

MYS2 多态位点在中国 26 个人群中分布的研究

薛雅丽^{1,3}, 包维东², 傅松滨¹, Tatiana Zerjal³, 朱苏玲², 徐玖瑾²
杜若甫², 张贵寅¹, 李 璞¹, Chris Tyler-Smith³

(1. 哈尔滨医科大学生物学教研室, 哈尔滨 150086; 2 中国科学院遗传学研究所, 北京 100101;
3. Department of Biochemistry, University of Oxford, South Park Road OX1 3QU, UK)

摘要: 对中国 26 个人群中 Y 染色体小卫星位点 MYS2 的分布进行了研究, 结果显示: MYS2*4 是常见的等位基因类型, MYS2*3 在中国不同人群中分布差异很大(0—32.4%); 并对该位点在不同人群中分布的频率进行了 AMOVA 分析, 结果为: 不分组、南方人群和北方人群的遗传学差异分别为 7.76%、8.11% 和 1.42%。南北方人群间、东北和西北人群间的差异分别为 4.34% 和 1.57%。结论: MYS2 是研究中国人群遗传结构有价值的多态位点, 中国不同人群之间存在一定的遗传学差异, 中国南方人群间的遗传学差异较大, 再次证实了中国人群的南北差异。

关键词: Y 染色体; MYS2; AMOVA; 中国人群

中图法分类号: Q987 文献标识码: A 文章编号: 1000-3193(2003)01-0063-06

人类 Y 染色体多态性在人类遗传学研究中具有重要的意义, 而且得到了广泛的应用^[1]。MYS2(minisatellite Y 2, DYS440) 是最近发现的人类 Y 染色体上的第二个小卫星多态性位点^[2]。它由长度为 99—110bp 的 3 或 4 重复单位组成, 位于 DBY 基因上游大约 1kb 处。常见的等位基因类型是 4 个重复单位。在大猩猩和黑猩猩中的重复单位为 2。该位点的进化速率较慢, 因而在人类进化研究中成为又一个有价值的多态性位点。

本研究检测了中国各地区的 26 个人群中的多态性位点的分布, 以研究该位点在中国人群中的分布特征和其在人类进化中的作用。

1 材料和方法

1.1 样品的采集和基因组 DNA 的制备

共采集了覆盖全中国的 26 个不同的人群, 共计 868 份的男性个体血样。这些个体之间无亲缘关系, 三代是同一民族。采样地详见表 1。所有血液样本均用 ACD 抗凝, -70℃ 保存直至提取基因组 DNA。按常规的苯酚氯仿法提取基因组 DNA。

收稿日期: 2001-10-22; 定稿日期: 2002-01-09

基金项目: 国家自然科学基金重大项目资助(39993420); 英国皇家学会资助项目(Q740); 黑龙江省科技攻关项目资助(G99C19-F1)

作者简介: 薛雅丽(1967-), 女, 医学博士, 副教授。哈尔滨医科大学生物学教研室, 主要从事人类基因组学及人类基因组多样性研究。

表 1 样品采集情况表
26 populations sampled in the current study

人 群 Population	人 数 Sample size	采样地点 Sampling place	经纬度 Location	语系和语族 Language Family
汉族(Han)	21	内蒙古海拉尔市	51.0, 126.0	阿尔泰语系- 蒙古语族
	35	黑龙江哈尔滨	45.7, 126.6	阿尔泰语系- 蒙古语族
	30	甘肃兰州	36.0, 103.7	阿尔泰语系- 蒙古语族
	32	新疆伊宁	43.3, 80.5	阿尔泰语系- 蒙古语族
	35	四川成都	30.6, 104.1	阿尔泰语系- 蒙古语族
	35	广州梅县	24.3, 116.1	阿尔泰语系- 蒙古语族
	45	内蒙古海拉尔市	51.0, 126.0	阿尔泰语系- 蒙古语族
蒙古(Mongol)	45	内蒙古海拉尔市	51.0, 126.0	阿尔泰语系- 蒙古语族
满族(Manchu)	35	辽宁岫岩	40.2, 123.2	阿尔泰语系- 满语语系
朝鲜族(Korean)	27	吉林延吉	42.0, 121.4	未定
达斡尔族(Daur)	31	内蒙古莫力达瓦旗	48.0, 118.6	阿尔泰语系- 蒙古语族
赫哲族(Hezhe)	45	黑龙江街津口	46.7, 132.5	阿尔泰语系满- 通古斯语族
鄂温克(Ewenki)	27	内蒙古莫力达瓦旗	48.0, 118.6	阿尔泰语系- 蒙古语族
鄂伦春(Oroqen)	31	内蒙古阿里河	53.0, 131.0	阿尔泰语系满- 通古斯语族
回族(Hui)	35	宁夏同心	36.9, 105.9	阿拉伯语系
维吾尔(Uyghur)	39	新疆伊宁	43.3, 80.5	阿尔泰语系- 突厥语族
维吾尔(Uyghur)	31	新疆乌鲁木齐	43.8, 88.6	阿尔泰语系- 突厥语族
锡伯族(Xibo)	42	新疆察布查尔	43.3, 80.5	阿尔泰语系满- 通古斯语族
哈萨克(Kazark)	17	新疆伊宁	43.3, 80.5	阿尔泰语系- 突厥语族
畲族(She)	34	福建罗源	26.4, 119.5	阿尔泰语系- 汉藏语系
瑶族(Yao)	35	广西巴马	24.1, 107.2	阿尔泰语系- 汉藏语系
	35	广东连南	24.7, 112.2	阿尔泰语系- 汉藏语系
黎族(Li)	34	海南通什	18.7, 109.5	阿尔泰语系- 汉藏语系
布依族(Buyi)	35	贵州镇宁	26.0, 105.7	阿尔泰语系- 汉藏语系
哈尼族(Hani)	34	云南元江	23.5, 101.9	阿尔泰语系- 汉藏语系
羌族(Qiang)	33	四川毛温	31.6, 103.8	阿尔泰语系- 汉藏语系
藏族(Tibetan)	35	西藏拉萨	29.6, 91.1	阿尔泰语系- 汉藏语系

