文章编号:1000-5404(2006)21-2119-04

论著

利用毛细管电泳技术分析云南回族人群7个STR基因座的遗传多态性

吴恒娟,文富强,应斌武,刘代顺,邱 容,陈亚娟,徐 丹 (四川大学华西医院呼吸内科,成都 610041)

提 要:目的 对云南回族群体的D3S1358/THO1/D21S11/D18S51/D5S818/D13S317/D7S820等7个STR基因座的遗传多态性进行群体遗传学研究。**方法** 利用荧光标记复合扩增及毛细管电泳自动荧光检测的方法,对204例无关个体的7个STR基因座进行检测。**结果** 获得7个基因座的各基因座等位基因的分布频率和基因型频率,结果均符合Hardy-Weinberg平衡。计算了各基因座的杂合度(He)、个人识别能力(PD)、偶合率(PM)、非父排除率(PE)和多态性信息总量(PIC)等群体遗传学数据。**结论** 利用毛细管电泳技术可对这7个STR基因座进行很好的检测,这些基因座多态性丰富、灵敏度高,可用于人类遗传分析及法医学中的亲子鉴定和个人识别。

关键词: 毛细管电泳;短串联重复序列(STR);复合扩增;群体遗传学
中图法分类号: R394.5;R446.9 **文献标识码:** A

Genetic polymorphism of 7 STR loci in Hui population of Yunnan province by capillary electrophoresis

GUO Heng-juan, WEN Fu-qiang, YING Bin-wu, LIU Dai-shun, QIU Rong, CHEN Ya-juan, XU Dan (Department of Respiratory Diseases, West China Hospital, Sichuan University, Chengdu 610041, China)

Abstract: Objective To investigate the allele frequency of seven short tandem repeats (STR) loci: D3S1358/THO1/D21S11/D18S51/D5S818/D13S317/D7S820 in Hui population of Yunnan province. **Methods** Totally 204 unrelated individuals from Hui population in Yunnan province were inspected for 7 STR polymorphism D3S1358/THO1/D21S11/D18S51/D5S818/D13S317/D7S820 by means of capillary electrophoresis and fluorescence multiple amplification. **Results** The genotype distributions of each locus were in accordance with Hardy-Weinberg equilibrium. Some population genetic data, including the measured values of heterozygosity, the polymorphism information content (PIC), the discrimination power (PD), the power of exclusion (PE) were calculated. **Conclusion** Our results show that 7 STR loci are valuable STR loci genetic markers and useful in human genetics and forensic medicine.

Key words: capillary electrophoresis; short tandem repeat; multiplex polymerase chain reaction; population genetics

短串联重复序列(short tandem repeats, STRs)具有较高的多态性,存在于人类基因组DNA中,是由2~7 bp重复单元构成的DNA序列,其重复单元的数目变化表现出长度多态性。STR作为一类重要的遗传标记之一,已被广泛用于人类进化研究、法医个体识别和亲子鉴定及疾病基因作图等。

本研究利用四色荧光标记复合扩增及毛细管电泳自动荧光检测的方法,对204例云南回族无关个体的D3S1358、THO1、D21S11、D18S51、D5S818、D13S317、D7S820等7个STR基因座的遗传多态性进行研究,获得了云南回族群体7个STR基因座的群体遗传学数据,为人类疾病基因作图、法医学群体基因数据库的建立、人类进化研究提供重要的数据基础。

基金项目: 国家自然科学基金杰出青年基金资助项目(30425007),国家自然科学基金资助项目(30370627);美国纽约CMB基金资助项目,中国博士后基金资助项目(2005038268)
Supported by the National Natural Science Foundation for Young Scholars (30425007), the National Natural Science Foundation of China (30370627), the New York CMB Fund of USA and the Post-Doctoral Foundation of China (2005038268)

作者简介: 吴恒娟(1979-),女,山东省济南市人,硕士研究生,主要从事呼吸疾病遗传学方面的研究。电话:(028)85446090,E-mail:guohengjuanzq@126.com
通讯作者: 文富强,电话:(028)85164087,E-mail:wenfuqiang@126.com

