变异和 Marker 变异基因。

在既往研究中,对于 GS 的遗传模式存在显性和隐性等争议,白种人普遍存在 c.-40_-39insTATA 纯合变异,但亚裔人群除 c.-40_-39insTATA 纯合变异、c.211A 等纯合变异外,尚存在 c. -40_-39insTA 和 c.211G>A 复合杂合变异,且 c. -40_-39insTA 和 c.211G>A 的杂合变异也广泛分布在正常人群中。在本研究中,最重要的是 c.-1352C>A 变异的鉴定以及该 SNP 与 SNP c.-3279T>G 的连锁关系的鉴定。不知何种原因,既往研究较少涉及启动子 DE,而启动子 DE 被鉴定的当时,及在该启动子上鉴定了 c.-1352C>A 变异,含 c.-1352A 等位启动子的 UGT1A1 基因的转录活性为含 c.-1352C 等位启动子的 UGT1A1 基因转录活性的 85%[10]。本研究的结果显示,c.-1352C>A 变异与上

述 c.-3279T>G 变异之间存在完全连锁不平衡的关系,且 c.-1352A 变异等位和 c.-3279G 变异等位总是位于不同的同源染色体上,在同一条染色体上,他们非此即彼。且在亚裔人群 GS 患者中除与白种人相同的 c.-40_-39insTA 变异等位外,最常见变异是可使胆红素结合的结构域肽段发生构象改变并使 UGT1A1 酶的活性下降 40%[17] 的 c.211G>A 变异,c.211A 与 c.-1352A 构成单体型 H5。同样,与 c.-3279G 类似的是,只包含 c.-1352A 等位基因的单体型 H2 在对照组中也同样占有明显优势,其频率亦远远高于 GS 组(0.381 和 0.087,$P=4.557\times10^{-7}$);在病例组中占明显优势的是单体型 H5(0.315 和 0.169,$P=0.007$)。因此,本研究通过纳入既往病例对照研究中被忽略的 c.1352A 重要变异等位,阐明了在 GS 中最常见的变异 Maker c.-40_-39insTA 和 c.211G>A 来源于不同的亲体,说明 c.-40_-39insTA 和 c.211G>A 复合杂合变异的两个变异来源于不同的亲本,大多数患者表现为"隐性"遗传模式,进一步解释了 c.-40_-39insTA 和 c.211G>A 的杂合变异在正常人群中广泛分布的原因。

在本课题组既往的家系研究中也充分证实了通过连锁关系分析所得出的各基因位点的传递关系,c.-40_-39insTA 和 c.211G>A 总是来源于不同的亲本,且总是分别与 c.-3279G 和 c.-1352A 各自相连的连锁关系[18]。同时,本研究进一步分析了千人基因组(发布版本 20110521,下载日期为 2019 年 8 月 9 日)中的数据,SNP c.-3279T>G 和 c.-1352C>A 之间仍然存在完全连锁不平衡的关系,且与 c.-3279G 和 c.-1352A 各自相连的连锁关系,各等位之间的连锁关系也与本研究结果一致。

本研究进一步分析了 GS 患者的单体基因型(双体型),在 GS 组及对照组中一共存在 17 种。进一步分析表明只包括 c.3279G 和 c.-1352A 的 H1 和 H2 所构成的单体型(H1/H1、H1/H2、H2/H2)在对照组中的比例为 0.537,而在 GS 组中的比例仅为 0.087;但 c.-40_-39insTA 和 c.211A 与之连锁后分别组成的单体型 H3、H4、H5、H6 和 H7 构成的双体型(H3/H3、H3/H4、H3/H5、H3/H6、H3/H7、H4/H4、H5/H5、H5/H6、H5/H7、H7/H7)的频率在对照组中仅为 0.05,而在 GS 组中高达 0.717。

综上所述,尽管既往研究及本研究在 GS 患者中测出了众多相关的 UGT1A1 变异,但通过完整的包括被忽略的重要启动子 DE 的纳入和分析,本研究进一步阐明了这些变异之间的连锁关系,证明了 GS 的主要致病变异 Marker 为来源于不同同源染色体的 c.-40_-39insTA 变异等位和 c.211A 变异等位,他们的两两变异是 GS 发生的主要遗传机制;c.3279G 和 c.-1352A 仅可能发挥剂量效应作用,他们的单独变异不足以导致 GS 的发生,排除了次要或无关变异位点,本研究的实施对于 GS 的基因诊断的广泛开展和筛查以及临床转化、兼备诊断等具有重要意义。

参考文献:

[1] BENTIVEGNA A, SANTAMBROGIO J, CLERICI M. UGT1A1 mutations and psychoses: towards understanding the relationship with unconjugated bilirubin[J]. CNS Spectr, 2019: 1-3. DOI: 10.1017/S1092852919001251.

[2] HA V H, JUPP J, TSANG R Y. Oncology drug dosing in gilbert syndrome associated with UGT1A1: A summary of the literature[J]. Pharmacotherapy, 2017, 37(8): 956-972. DOI: 10.1002/phar.1946.

[3] PANAGOPOULOS P, MALTEZOS E, HATZAKIS A, *et al*. Hyperbilirubinemia in atazanavir treated HIV-infected patients: the impact of the UGT1A1 * 28 allele[J]. Pharmgenom Pers Med, 2017, 10: 205-208. DOI: 10.2147/PGPM.S107152.

