

广西仫佬族 9 个 STR 的遗传多态性研究

邓琼英¹, 徐林¹, 周丽宁¹, 李松峰¹

(1. 广西医科大学人体解剖学教研室, 南宁 530021)

摘要: 本文采用 PCR-STR 及基因分型技术, 研究广西仫佬族 183 例无关个体 9 个 STR 位点的遗传多态性分布, 建立仫佬族群体的遗传学数据库。经统计分析, 在 9 个 STR 位点共检出 70 种等位基因, 其频率分布在 0.0027~0.5301 之间; 207 种基因型, 其频率分布在 0.0055~0.3388 之间; 平均杂合度为 0.7298, 平均多态信息总量为 0.7016, 累积个体识别力达 0.99999999, 累积非父排除率达 0.999098。与不同民族比较结果显示: 广西仫佬族与广西苗、回族及云南、北方各民族之间绝大多数基因座存在显著差异, 而与广西壮族和湖南汉族之间绝大多数基因座均无差异。以上数据可为群体遗传学、法医学及人类学等研究提供重要的资料。

关键词: 仫佬族; PCR-STR; 遗传多态性

中图法分类号: Q987

文献标识码: A

文章编号: 1000-3193 (2005) 03-0215-06

生物界乃至人类个体间之所以千差万别, 其本质是在 DNA 上的差异。这些差异是由于不编码蛋白质的区域和没有重要调节功能的区域发生了中立突变, 这些中立突变构成的 DNA 变异称为 DNA 多样性^[1]。在 70 年代以前, 人们对于这些差异的认识首先始于其表达产物性状上的差异, 如抗原、蛋白质和酶等基因产物是过去人们对人类遗传多态性研究的主要对象。70 年代后, 人们逐渐认识到基因产物多态性源于结构基因的多态性, 自此后, 国内外学者对 DNA 多态性便展开了经久不衰的研究。随着分子生物学技术的发展, 对 DNA 多态性的研究经历了三个发展阶段, 先后有限制性片段长度多态性 (Restriction Fragment Length Polymorphism, RFLP)、可变数目串联重复序列 (Variable number of tandem repeats, VNTR)、短串联重复序列 (Short Tandem Repeats, STR) 和单核苷酸多态性 (Single Nucleotide Polymorphism, SNPs) 等四种遗传标记物被分离和应用^[2]。其中 STR 因其在人类基因组中分布广泛, 片段小, 易扩增, 所需检材微量及其高度杂合性、高度多态性而受到各国学者的高度重视。此外, 采用 PCR-STR 技术及全自动测序仪检测 STR 方法简便, 自动化, 判型准确, 成为目前国内外在群体遗传学、法医学、基因定位和作图以及遗传疾病研究方面最常用的遗传标记物^[3]。

近年来, 对不同民族 STR 基因座多态性研究的报道很多^[4-12], 但有关仫佬族的 STR 基因座的研究尚未见报道。因此, 本研究采用分子生物学技术, 探讨广西仫佬族人群的 9 个 STR 基因座的遗传多态性, 并与不同民族进行差异性比较, 为群体遗传学、法医学个体识别和亲子鉴定等研究提供资料。

收稿日期: 2004-10-12; 定稿日期: 2005-06-02

基金项目: 国家自然科学基金 (30260044) 资助

作者简介: 邓琼英 (1971-), 女, 在读博士生, 主要从事肿瘤免疫学研究。

1 材料和方法

1.1 材料

1.1.1 样本采集

随机选取广西罗城三代以上均为仫佬族的无关个体共 183 例, 分别采集其静脉血样约 2ml, 枸橼酸钠抗凝, 分装于 5ml 的试管中, -20℃ 保存备用。

1.1.2 主要试剂

TaqGold DNA polymerase, 荧光标记 16 位点的 PCR 复合扩增试剂盒 (AmpFISTR Identifier™ PCR Amplification Kit), 电泳分离检测试剂 (Hi-Di Formamide, GeneScan™-500LIZ™ Size Standard) 均购于美国 ABI 公司。