收稿日期: 2006-01-17;**修回日期:** 2006-06-24

1 材料与方法

1.1 材料

样本种类为EDTA-抗凝血样，采自于云南回族的204例无血缘关系个体。

1.2 方法

1.2.1 DNA的提取 用Chelex-100方法提取DNA[1]。

1.2.2 复合扩增 扩增在10 μl反应体系中进行。反应体系包括缓冲液4 μl，引物2 μl，*Taq*酶0.2 μl(5 U/μl)，模板DNA 1 μl(DNA 1～2.5 ng)，用无菌去离子水补足体系。热循环参数为95 ℃ 11 min；94 ℃ 1 min、59 ℃ 1 min、72 ℃ 1 min，28个循环；60 ℃ 45 min。

1.2.3 扩增产物的检测 12 μl去离子甲酰胺，0.2 μl Genescan2500 Rox内标，1 μl扩增产物。95 ℃变性3 min后速置冰水中保温至少3 min，47 cm×50 μm毛细管，POP24液态胶，恒温60 ℃，电压15 kV。ABI3100型基因分析仪中自动进行电泳分离并检测。用Genescan3.1和Genotyper2.0分析软件，以标准阶梯等位基因为标准进行基因分型。

1.3 统计学分析

记录各无关个体7个STR位点的基因型，计算基因频率，并对基因型分布作χ^2检验[2]。据此，利用Powerstats软件分别计算云南回族人群STR位点各基因座的预期杂合度(He)、偶合率(PM)、非父排除率(PE)、个人识别率(PD)、多态性信息总量(PIC)[3-6]。

2 结果

2.1 云南回族人群7个STR基因座上的等位基因和基因型分布

云南回族群体在D3S1358、THO1、D21S11、D18S51、D5S818、D13S317、D7S820等7个基因座上分别检出等位基因7、7、14、15、9、7和7个。群体分布经χ^2检验符合Hardy-Weinberg平衡，见表1。D3S1358、THO1、D21S11、D18S51、D5S818、D13S317和D7S820在群体中分别检测基因型数目为18、20、42、52、24、22、21。D3S1358的各等位基因频率范围为0.2%(等位基因13)～37%(等位基因15)，基因型频率最高是25.9%(基因型15－16)。THO1的各等位基因频率范围为0.2%(等位基因12)～47.8%(等位基因9)，基因型频率最高是30.4%(基因型7－9)。D21S11的各等位基因频率范围为0.2%(等位基因35.2/30.2)～27.2%(等位基因29)，基因型频率最高是14.7%(基因型29－30)。D18S51的各等位基因频率范围为0.2%(等位基因25)～23.3%(等位基因14)，基因型频率最高是9.3%(基因型13－14)。D5S818的各等位基因频率范围为0.2%(等位基因15)～27.7%(等位基因11)，基因型频率最高是16.2%(基因型11－12)。D13S317的各等位基因频率范围为1.2%(等位基因14)～32.6%(等位基因8)，基因型频率最高是13.2%(基因型8－11)。D7S820的各等位基因频率范围为0.5%(等位基因14)～33.8%(等位基因11)，基因型频率最高是16.7%(基因型11－12)(见表1～8)。在表2～7中详细罗列了7个基因座的基因型分布状况。

表1 7个STR基因座在云南回族群体中等位基因及基因频率

	基因座频率(%)						
等位基因	D3S1358	THO1	D21S11	D18S51	D5S818	D13S317	D7S820
6		10.0					
7		27.9			2.7		
8		6.4			0.5	32.6	16.7
9		47.8			6.9	12.7	6.9
9.3		3.4					
10		4.2			22.3	13.0	14.5
11				0.7	27.7	22.5	33.8
12		0.2		5.6	25.0	14.7	24.3
13	0.2			15.0	13.5	3.2	3.4
14	4.4			23.3	1.2	1.2	0.5
15	37.0			19.6	0.2		
16	32.6			12.3			
17	19.6			6.4			
18	5.1			6.1			
19	1.0			3.9			
20				1.7			
21				1.7			
22				1.7			
23				0.7			
24				1.0			
25				0.2			
28			4.2				
28.2			0.5				
29			27.2				
30			27.0				
30.2			0.7				
31			11.5				
31.2			6.9				
32			3.4				
32.2			12.3				
32.3			0.2				
33			0.7				
33.2			4.7				
34.2			0.5				
35.2			0.2				