1.2 MYS2 的 PCR 扩增

PCR 扩增用 DNA 热循环仪(GENE 2000), 上游引物 P₁ 序列为: 5' CTGGCITTGGG-GACGTAGTA 3', 下游引物 P₂ 序列为: 5' CTGCGCITTATGAGCATACG'; 12.5μl 反应体系中含有: 1× Bioline buffer, 200 μM dNTP, 1.5 mM MgCl₂, 引物各 1 μM, 0.5 μ TaqDNA 聚合酶和 5—10ng 模板 DNA。扩增条件: 94℃ 预变性 3 min, 30 个循环扩增(94℃ 35s, 64℃ 35s, 72℃ 40s), 最后 72℃ 延伸 5 min。PCR 产物经 2% 琼脂糖电泳检验扩增结果。

1.3 数据分析

采用 Arlequin 软件(<http://anthro.unige.ch/arlequin>) 对该多态位点在不同人群中的分布进行 AMOVA (analysis of molecular variance) 分析, 将所研究的人群分为五组: 不分组南北人群组、南方人群、北方人群(以长江为界)及西北和东北分组(以黄河为界), 以分析不同地域人群的遗传结构差异。

2 结果

2.1 MYS2 的 PCR 扩增结果

经 PCR 扩增后,可获得两种等位基因类型,分别是 MYS^{*} 4 和 MYS^{*} 3。PCR 片段的大小为~ 400bp 和~ 300bp(图 1)。

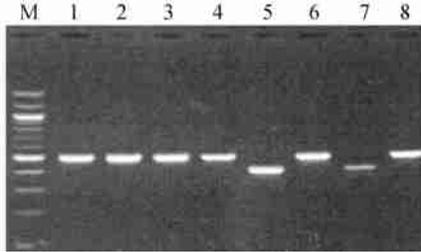


图 1 MYS2 在 2% 琼脂糖的检测结果

The agarose gel image of MYS2

M: 100bp DNA Ladder; 1, 2, 3, 4, 6, 8: MYS^{*} 4; 5, 7: MYS^{*} 3

2.2 MYS2 在 26 个人群中的分布

在所检测的 868 个男性个体中, MYS2^{*} 4 是常见的等位基因类型, 占 94.6% (821/868), 在中国的北方人群中频率很高(82.9% —100%), 而 MYS2^{*} 3 在中国的北方人群中除回族外的其他人群中出现的频率较低, 而在较多的南方人群中的频率较高, 在海南黎族中最高, 高达 32.4%。MYS2 的在 26 个人群中的分布见表 2。

2.3 AMOVA 分析结果

分别对 MYS2 在不同人群中分布的频率分别按地域分布分组并进行 AMOVA 分析, 分析结果见表 3。不分组时不同人群之间的差异为 7.76%, 差异显著。南北方人群中存在 4.34% 的差异, 且该差异显著, 说明南北人群具有一定的遗传学差异。南方不同人群间的差异为 8.11%, 且差异显著, 而北方人群间的差异为 1.42%, 说明南方人群间的差异比北方人群间的差异大。但以黄河为界将北方人群分为东北人群和西北人群进行分析时, 发现存在一定的差异, 但这种差异不是很显著。