1.1.3 主要仪器

ABI3100 型 DNA 测序仪、9700 型 PCR 扩增仪、GeneScan Analysis 3.7 和 Genotyper 3.7 分析软件由美国 ABI 公司提供。台式高速冷冻离心机为 Hettich 公司的产品。

1.2 方法

1.2.1 DNA 提取

采用 Chelex-100 法提取 DNA。

1.2.2 PCR 扩增

按照 AmpFISTR Identifier™ PCR Amplification Kit 试剂盒上的说明书, 将 25μl 的反应体系改良为 10μl 反应体系 (分别取 PCR Reaction MIX 4.4μl, Identifier™ PrimerSet 2.2μl, TaqGold DNA polymerase 0.2μl 以及 H₂O 3.3μl 混匀后, 取 9μl 混合液装于 EP 管中, 再加入模板 DNA 1μl), 在 ABI9700 扩增仪中进行扩增。反应条件: 95℃ 预变性 11 min, 然后 94℃ 1 min, 59℃ 1 min, 72℃ 1 min, 循环 10 次, 再 60℃ 60 min 后 4℃ 保存。扩增产物于 4℃ 避光保存备用。

1.2.3 扩增产物的电泳分离及检测

电泳样品的制备: 取 Hi-Di Formamide 12μl, GS-500LIZ 分子量标记物 0.3μl, PCR 扩增产物 1μl, 混匀, 于 96℃ 变性 3min 后立即冰浴。上样: 取变性液 10μl, 加入样本板中, 经 ABI3100 遗传分析仪进行毛细管电泳, Datacollection101 收集信息后, 用 GeneScan Analysis 3.7 和 Genotyper 3.7 分析软件进行等位基因分型。

1.2.4 统计分析

用直接记数法计算各基因座的基因频率和基因型频率, 并按文献[13] 计算各基因座的杂合度(H)、个体识别力(DP)、非父排除率(EPP)、多态信息总量(PIC) 以及 9 个 STR 基因座的累积非父排除率(CEP)、累积多态信息总量等群体遗传学参数。按 Rostedt 等^[14] 的方法将基因型期望值小于 5 的合并为一项, 再采用 χ^2 检验进行 Hardy-Weinberg 平衡吻合度检验。用 SPSS 统计软件进行民族间的遗传差异的比较。

2 实验结果

2.1 广西仫佬族 STR 位点基因频率分布

仫佬族 9 个 STR 基因座的等位基因频率分布见表 1。

表 1 仫佬族 9 个 STR 基因座的等位基因频率分布
STR allele frequencies in Mulao population (n= 366)

等位基因 Allele	频率(Frequency)								
	D3S1358	D5S818	D7S820	D13S317	CSF1PO	TPOX	TH01	VWA	FGA
6							0.0710		
7		0.0519			0.0164		0.2951		
8			0.1721	0.3033		0.5301	0.0492		
9		0.0628	0.0410	0.1366	0.0383	0.1339	0.5027		
9.3							0.0464		
10		0.1967	0.1721	0.1530	0.2404	0.0191	0.0355		
11		0.3361	0.4290	0.2842	0.1885	0.2978			
12	0.0055	0.1967	0.1530	0.0929	0.3880	0.0191			
13		0.1339	0.0328	0.0246	0.1148				
14	0.0546	0.0137		0.0055	0.0109			0.3497	
15	0.2486	0.0082			0.0027			0.0109	
16	0.3197							0.1066	0.0027
16.2									0.0055
17	0.3087							0.2350	
18	0.0601							0.1858	0.0328
19	0.0027							0.0902	0.0765
20								0.0191	0.0492
21									0.1011
22									0.2077
22.2									0.0055
23								0.0027	0.1803
23.2									0.0137
24									0.1612
24.2									0.0027
25									0.0656
26									0.0628
27									0.0246
28									0.0082
SUM	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

