

微卫星位点 DYS19 在中国人群中的多态研究

包维东 徐玖瑾 朱苏玲 许丽萍 杜若甫

(中国科学院遗传研究所, 北京 100101)

摘 要

以人 DNA 为模板, 经 PCR 扩增后, 用聚丙烯酰胺凝胶电泳分离扩增片段, 再经高灵敏度银染着色, 对中国陕西汉族、广东汉族、宁夏回族、辽宁满族、四川彝族、西藏藏族、广西壮族、广西瑶族、新疆维吾尔族、湖南土家族等 10 个人群 535 个个体的 Y 染色体上微卫星位点 DYS19 的遗传多态性进行了研究。结果表明: 中国人群中以等位基因 C (194bp) 为主要等位基因, 基因频率范围在 0.25—0.61; 等位基因 B (190bp)、D (198bp) 次之, 基因频率范围分别为 0.08—0.36、0.06—0.42; 而等位基因 A (186bp) 和 E (202bp) 频率较低, 频率范围分别为 0—0.07 和 0—0.38。在壮族中还检测出了一名携带 F (206bp) 等位基因的个体。中国人群 DYS19 等位基因的分布与蒙古人种群体以 C 型为主的结果相一致。 χ^2 成对比较表明瑶族、藏族与大多数其它中国民族间 DYS19 表型分布存在差异 ($P < 0.05$) 或显著性差异 ($P < 0.01$)。中国人群 DYS19 的基因频率至今在文献中尚未见报道。

关键词 中国民族, Y 染色体, 微卫星 DNA, DYS19, 多态性

1 前 言

微卫星 (Microsatellites) 也称短串联重复序列 (STRs), 是一类广泛存在于基因组中以 1—6bp 为重复单位的串联重复序列, 重复次数可达 100 次 (Tautz, 1993; Litt *et al.*, 1989)。

微卫星 DNA 以其在基因组中广泛分布、高度的多态性以及快捷简单的检测方法而在遗传图谱的构建中占有重要地位, 成为继 RFLP、RAPD、小卫星 DNA 之后的新一代分子标记。在人类基因组中它们大多表现出高度多态性, 并可运用 PCR 的方法加以检测 (Weber *et al.*, 1989)。DYS19 是存在于 Y 染色体短臂上一个以 GATA 为重复单位的微卫星 (Roewer *et al.* 1992)。它的 5 种常见的等位基因为 A 型 (186bp), B 型 (190bp), C 型 (194bp), D 型 (198bp), E 型 (202bp), 依次相差一个 GATA 单位, 此外, 还发现了 Z 型 (182bp), F 型 (206bp), BCE 和 “无” 这 4 种变异型。其中 BCE 代表一个体携带有 B、C、E 3 种变异型而一般个体仅带一种变异型。“无” 代表因引物配对区域碱基突变而致使 PCR 扩增成阴性 (Santos *et al.*, 1996)。由于该多态位点的限雄遗传特性, 它能提供有关人类

父系祖先的线索, 在人类进化研究上有重要意义 (Hammer and Horai, 1995)。DYS19 位点在法医学、亲子鉴定、民族间的亲缘关系等方面也有十分重要的意义。

2 材料与方法

本文研究的中国 10 个人群的血液样本采血地点和人数见表 1。每个群体随机抽取 50 名左右相互之间无直接亲属关系、3 代以内均为居住在当地的该民族的男性健康人。献血者多是该民族的大、中学校学生、回族献血者为工人。DNA 样品从人外周血液中或部分从该民族永生细胞株提取, 方法见 Sambrook *et al.* (1989)。PCR 扩增目标片段, 一对引物由美国斯坦福大学 L. L. Cavalli-Sforza 教授赠送。引物序列为: Y27H39.1 5'-CTA CTG AGT TTC TGT TAT ATG-3' Y27H39.2 5'-ATC GCA TGT AGT GAG GAC-A-3'。PCR 在 25 μ l 体系中进行 (含 PCR Buffer 2.5 μ l (100mM Tris-HCl pH8.8, 15mM MgCl₂, 500mM KCl), 模板 DNA 100ng, Taq DNA 聚合酶 1 单位, 引物各 10 μ M, 每种 dNTP 250 μ M。30