[4] MEHBOOB H, TAHIR I M, IQBAL T, *et al*. Effect of UDP-glucuronosyltransferase (UGT) 1A polymorphism (rs8330 and rs10929303) on glucuronidation status of acetaminophen[J]. Dose Response, 2017, 15(3): 1559325817 723731. DOI: 10.1177/1559325817723731.

[5] LEGER P, CHIRWA S, NWOGU J N, *et al*. Race/ethnicity difference in the pharmacogenetics of bilirubin-related atazanavir discontinuation[J]. Pharmacogenet Genomics, 2018, 28(1): 1-6. DOI: 10.1097/FPC.0000000000000316.

[6] COLTELL O, ASENSIO E M, SORLÍ J V, *et al*. Genome-wide association study (GWAS) on bilirubin concentrations in subjects with metabolic syndrome: sex-specific GWAS analysis and gene-diet interactions in a Mediterranean population[J]. Nutrients, 2019, 11(1): E90. DOI: 10.3390/nu11010090.

[7] HUANG M J, CHEN Y C, HUANG Y Y, *et al*. Effect of UDP-glucuronosyltransferase 1A1 activity on risk for developing Gilbert's syndrome[J]. Kaohsiung J Med Sci, 2019, 35(7): 432-439. DOI: 10.1002/kjm2.12077.

[8] KING D, ARMSTRONG M J. Overview of Gilbert's syndrome[J]. Drug Ther Bull, 2019, 57(2): 27-31. DOI: 10.1136/dtb.2018.000028.

[9] RITTER J K, CHEN F, SHEEN Y Y, *et al*. A novel complex locus UGT1 encodes human bilirubin, phenol, and other UDP-glucuronosyltransferase isozymes with identical carboxyl termini[J]. J Biol Chem, 1992, 267(5): 3257-3261.

[10] UEYAMA H. Analysis of the promoter of human bilirubin UDP-glucuronosyltransferase gene (UGT1 * 1) in relevance to Gilbert's syndrome[J]. Hepatol Res, 1997, 9(2/3):

152-163. DOI: 10.1016/s1386-6346(97)00097-1.

[11] SUGATANI J, KOJIMA H, UEDA A, *et al*. The phenobarbital response enhancer module in the human bilirubin UDP-glucuronosyltransferase UGT1A1 gene and regulation by the nuclear receptor CAR[J]. Hepatology, 2001, 33(5): 1232-1238. DOI: 10.1053/jhep.2001.24172.

[12] LÉVESQUE E, GIRARD H, JOURNAULT K, *et al*. Regulation of the UGT1A1 bilirubin-conjugating pathway: role of a new splicing event at the UGT1A locus[J]. Hepatology, 2007, 45(1): 128-138. DOI: 10.1002/hep.21464.

[13] SUGATANI J, YAMAKAWA K, YOSHINARI K, *et al*. Identification of a defect in the UGT1A1 gene promoter and its association with hyperbilirubinemia[J]. Biochem Biophys Res Commun, 2002, 292(2): 492-497. DOI: 10.1006/bbrc.2002.6683.

[14] BOSMA P J, CHOWDHURY J R, BAKKER C, *et al*. The genetic basis of the reduced expression of bilirubin UDP-glucuronosyltransferase 1 in Gilbert's syndrome[J]. N Engl J Med, 1995, 333(18): 1171-1175. DOI: 10.1056/NEJM199511023331802.

[15] PASTERNAK A L, CREWS K R, CAUDLE K E, *et al*. The impact of the UGT1A1 * 60 allele on bilirubin serum concentrations[J]. Pharmacogenomics, 2017, 18(1): 5-16. DOI: 10.2217/pgs-2016-0135.

[16] SUN L, LI M, ZHANG L, *et al*. Differences in UGT1A1 gene mutations and pathological liver changes between Chinese patients with Gilbert syndrome and Crigler-Najjar syndrome type II[J]. Medicine (Baltimore), 2017, 96(45): e8620. DOI: 10.1097/MD.0000000000008620.

[17] 陈志威. UGT1A1 基因与成人高胆红素血症:风险研究与定量比较[D]. 上海交通大学,2012.
CHEN Z W. UGT1A1 gene and adult hyperbilirubinemia: risk study and quantitative comparison[D]. Shanghai Jiaotong University,2012.

[18] YAMAMOTO K, SATO H, FUJIYAMA Y, *et al*. Contribution of two missense mutations (G71R and Y486D) of the bilirubin UDP glycosyltransferase (UGT1A1) gene to phenotypes of Gilbert's syndrome and Crigler-Najjar syndrome type II[J]. Biochim Biophys Acta, 1998, 1406(3): 267-273. DOI: 10.1016/s0925-4439(98)00013-1.

[19] 李露锋, 邓国宏, 谭文婷, 等. 8 个 Gilbert 综合征家系 UGT1A1 变异谱调查及分析[J]. 第三军医大学学报, 2014, 36(23): 2408-2412. DOI: 10.16016/j.1000-5404.2014.23.019.
LI L F, DENG G H, TAN W T, *et al*. Investigation and analyses of UGT1A1 variant spectrum for 8 Gilbert syndrome families[J]. J Third Mil Med Univ, 2014, 36(23): 2408-2412. DOI:10.16016/j.1000-5404.2014.23.019.

(**收稿**:2019-08-15;**修回**:2019-09-22)

(编辑 邓强庭)