注: n 为样品数

2.2 有关群体遗传学指标

仫佬族 9 个 STR 基因座的杂合度(H)、个体识别力(DP)、非父排除率(EPP)、多态信息总量(PIC)等群体遗传学参数见表 2。

2.3 与其他人群基因频率分布的对比分析

将本研究的结果与文献[4—12]所报道的广西壮族、苗族和回族、云南傈僳族、怒族和普米族以及湖南汉族、新疆维吾尔族、青海撒拉族、甘肃东乡族的 9 个 STR 位点的等位基因频率分布分别进行对比,经 SPSS 统计软件检验得到 P 值,见表 3。

表 2 仫佬族 9 个 STR 基因座群体遗传学指标
The statistical indexes of 9 STR loci in Mulao population

基因座 Locus	杂合度 H	个体识别力 DP	非父排除率 EPP	多态信息总量 PIC
D3S1358	0.7104	0.8857	0.4920	0.6867
D5S818	0.7814	0.9177	0.5866	0.7546
D7S820	0.7049	0.9191	0.5178	0.6941
D13S317	0.7377	0.9166	0.5688	0.7417
CSF1PO	0.6776	0.8953	0.5347	0.7015
TPOX	0.6230	0.7727	0.3484	0.5480
TH01	0.6175	0.8313	0.4079	0.5978
VWA	0.8251	0.9800	0.5573	0.7335
FGA	0.8907	0.9638	0.7402	0.8564
平均值	0.7298	0.8889	0.5282	0.7016
累积概率		0.99999999	0.999098	0.999987

表 3 广西仫佬族与不同人群 STR 基因座的基因频率分布比较结果
The result of comparison of 9 STR loci allele frequencies between
Mulao and the other population (the value of P)

基因座	广西 壮族	广西 苗族	广西 回族	云南 傣族	云南 怒族	云南 普米族	湖南 汉族	新疆 维吾尔族	青海 撒拉族	甘肃 东乡族
D3S1358	> 0.05	< 0.005	< 0.005	< 0.005	< 0.005	< 0.005	> 0.05	> 0.05	< 0.05	< 0.05
D5S818	> 0.05	< 0.005	< 0.05	< 0.005	< 0.005	< 0.05	> 0.05	< 0.005	> 0.05	> 0.05
D7S820	> 0.05	< 0.005	< 0.05	< 0.005	< 0.005	< 0.005	> 0.05	< 0.005	< 0.05	< 0.05
D13S317	> 0.05	> 0.05	< 0.05	< 0.005	< 0.005	< 0.005	> 0.05	< 0.005	< 0.005	< 0.005
CSF1PO	> 0.05	< 0.05	< 0.005	> 0.05	< 0.005	< 0.005	> 0.05	< 0.005	< 0.005	< 0.05
TPOX	> 0.05	> 0.05	> 0.05	> 0.05	< 0.005	< 0.005	> 0.05	< 0.005	> 0.05	> 0.05
TH01	< 0.005	> 0.05	< 0.005	> 0.05	< 0.05	< 0.005	> 0.05	< 0.005	> 0.05	< 0.05
VWA	> 0.05	< 0.005	> 0.05	< 0.005	< 0.005	< 0.005	< 0.005	< 0.005	< 0.005	< 0.005
FGA	> 0.05	< 0.005	< 0.005	< 0.005	< 0.005	< 0.005	< 0.05	< 0.005	< 0.005	< 0.05