表 1 采样地点和群体

Sampling place and ethnic groups

民族 Ethnic groups	人数 Number tested	地点 Sampling place	经纬度 Longitude and latitude
维吾尔族 Uygur	61	新疆维吾尔自治区乌鲁木齐市 Urumqi City, Xinjiang Uygur Autonomous Region	(87.6E, 43.8N)
藏族 Tibetan	55	西藏自治区拉萨市 Lhasa City, Tibet Autonomous Region	(91.1E, 29.6N)
彝族 Yi	43	四川省布拖县 Butuo County, Sichuan Province	(102.8E, 27.7N)
回族 Hui	54	宁夏回族自治区同心县 Tongxin County, Ningxia Hui Autonomous Region	(105.9E, 36.9N)
瑶族 Yao	66	广西壮族自治区巴马县 Bama County, Guangxi Zhuang Autonomous Region	(107.2E, 24.1N)
北方汉族 Han (Northern)	44	陕西省麟游县 Linyou County, Shanxi Province	(107.8E, 34.6N)
壮族 Zhuang	45	广西壮族自治区武鸣县 Wuming County, Guangxi Zhuang Autonomous Region	(108.2E, 23.1N)
土家族 Tujia	64	湖南省吉首市 Jishou City, Hunan Province	(109.7E, 28.3N)
南方汉族 Han (Southern)	50	广东省广宁县 Guangning County, Guangdong Province	(112.4E, 23.6N)
满族 Manchu	53	辽宁省岫岩县 Xiuyan County, Liaoning Province	(123.2E, 40.2N)

次循环 (94 变性 1 分钟, 55 复性 1 分钟, 72 延伸 1 分钟)。在每一次 PCR 时, 包括以下对照: a. 只有超纯水 (无模板 DNA) 的阴性对照。b. 用已知能扩增的模板 DNA 的阳性对照。用 1.5% 琼脂糖凝胶电泳检测 PCR 扩增产物, 然后, 取 2.5 μ l PCR 产物在 10% 聚丙烯酰胺凝胶上 (16 厘米 \times 0.4 毫米厚), 150 伏恒压电泳 8 小时。电泳缓冲液是 TBE (0.18M Tris-borate, 0.004M EDTA, pH8.3)。电泳结束后, 用 Promega 公司 Silver Staining System 进行染色, 每一个体表型都经两次确认。

3 结果与讨论

聚丙烯酰胺凝胶电泳后银染色结果见图 1。

为判定 10 个人群、535 个个体的 DYS19 每一表型, 在每次 PAGE 凝胶电泳的胶板加样时, 加上一个已知表型碱基数大小的样本, 用来与 DNA Marker 标准条带的位置相比较, 依照这二条标记条带, 确定各人群的待测样品的各种表型。图 5 中 PAGE 电泳用的 DNA Marker 是 PBR322/Hae。图中有本研究所观察到的 5 种表型, 即 A (186bp)、B (190bp)、C (194bp)、D (198bp)、E (202bp)。



图 1 DYS19 位点 5 种等位基因的 PAGE 电泳图

Five different alleles of DYS19 visualized by silver staining

1. 7 为 DNA 分子量标记, PBR 322/Hae, 分子量由小到大分别为 184 bp、192 bp、213 bp、234 bp、267 bp。2 为 E 等位基因 (202 bp), 3 为 D 等位基因 (198 bp), 4 为 C 等位基因 (194 bp), 5 为 B 等位基因 (190 bp), 6 为 A 等位基因 (186 bp)

1, 7: DNA molecular marker PBR 322/Hae

with size: 184bp、192bp、213bp、234bp、267bp.

2: E allele (202bp), 3: D allele (198bp).

4: C allele (194bp), 5: B allele (190bp), 6: A allele (186bp)

表 2 是中国 10 个人群 DYS19 各人群表型分布及等位基因的频率。可以看出:

在中国 10 个人群中检出了 DYS19 的 6 种表型, 即 A、B、C、D、E、F。在 10 个人群中都检出了 B、C、D 3 种表型, 在南方汉族及瑶族中没有检出 A 型, 在回族、藏族中没有检出 E 型, 在壮族中发现了一例 F 型。

表 2 DYS19 等位基因频率在不同人群中的分布
DYS19 Allele frequencies distribution in different populations