表2 D3S1358基因座的等位基因频率和基因型分布

等位基因	13	14	15	16	17	18	19	频率(%)
13								0.2
14	1	1						4.4
15		5	29					37.0
16		6	53	20				32.6
17		3	27	28	7			19.6
18		1	8	4	7			5.1
19				2	1	1		1.0

表3 THO1基因座的等位基因频率和基因型分布

等位基因	6	7	8	9	9.3	10	12	频率(%)
6	3							10.0
7	13	11						27.9
8	5	10	1					6.4
9	15	62	8	47				47.8
9.3	1	6	1	5				3.4
10	1	1		10	1	2		4.2
12				1				0.2

表 4　D21S11 基因座的等位基因频率和基因型分布

等位基因	28	28.2	29	30	30.2	31	31.2	32	32.2	32.3	33	33.2	34.2	35.2	频率(%)
28	1														4.2
28.2	1														0.5
29	3	1	16												27.2
30	5		30	11											27.0
30.2			1	1											0.7
31	2		18	9		1									11.5
31.2			5	13	1	3	1								6.9
32			1	6		2		1							3.4
32.2	3		12	16		6	4	2	2						12.3
32.3			1												0.2
33			2						1						0.7
33.2	1		3	8		5		1	1						4.7
34.2			1						1						0.5
35.2			1												0.2

表 5　D18S51 基因座的等位基因频率和基因型分布

等位基因	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	频率(%)
11																0.7
12																5.6
13	1	2	5													15.0
14		8	19	10												23.3
15	1	3	10	14	10											19.6
16	1	2	9	15	9	3										12.3
17			2	7	6	4	1									6.4
18			3	4	7		3	3								6.1
19		3	1	3	4	2	1		1							3.9
20		1		2	2			1								1.7
21			1	1	3		1			1						1.7
22			3		1	2										1.7
23		2		1												0.7
24		2						1				1				1.0
25				1												0.2

表 6　D5S818 基因座的等位基因频率和基因型分布

等位基因	7	8	9	10	11	12	13	14	15	频率(%)
7										2.7
8										0.5
9	1									6.9
10	3	2	5	8						22.3
11	3		13	26	10					27.7
12	2		8	19	33	12				25.0
13	2		1	17	16	15	2			13.5
14				3	1	1				1.2
15					1					0.2

表 7　D13S317 基因座的等位基因频率和基因型分布

等位基因	8	9	10	11	12	13	14	频率(%)
8	19							32.6
9	22	2						12.7
10	22	7	1					13.0
11	27	11	14	8				22.5
12	15	7	5	19	7			14.7
13	6	1	2	4				3.2
14	3		1	1				1.2

表 8　D7S820 基因座的等位基因频率和基因型分布

等位基因	8	9	10	11	12	13	14	频率(%)
8	4							16.7
9	6	1						6.9
10	7	4	6					14.5
11	28	7	23	21				33.8
12	15	9	12	34	11			24.3
13	3		1	4	6			3.4
14	1				1			0.5

2.2　群体遗传学指标

利用软件 Powerstats 分别计算云南回族人群在这 7 个 STR 位点上的杂合度、多态信息量、个体识别率、非父排除率和偶合率，结果见表 9。

表 9　7 个 STR 基因座在云南回族群体的遗传学数据

遗传指标		D3S1358	THO1	D21S11	D18S51	D5S818	D13S317	D7S820
偶合率(PM)		0.140	0.167	0.061	0.040	0.084	0.079	0.090
个体识别率(PD)		0.860	0.833	0.939	0.960	0.916	0.921	0.910
多态信息含量(PIC)		0.660	0.630	0.790	0.840	0.750	0.760	0.740
非父排除率(PE)		0.461	0.407	0.672	0.672	0.681	0.634	0.579
杂合度(He)		72.1	68.6	83.8	83.8	84.3	81.9	78.9
HWE								
df = 1	χ^2	0.027 248 9	0.062 686 4	0.612 535 2	0.685 69	3.568 393 7	1.081 511 4	0.269 446 2
	P	0.868 887 1	0.802 299 3	0.433 834 9	0.407 634	0.058 889 3	0.298 36	0.603 703 5