3 讨 论

3.1 广西仫佬族群体 9 个 STR 位点的频率分布特点

中国仫佬族绝大多数分布于广西罗城境内,是广西土著民族,他们没有自己民族的语言和文字,俗语属汉藏语系—壮侗语系—侗水语支,与毛南语、侗语非常接近。其聚居地多为山区,交通极为不便,经济文化相对落后,很少与外界通婚,可视为隔离人群,在长期的生活环境中应形成了自己的体质特征和遗传基因。本研究结果显示,仫佬族的 9 个 STR 位点共检出 70 种等位基因,其频率分布在 0.0027~ 0.5301 之间(表 1);高频基因分布分别为 D3S1358/A16 为 0.3197, D5S818/A11 为 0.3361, D7S820/A11 为 0.4290, D13S317/A8 为 0.3033, CSF1PO/A12 为 0.3880, TPOX/A8 为 0.5301, TH01/A9 为 0.5027, VWA/A14 为 0.3497, FGA/A22 为 0.2077,与已报道的国内各群体的常见高频基因基本一致;共检出 207 种基因型,其频率分布在 0.0055~ 0.3388 之间;对 9 个 STR 位点基因型的观察值和期望值进行 χ^2 检验,均符合 Hardy-Weinberg 平衡定律($P > 0.05$),说明本次调查资料的可靠性和准确性。

3.2 群体 STR 的遗传多态性

遗传标记的多态性程度及其应用价值通常用 H、DP、EPP、PIC 等指标来评价。其中 H 能客观反映出群体的遗传变异水平, 平均 H 越大, 表明群体内遗传差异越大; DP 和 EPP 则反映该遗传标记在法医学个体识别及亲子鉴定中的能力, 一般 $DP > 0.8$ 、 $EPP > 0.5$ 时, 属于高度多态性遗传标记, 具有较高的应用价值; 而 PIC 是评价一个多态性基因座的使用价值的一个量化指标, 它能直接反映出一个遗传标记所包含或所能提供的遗传信息容量, 当 $PIC > 0.5$ 时, 常认为该标记具有高度的可提供信息量^[15]。本研究结果(表 2)显示, 仫佬族群体 9 个 STR 基因座平均 $H > 0.7$ 、 $DP > 0.8$ 、 $EPP > 0.5$ 、 $PIC > 0.7$, 累积 $DP = 0.999999999$, $EPP = 0.999098$, $PIC = 0.999987$, 表明是一组具有高杂合性、高鉴别力及高度多态性的遗传标记, 反映出仫佬族具有高度遗传多态性的特点, 因此, 实用价值较高, 可用于法医学个体识别和亲子鉴定及群体遗传学等研究。

3.3 民族间差异的比较

将广西仫佬族与其他十个民族相应位点的基因频率分布分别进行比较(表 3), 结果发现广西仫佬族与广西壮族和湖南汉族除了一两个位点之间存在差异外, 其余位点均无差异 ($P > 0.05$), 推测广西以壮族为主, 壮族人口在全区各处均有散居, 而仫佬族聚居地广西罗城临近湖南, 故与两者基因交流机会相对较多, 差异较其它民族小; 仫佬族与广西苗族、广西回族、云南傈僳族、云南怒族、云南普米族、新疆维吾尔族、青海撒拉族、甘肃东乡之间, 绝大多数位点均存在显著性差异 ($P < 0.05$), 究其原因可能是云南三少数民族多聚居在山区, 交通极为不便, 很少与外界通婚, 而北方三个民族与广西距离较远, 因此他们与仫佬族之间的基因交流机会非常小, 故绝大多数位点等位基因频率分布存在很大的差异; 而广西苗族和回族与仫佬族相似, 他们各聚居在不同的山区, 可能相互之间基因交流机会也不多, 所以大多数位点均存在差异。

因此, 分析不同民族之间 STR 位点的差异性对探讨民族的起源、迁徙、进化关系和民族学的研究均具有重要意义。

参考文献:

- [1] 高树辉, 王新淮. 浅谈 DNA 鉴定在法庭科学中的应用研究[J]. 公安大学学报(自然科学版), 2001; 6(6): 4—8.
- [2] 陈雪玲, 袁育康, 黄辰, 等. 短串联重复序列的研究进展[J]. 医学综述, 2002; 8(2): 72—74.
- [3] 袁文涛, 徐红岩, 赵进英, 等. 微卫星位点在基因组扫描中的信息表现[J]. 中华医学遗传学杂志, 2000, 17(2): 65—71.
- [4] 金天博, 高雅, 赖江华, 等. 中国广西壮族 9 个 STR 基因座遗传多态性研究[J]. 遗传学报, 2002, 29(12): 1052—1056.
- [5] 高树辉, 王新淮, 赖江华, 等. 云南怒族 STR 基因座遗传多态性研究[J]. 遗传, 2002, 24(2): 125—130.
- [6] 赖江华, 陈艳炯, 朱波峰, 等. 中国普米族、傈僳族 STR 遗传多态性研究[J]. 遗传学报, 2002, 29(11): 959—965.
- [7] 陈腾, 张琳琳, 赖江华, 等. 中国东乡族 9 个 STR 基因座遗传多态性研究[J]. 遗传, 2002, 24(3): 247—250.
- [8] 刘长晖, 杨电, 汪萍, 等. 广西地区回族人群 15 个 STR 基因座的多态性调查[J]. 中国法医学杂志, 2002, 17(5): 291—292.
- [9] 刘超, 杨电, 刘长晖, 等. 广西苗族人群 15 个 STR 基因座的多态性调查[J]. 法医学杂志, 2003, 19(4): 204—206.
- [10] 王占海, 陈腾, 安卫国, 等. 中国撒拉族 9 个 STR 基因座的遗传多态性[J]. 中国法医学杂志, 2004, 19(3): 171—172.
- [11] 刘秋玲, 吕德坚, 陈丽娴, 等. 湖南汉族人群 15 个 STR 基因座的遗传多态性[J]. 中国法医学杂志, 2004, 19(2): 98—99.
- [12] 金枝, 郭微, 张颖, 等. 新疆维吾尔族 15 个 STR 基因座的遗传多态性[J]. 刑事技术, 2004(3): 14—17.

- [13] 郑秀芬. 法医 DNA 分析[M]. 北京: 中国人民公安大学出版社, 2002, 383—387.
- [14] Rostedt I, Lahu K, Lukka M, *et al.* Genotyping of five short tandem repeat loci via triplex and duplex PCR[J]. *Forensic Sci Int*, 1996, 82: 217—226.
- [15] 赖江华, 张保华, 郑海波, 等. 中国纳西族 STR 遗传结构研究[J]. *遗传学报*, 2002, 29(7): 576—580.

Genetic Polymorphisms of 9 STR Loci in the Mulao Ethnic Group from Guangxi Province

DENG Qiong-ying¹, XU Lin¹, ZHOU Li-ring¹, LI Song-feng¹

(1. *Department of Anatomy, Guangxi Medical University, Nanning 530021*)

Abstract: Using PCR-STR, the polymorphism distributions of nine STR loci were investigated in 183 unrelated Mulao individuals from Guangxi Province, and as part of this work the genetic database of Mulao population was established. There were 70 STR alleles and 207 genotypes were observed in the nine STR of the Mulao ethnic group with statistical frequencies ranging from 0.0027~0.5301 and 0.0055~0.3388 respectively. The average heterozygosity was 0.7298; polymorphism information content was 0.7016; accumulative discrimination power was 0.999999999, and the probability of paternity exclusion was 0.999098. The result of a comparison between the Mulao and other ethnic groups showed that there were statistically significant differences between the Mulao and Miao, Hui minorities of Guangxi and the minorities of Yunnan and North China, but with there was no difference between the Mulao and Guangxi Zhuang and Hunan Han minorities. The obtained data can be useful for its application to genetics, forensic science and anthropology.

Key words: Mulao minority; PCR-STR; Genetic polymorphisms