群 体 Population	人数 No.	A 186bp	B 190bp	C 194bp	D 198bp	E 202bp	其它 other	文 献 Reference
彝族(Yi)	43	3(0.07)	15(0.35)	16(0.37)	8(0.19)	1(0.02)		
回族(Hui)	54	2(0.04)	16(0.29)	26(0.48)	10(0.19)			
藏族(Tibetan)	55	4(0.07)	11(0.20)	33(0.60)	7(0.13)			
土家族(Tujia)	64	3(0.05)	9(0.14)	39(0.61)	9(0.14)	4(0.06)		本文
满族(Man)	53	3(0.06)	12(0.23)	20(0.38)	14(0.26)	4(0.07)		
北方汉族(Han N.)	44	2(0.05)	8(0.18)	18(0.41)	11(0.25)	5(0.11)		
南方汉族(Han S.)	50		13(0.26)	17(0.34)	12(0.24)	8(0.16)		
维吾尔族(Uygur)	61	2(0.03)	22(0.36)	15(0.25)	20(0.33)	2(0.03)		
壮族(Zhuang)	45	1(0.02)	5(0.11)	17(0.38)	19(0.42)	2(0.04)	F1(0.02)	
瑶族(Yao)	66		5(0.08)	32(0.48)	4(0.06)	25(0.38)		
蒙古(Mongolia)	48	1(0.02)	16(0.33)	14(0.29)	13(0.27)	2(0.04)	F2(0.04)	Santos 等, 1996
朝鲜(Korea)	33		8(0.24)	15(0.45)	7(0.21)	3(0.09)		Gomolka 等, 1994
日本(Japan)	136	10(0.07)	4(0.03)	66(0.49)	33(0.24)	22(0.16)	F1(0)	Hammer 等, 1995
泰国(Thai)	42		3(0.07)	29(0.69)	6(0.14)	4(0.10)		Gomolka 等, 1994
英国(UK)	41	2(0.05)	33(0.80)	4(0.10)	1(0.02)	1(0.02)		Santos 等, 1996
德国(Germany)	306	21(0.07)	136(0.44)	74(0.24)	59(0.19)	16(0.05)		Muller 等, 1994
斯洛伐克(Slovakia)	81	4(0.05)	17(0.21)	16(0.20)	27(0.33)	17(0.21)		Muller 等, 1994
巴西(Brazil)	252	41(0.16)	135(0.54)	59(0.23)	14(0.06)	2(0.01)	BCE1(0)	Santos 等, 1993

在中国人群中等位基因 C 是主要的等位基因。在 10 个人群中, 有 8 个人群中的等位基因 C 的频率都是最高的, 其中藏族、土家族 C 等位基因频率分别高达 0.60 及 0.61, 但仍低于亚洲泰国人中检测到的 C 型 0.69 的值。北方汉族、瑶族及回族 C 等位基因频率都在 0.40 以上。这一特点与亚洲人以 C 为主的结论相一致。B、D 等位基因频率在各人群中居中, B 等位基因频率范围在 0.08—0.36, 多数人群在 0.20 左右, 只有新疆维吾尔族 B 等位基因较高, 为 0.36。D 等位基因频率在人群中的分布为 0.06—0.42, 以南方的壮族为最高。A、E 等位基因频率在人群中较低, 有的人群甚至为 0。比较特殊的是瑶族, 瑶族 E 等位基因频率为 0.38, 而 C 等位基因频率是 0.48, 没有发现 A 等位基因, B、D 等位基因频率只有 0.08 和 0.06。

这 10 个人群中, 除瑶族和维吾尔族外, 余下 8 个人群中基因频率最高的前两个等位基因都是相邻的两个等位基因如: CB、CD、DC 等, 而瑶族和维吾尔族的却不相邻。瑶族的

等位基因主要分布在两个相隔的等位基因 C 和 E 上。维吾尔族也同瑶族一样, 等位基因的分布不同于其余 8 个人群的“钟形”分布, 而呈“M”形分布, 只是表现得不如瑶族明显。

瑶族的另一个特点是 E 等位基因频率在这 10 个人群中是最高的, 占 0.38, 就是在世界上, 这一频率之高也是罕见的。这些特点似乎提示瑶族(这里特指巴马县的瑶族样本), 可能包含两个不同的群体分支。维吾尔族被认为有部分高加索人的血源。西欧高加索人 DYS19 是以等位基因 B 为主要等位基因的。维吾尔族 B 型占大多数和它的“M”型分布, 也支持维吾尔族含有部分高加索人的血缘。