3　讨论

由于 STR 在人类基因组中出现的数目和频率不同,从而表现多态性。其本质是人类在进化过程中由于各种原因引起染色体 DNA 中核苷酸排列顺序发生了改变,即产生个体间 DNA 片段长度或 DNA 序列差异。STR 基因位点,因其大多具有多态性,且核心序列重复次数和个体间差异,在基因传递过程中一般遵循孟德尔显性遗传规律[7]。由于 DNA 片段短,造成错配和优势扩增的可能性小,提高了 PCR 扩增的成功率和灵敏度,而且标本用量少(仅仅需要 0.1 ng 模板 DNA),扩增结果稳定,重复性好,并可进行自动化分型。因此可以做为法医学鉴定、疾病遗传学分析、人类学研究以及基因身份证项目的研究手段。

用毛细管电泳和荧光标记复合扩增的方法对 204 例无关云南回族人群进行 D3S1358、THO1、D21S11、D18S51、D5S818、D13S317、D7S820 等 7 个 STR 基因座遗传多态性研究，从中分别显示 7、7、14、15、9、7 和 7 个等位基因,扩增片段长度均在 100 ~ 350 bp 之间,便于检测。所有基因座的基因型分布已经 χ^2 检验，结果均符合 Hardy-Weinberg 平衡(表 1)。每个基因座的预期杂合度(He)、偶合率(PM)、非父排除率(PE)、多态性信息总量(PIC)、个人识别率(PD)等的详细数据见表 9。由表 9 看出：7 个基因座的 PD 值都在 0.833 ~

0.960之间，累计PD值达0.999 999 999 9；PM在0.040～0.167之间；PE在0.407～0.681之间；PIC值在0.630～0.840之间。按Gill等建立的标准，PD≥0.9，H≥0.7的基因座是高鉴别能力遗传体系，在本研究中的7个基因座基本属于高鉴别力、高杂合度、高信息量的基因座（H≥0.7、PD≥0.9、PIC≥0.750）。从表2～7中，我们可以看出7个STR基因座的基因型分布状况良好，PIC值在0.630～0.840之间，说明这7个STR基因座在云南回族群体中的群体分布很好，多态信息含量高。其中D18S51基因座发现的等位基因数最多，D18S51基因座的H、PD、PE及PIC的指标最高，这7个STR基因座均适合法医学、人类进化学、基因作图、群体遗传学的研究与应用。

参考文献：

[1] WALSH P S, Metzger D A, Higuchi R. Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material [J]. Biotechniques, 1991, 10(4): 506－513.

[2] HOU Y, PRINZ M, STAAK M. Comparison of different tests for deviation from Hardy-Weinberg equilibrium of AMPFLP population data [M]//BAR W, FIORI A, ROSSI U. Advances in Forensic Haemogenetics 5. Berlin: Springer Verlag, 1994: 511－514.

[3] TEREBA A. Tools for analysis of population statistics [J]. Profiles In DNA, 1999, 2(3): 14－16.

[4] BOTSTEIN D, WHITE R L, SKOLNICK M, *et al*. Construction of a genetic linkage map in man using restriction fragment length polymorphisms[J]. Am J Hum Genet, 1980, 32(3): 314－331.

[5] BRENNER C, MORRIS J W. Paternity index calculations in single locus hypervariable DNA probes: validation and other studies [M]// Proceedings for The International Symposium on Human Identification 1989. Wisconsin: Promega Corporation, 1990: 21－53.

[6] YING B W, WEI Y G, SUN X M, *et al*. STR data for the AmpFlSTR profiler plus from western China[J]. J Forensic Sci, 2005, 50(3): 716－717.