3 讨论

越来越多的研究显示: Y 染色体多态是法医学、人类遗传学及人类进化等研究中的非常有价值的研究工具^[1]。这类多态性位点包括突变率较高的微卫星位点和突变率低的双等位基因。微卫星位点由于突变率较高, 因而对较远的历史事件缺乏分辨力^[3-4]。突变率较低的双等位基因虽然能够代表稳定的谱系关系, 但是却常常产生偏倚而低估了人群的多样性。例如 Sb bin 等人所表发的双等位基因在中国北方人群中的变异非常小^[5], 但以往的研究显示, 这些人群的遗传结构并非如此相近^[6]。因而, 需要一种既在人群中存在一定的变异, 又不会产生偏倚的 Y 染色体多态性标记, 而且其突变率应足够低以保证在和其他多态性位点使用时不产生类同效应(homoplasy)。MYS2 就是这样一个 Y 染色体多态性位点^[2]。

表 2 MYS2 在中国 26 个人群中的分布
The distribution of MYS2 in 26 Chinese populations

人群 (Population)	人数 (Sample size)	MYS2* 3	MYS2* 4
汉族 (Han)/ 内蒙古	21	0	21 (100%)
汉族 (Han)/ 黑龙江	35	1 (2.9%)	34 (97.1%)
汉族 (Han)/ 新疆	33	4 (12.1%)	29 (87.9%)
汉族 (Han)/ 甘肃	30	2 (6.7%)	28 (93.4%)
汉族 (Han)/ 四川	35	6 (17.1%)	29 (82.9%)
汉族 (Han)/ 广东	35	7 (20.0%)	28 (80.0%)
蒙古 (Mongol)/ 内蒙古	45	0	45 (100%)
满族 (Manchu)/ 辽宁	35	1 (2.9%)	34 (97.1%)
朝鲜族 (Korean)/ 吉林	27	0	27 (100%)
达斡尔族 (Daur)/ 黑龙江	39	2 (5.1%)	37 (94.9%)
赫哲族 (Hezhe)/ 黑龙江	45	0	45 (100%)
鄂温克 (Ewenki)/ 内蒙古	26	2 (7.7%)	24 (92.3%)
鄂伦春 (Oroqen)/ 内蒙古	31	0	31 (100%)
回族 (Hui)/ 宁夏	35	3 (17.1%)	32 (82.9%)
维吾尔 (Uyghur)/ 新疆伊宁	39	0	39 (100%)
维吾尔 (Uyghur)/ 新疆乌鲁木齐	31	1 (3.2%)	30 (96.8%)
锡伯族 (Xibe)/ 新疆	43	3 (7.0%)	40 (93.0%)
哈萨克 (Kazark)/ 新疆	17	0	17 (100%)
畲族 (She)/ 福建	34	6 (17.6%)	28 (82.4%)
瑶族 (Yao)/ 广西	35	1 (2.9%)	34 (97.1%)
瑶族 (Yao)/ 广东	35	3 (8.6%)	32 (91.2%)
黎族 (Li)/ 海南	34	11 (32.4%)	23 (67.6%)
布依族 (Buyi)/ 贵州	35	1 (2.9%)	34 (97.1%)
哈尼族 (Hani)/ 云南	34	0	34 (100%)
羌族 (Qiang)/ 四川	33	5 (15.2%)	28 (84.8%)
藏族 (Tibetan)/ 西藏	35	0	35 (100%)

表 3 MYS2 遗传变异的分配
The apportion of genetic variation

	群体内变异	组内群体间变异	组间变异	P 值
不分组	92.24%	7.76%		< 0.01
南方组	91.89%	8.11%		< 0.01
北方组	98.58%	1.42%		0.11
南/北分组	90.22%	5.44%	4.34%	< 0.01
东北/西北分组	97.84%	0.59%	1.57%	> 0.05

MYS2 和以往发现的 Y 染色体小卫星 MYS1 不同^[7], 其重复单位是已知的小卫星中最长的一个而多样性却很低^[2]。MYS2 中的每个重复单位都包含一个回文结构 CCTAGG, 亦可形成发夹结构而维持其结构稳定, MYS2 的 GC 含量介于 MYS1 和常染色体小卫星之间, 约为 45%, 其变异主要是通过链内滑动和不等的姐妹染色体互换产生, 因而其突变率较低^[8]。以往的研究推测: 最初的很可能只是包含 CCTAGG 的一个重复单位, 出现在巨猿分歧的时期, 而后在猩猩从非洲猿中分化出来时发生了复制而成为 2 个重复单位。但是在人类 MYS2^{*}4 的祖先等位基因, 而 MYS2^{*}3 是突变的等位基因。如果该位点的重复单位的增加和减少的机率是相同的话, 我们还应该发现另一种等位基因, MYS2^{*}5。但事实上并非如此。这可能是由于该位点位于在 DBY 基因上游 1kb 的位置^[9], 其重复单位的增加可影响到 DBY 基因的表达而导致不育, 使其受到选择的原因, 这种推测还有待进一步研究。