经 χ^2 检验发现: 藏族除了与彝族、回族、土家族外, 与其他 6 个人群都有显著差异或极显著差异 ($P < 0.05$ 或 $P < 0.01$), 瑶族与其它 9 个人群有极显著的差异 ($P < 0.01$)。北方汉族和南方汉族之间无明显差异 ($P > 0.05$)。

世界范围内 DYS19 微卫星位点的多态分布呈现出较强的人群特异性, 西欧人以 B 型 (190bp) 为主, 亚洲人基本以 C 型 (194bp) 为主, 而美洲的印第安人以 A 型为主, 非洲人也以 C 型为主。英国人、荷兰人、德国人、意大利人和巴西白人都以 B 型为主。其中巴西白人源自从 16 世纪起迁到巴西的葡萄牙人和一部分意大利人。这些国家的居民都与欧洲古代民族中的凯尔特人、日耳曼人有关。因此有理由认为古代日耳曼民族和凯尔特人是 B 型为主, 另外据 Santachiara-Benerecetti *et al.* (1995) 报道, 西班牙的巴斯克人 B 型也很高。巴斯克人是比利牛斯山西部地区最古老的土著居民。看来在早期人类的大迁移中占据西欧、北欧这一带的人群可能是以 B 型为主的。而且越是地处大陆边缘地带和岛屿等地的人群越容易保存该人群的主要等位基因, 比如地处岛国的英国人的 B 型占 80%, 地处大陆边缘的荷兰人 B 型占 71%, 而地处中央地带的德国 B 型最高也不过为 54%, 南意大利和希腊人虽然位于地中海一带, 但历史上这一带正是战争频繁发生, 人口融合交流程度大的地区。所以 B 型频率仍低于英国和荷兰。另外从中国的藏族拥有 60% 的 C 型, 本州岛日本人拥有 55% 的 C 型来看, 似乎也存在这样一种趋势: 在大陆边缘和地理位置相对隔绝的地区生活的民族较那些地处大陆中央的民族更易保持其群体的等位基因频率。

关于微卫星 DNA 的突变率问题, 由于利用其基因歧化可以估计分化时间, 所以一直备受研究者们关注, 遗憾的是这方面的了解还相当的粗浅, 在一些问题上还没有统一的认识。比如在核心序列碱基数与突变率的关系上, 有人认为核心序列碱基数越少突变率越低 (Rienzo, 1994), 有人认为情况正好相反 (Chakraborty, 1997)。

现在的证据似乎暗示 Y 染色体上的微卫星 DNA 突变率偏低。第一, DYS19 在世界范围内的频率分布具有很大的异质性, 美洲印第安人以 A 型为主, 西欧的高加索人以 B 型为主, 亚洲人以 C 型为主。第二, 各地区不同部落的美洲印第安人绝大多数 (78%) 属于 A 型 (α 型 + DYS19A 型) (Santos *et al.*, 1996)。如果 DYS19 以同常染色体上的微卫星 DNA 相同的 10^{-4} — 10^{-5} 的速度进化, 那么“始祖效应”(founder effect) 早就会被稀释掉, 尤其是对于其先民是从距今两万年前的旧石器时代末期陆续从亚洲大陆迁入美洲大陆的印第安人而言。所以用 DYS19 这个位点的基因频率数据来探讨群体的基因流动时可不考虑基因突变的影响。

4 小 结

1. 中国 10 个人群 Y 染色体微卫星位点 DYS19 的等位基因频率, 在世界上均属首次报道。这些结果为我国人类群体遗传学研究提供了基本数据, 尤其对探讨我国各民族的源与流提供了重要资料。

2. DYS19 位点在我国人群中具有很强的遗传多态性, 共发现了 6 种等位基因, 每一等位基因相差一个“GATA”重复串联序列, 即 A (186bp)、B (190bp)、C (194bp)、D (198bp)、E (202bp) 及 F (206bp)。DYS19 是研究 Y 染色体 DNA 多态性十分理想的微卫星位点。

3. 在我国人群中检出的 6 种等位基因以 C 等位基因为主, 藏族、土家族 C 型频率分别高达 60% 及 61%。与蒙古人种中 DYS19 以 C 等位基因频率较高的结果相一致。维吾尔族则以 B 型为主, 而南方的壮族则以 D 型频率 0.42 为最高。