[7] ALFORD R L, HAMMOND H A, COTO I, *et al*. Rapid and efficient resolution of parentage by amplification of short tandem repeats[J]. Am J Hum Genet, 1994, 55(1): 190－195.

（编辑 陈聪连）

（上接2114页）

表1 正常大鼠血液常规的测量值（$\bar{x}\pm s$）

指标	测量值
白细胞总数 WBC（$\times10^9$/L）	36.2±14.3
小细胞百分数 W-SCR（%）	73.3±10.2
大细胞百分数 W-LCR（%）	26.1±8.9
小细胞绝对值 W-SCC（$\times10^9$/L）	25.8±9.1
大细胞绝对值 W-LCC（$\times10^9$/L）	10.4±7.2
红细胞数 RBC（$\times10^{12}$/L）	7.8±0.5
血红蛋白量 HGB（g/L）	138.0±9.4
红细胞压积 Het（%）	45.4±3.3
红细胞平均容积 MCV（fl）	58.2±2.2
红细胞平均血红蛋白量 MCH（pg）	17.7±0.6
红细胞平均血红蛋白浓度 MCHC（g/L）	303.8±10.4
红细胞体积分布宽度 RDW-SD（fl）	23.2±2.8
血小板数 PLT（$\times10^9$/L）	898.7±361.1
血小板体积分布宽度 PDW（fl）	8.4±0.8
血小板平均容积 MPV（fl）	7.6±0.3
大型血小板比率 P-LCR（%）	10.3±2.4

表2 正常大鼠血浆生物化学测量值（$\bar{x}\pm s$）

指标	测量值
总蛋白 TP（g/L）	61.5±5.0
白蛋白 Alb（g/L）	30.0±3.9
球蛋白 G（g/L）	31.4±3.4
白球比值 A/G	0.93－0.97
谷丙转氨酶 ALT（IU/L）	81.8±26.1
谷草转氨酶 AST（IU/L）	232.4±99.0
总胆红素 TBIL（μmol/L）	1.8±0.9
血尿素 UN（mmol/L）	7.5±1.0
血肌酐 Cr（μmol/L）	58.3±14.9
血糖 Glu（mmol/L）	11.3±2.9

3 讨论

一般临床观察和病理解剖均未见异常，说明所有大鼠都处于正常状况。血液常规2次测量值非常接近，统计学上无显著性差异，表明本测量方法重现性好。

从表1、2中的数据与传统手工操作法[1,2]相比较：①部分指标相近似，如RBC、Het、HGB、PLT、TP、Alb、G、A/G及UN等，进一步证实本方法测得数据准确可靠。②WBC偏高，白细胞分类只有小细胞群（淋巴细胞）和大细胞群（中性粒细胞），没有中间细胞群（单核细胞、嗜酸及嗜碱粒细胞），这三分群方法与手工操作五分法不好比较。③血浆酶由于测量方法和使用单位不同，数据无法比较。④增加新的指标有：W-SCC、W-LCC、MCV、MCH、MCHC、RDW-SD、PDW、MPV、P-LCR、TBIL及Cr等，为实验研究提供更多的观察指标。从②、③、④的情况能说明建立正常大鼠血液常规和生物化学指标新的参考值是必要。

（本研究得到我院检验科张波博士和检验员的技术帮助，在此深表谢意！）

关键词：全自动医学检验分析仪；血液常规；生物化学指标；大鼠

中图法分类号：R446.1 文献标识码：B

参考文献：

[1] 朱忠勇. 临床血液学实验室诊断进展[J]. 中华检验医学杂志, 2003, 26(12): 729－730.

[2] 王文娟、王佩佩、陈保德，等. LH750血液分析仪临床应用评价[J]. 中华检验医学杂志, 2005, 28(3): 319－321.

[3] 施新猷. 现代医学实验动物学[M]. 北京：人民军医出版社, 2000: 50－89.

[4] 石 岩，梅世昌. 医学动物实验实用手册[M]. 北京：中国农业出版社, 2002: 233－244.

（编辑 陈聪连）