通过对中国 26 个人群该多态位点的检测发现, MYS2^{*}3 在不同人数中分布差异很大, 说明该位点是分析中国人群遗传结构的一个有价值的多态位点。通过对南方人群、北方人群分析发现, 南方人群的差异大于总的人群间的差异, 说明南方不同人群间的遗传结构差异较大, 而北方人群间的人群则差异很小。这可能和中国的地理结构有关。中国南方地区山多, 存在很多地理屏障而交通不便, 使得人群间的基因交流较少。而中国的北方地区多为平原, 地理屏障较少, 另外由于战乱、饥荒及北方地区的少数民族多为游牧民族的原因, 使得人群间的基因流动较多, 因而北方地区人群遗传结构相对匀称。以长江为界将这些人群分为南方人群和北方人群进行分析时, 发现南北方人群存在着一定的遗传学差异, 这和以往的研究结果一致, 也是关于南、北蒙古人群的祖先由两条路线从非洲或中东到中国, 并在长江会合又一佐证^[10]。

总之, MYS2 是研究中国人群有用的 Y 染色体多态性位点。利用该位点中国人群的研究进一步证实了中国人群的南北分界, 为中国人群的遗传结构、不同人群间的相互关系和源流提供了重要的资料。

参考文献:

- [1] Jobling MA, Tyler Smith C. Fathers and Sons: the Y chromosome and human evolution[J]. *Trend Genet*, 1995, 11: 449—456.
- [2] Weidong Bao, Suling Zhu, Apita Pandya *et al.* MYS2: a slowly evolving minisatellite on the human Y chromosome which provides a useful polymorphic marker in Chinese populations[J]. *Gene*, 2000, 244: 29—33.
- [3] De Knijff P, Kayser M, Corach D *et al.* Chromosome microsatellites: population genetic and evolutionary aspects[J]. *Int J Legal Med*, 1998, 110: 134—149.
- [4] Foster EA, Jobling MA, Taylor PG *et al.* Jefferson fathered slave's last child[J]. *Nature*, 1998, 396: 27—28.
- [5] Su B, Xiao J, Underhill P *et al.* Y chromosome evidence for a northward migration of modern humans in East Asia during the last ice age[J]. *Am J Hum Genet*, 1999, 65: 1718—1724.
- [6] Chu JY, Huang W, Kuang SQ *et al.* Genetic relationship of populations in China[J]. *Proc Natl Acad Sci USA* 1998, 95: 11763—11768.
- [7] Jeffreys AJ, Tamaki K, Macleod A *et al.* Complex gene conversion events in germline mutation at human minisatellites[J]. *Nat Genet*, 1994, 6: 136—145.
- [8] Jeffreys AJ, Neumann R, Wilson V. Repeat unit sequence variation in minisatellites: a novel source of DNA polymorphism for studying variation and mutation by single molecule analysis[J]. *Cell*, 1994, 60: 473—485.
- [9] Lahn BT, Page DC. Functional coherence of the human Y chromosome[J]. *Science*, 1997, 278: 675—680.
- [10] 肖春杰, 杜若甫, Vallir Stozza LL 等. 中国人群基因频率的主成分分析[J]. *中国科学(C 辑)*, 2000, 30: 439—442.

THE DISTRIBUTION OF MYS2 POLYMORPHISM IN 26 CHINESE POPULATIONS

XUE Ya-li^{1,3}, BAO Wei-dong², FU Song-bin¹, Tatiana Zerjal³, ZHU Su-ling²,
XU Jiu-jin², DU Ruo-fu², ZHANG Guiyin¹, LI Pu¹, Chris Tyler-Smith³

(1. *Department of Biology, Harbin Medical University, Harbin 150086;*

2. *Institute of Genetics, Chinese Academy of Sciences, Beijing 100101;*

3. *Department of Biochemistry, University of Oxford, South Park Road OX1 3QU, UK)*

Abstract: The distribution of MYS2 polymorphism in 26 Chinese populations was analyzed by PCR. And the variations of intra- and between populations were calculated using AMOVA by Arlequin Package. MYS2*4 was the common allele while the incidence of MYS2*3 was much different in these populations. By AMOVA, 7.76%、8.11% and 1.42% of total variation were found among all 26, 10 Southern and 16 Northern Chinese populations respectively. And 4.34% of total variation was found between the Southern and Northern groups while 1.57% was found between northeast and northwest groups. The results showed that MYS2 polymorphism is very useful in analyzing the genetic structure of Chinese population. Some difference in genetic structure was seen among the different populations and a geographical distinction between South and North was confirmed. The distinction between Western and Eastern populations in the Northern China was also suggested.

Key words: Y chromosome; MYS2; AMOVA; Chinese populations