参 考 文 献

- Chakraborty R, Kimmel M, Stivers DN *et al.* 1997. Relative mutation rates at di-, tri- and tetranucleotide microsatellites loci. *Proc Natl Acad Sci USA*, 94: 1041—1046.
- Gomolka M, Hundrieser J, Nurnberg P 1994. Selected di- and tetranucleotide microsatellites from chromosomes 7, 12, 14 and Y in various Eurasian populations. *Hum Genet*, 93: 592—596.
- Hammer M F, Horai S. 1995. Y chromosomal DNA variation and the peopling of Japan. *Am J Hum Genet*, 56: 951—962.
- Litt M, Luty J A. 1989. A hypervariable microsatellite revealed by in vitro amplification of dinucleotide repeat with in the cardiac muscle actin gene. *Am J Hum Genet*, 44: 397—401.
- Muller S, Gomolka M, Walter H. 1994. The Y-specific SSLP of the locus DYS19 in four different European samples. *Hum Hered*, 44: 298—300.
- Rienzo A D, Peterson A C, Garza J C *et al.* 1994. Mutational processes of simple-sequence repeat loci in human populations. *Proc Natl Acad Sci USA*, 91: 3166—3170.
- Roewer L, Arnemann J, Spurr N K *et al.* 1992. Simple repeat sequences on the human Y chromosome are equally polymorphic as their autosomal counterparts. *Hum Genet*, 89: 389—394.
- Sambrook J, Frisch E F, Maniatis T. 1989. *Molecular Cloning. a laboratory manual*. Second edition. USA: Cold Spring Harbor, Cold Spring Harbor Laboratory Press.
- Santos F R, Bianchi N O, Pena S D J. 1996. Worldwide distribution of human Y-chromosome haplotypes. *Genome Res*, 6: 601—611.
- Santos F R, Penna S D J, Epplen J T. 1993. Genetic and population study of a Y-linked tetranucleotide repeat DNA polymorphism with a simple non-isotopic technique. *Hum Genet*, 90: 655—656.
- Santos F R, Gerelsaikhan T, Munkhtuja B *et al.* 1996. Geographic differences in the allele frequencies of the human Y-linked tetranucleotide polymorphism DYS19. *Hum Genet*, 97: 309—313.
- Tautz D. 1993. Notes on the defunction and nomenclature of tandemly repetitive DNA sequences. In: Pena S D J, Chakraborty R, Epplen J T *et al.* eds. *DNA Fingerprinting: State of the Science*. Basel: Birkhauser Verlag, 21—28.
- Weber J L, May P E. 1989. Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *Am J Hum Genet*, 44: 388—396.

A STUDY OF A POLYMORPHIC MICROSATELLITE LOCUS DYS19 IN CHINESE ETHNIC GROUPS

Bao Weidong Xu Jiujin Zhu Shuling Xu Liping Du Ruofu
(*Institute of Genetics, Chinese Academy of Sciences, Beijing 100101*)

Abstract

The distribution of a Y-chromosome specific polymorphic locus, DYS19, in 535 males of ten ethnic groups in China was studied by using PCR method followed by PAGE and silver staining. The ten ethnic groups are Han in Shanxi Province, Han in Guangdong Province, Hui in Ningxia Hui Autonomous Region, Man in Liaoning Province, Uygur in Xinjiang Uygur Autonomous Region, Yi in Sichuan Province, Tibetan in Tibet Autonomous Region, Zhuang in Guangxi Zhuang Autonomous Region, Yao in Guangxi Zhuang Autonomous Region, Tujia in Hunan Province. The result shows that allele C (194bp), frequency of which ranges from 0.25 to 0.61, is predominant in Chinese populations; alleles B (190bp) and D (198bp) take the second place with the frequency ranging from 0.08 to 0.36 and 0.06 to 0.42, respectively; while alleles A (186bp) and E (202bp) take the last place with the frequency ranging from 0 to 0.07 and 0 to 0.38, respectively. An individual carrying the allele F (206bp) was identified in Zhuang. The pairwise comparison with χ^2 test reveals that there exists difference ($P < 0.05$) or significant difference ($P < 0.01$) in DYS19 phenotype distribution between Yao, Tibetan and most other Chinese populations. The allele frequency of DYS19 in Chinese ethnic groups has not been reported so far.

Key words Chinese populations, Y-chromosome, Microsatellite DNA, DYS19, Polymorphism