

克里雅河下游封闭人群 DYS19 和 DYS390 多态性研究

段然慧¹, 刘伟强¹, 周 慧¹, 朱 泓¹

(1. 吉林大学考古 DNA 实验室, 长春 130023)

摘要: 本文以居住于塔克拉玛干沙漠当中克里雅河下游地区封闭人群(51 例男性)为研究对象, 采用基因扫描对其 DYS19 和 DYS390 两个 STR 基因座进行基因扫描研究其遗传多态性。对于 DYS19 基因座, 克里雅河下游的封闭人群等位基因分布呈现“M”形分布, 以 DYS19* 14 和 DYS19* 16 最常见, 基因频率分别为 0.353 和 0.510; 对于 DYS390 基因座, 其人群等位基因分布也并非呈现“钟形”分布, 而是以 DYS390* 21 和 DYS390* 24 两种基因型基因频率最高, 并且 DYS390* 21 为此人群等位基因重复次数最少的基因型, 基因频率分别为 0.235 和 0.431, 这可能是提示克里雅河下游的封闭人群的来源包含两个不同的群体分支。

关键词: Y 染色体; 封闭人群; STR 基因座; 多态性

中图法分类号: Q987 文献标识码: A 文章编号: 1000-3193(2004)04-0326-05

STR(short tandem repeat)是由 2—6bp 重复单位构成的微卫星 DNA, 广泛分别在人类基因组中。据统计在人类基因组中, 每 20kb 就含有一个 STR 基因座, 而且很多 STR 基因座都具有多态性, 使之成为一个重要的遗传标记。它具有扩增效率好和判断准确, 并且适用于微量甚至降解生物材料的 DNA 分型技术等优点^[1]。DYS19 基因座是 Y 染色体上以 GATA 为重复单位的微卫星, DYS390 基因座是 Y 染色体上以 GATA/GACA 为重复单位的微卫星, 目前 DYS19 和 DYS390 两基因座均发现 10 种单倍型^[2]。

克里雅河下游及附近, 是塔克拉玛干历史遗址集中分布地, 其东有尼雅、西有丹丹乌里克, 在下游地区还有喀拉墩、马坚勒克、圆沙古城遗址以及北部的龟兹国等遗址。据考证, 从西汉到南北朝时期, 这里曾是一个民间聚居地, 社会经济以狩猎和畜牧为主, 兼有农业, 冶炼和制陶。奇怪的是, 从唐末到清初, 在这里再也没有人类活动的记载。如今的克里雅人也没有承袭过去的经济活动, 可以认为不是古代克里雅河下游居民的后裔, 曾有报道克里雅居民是西藏古格王国逃来藏匿的遗民, 但从这里居民的宗教信仰和人种特点看和维吾尔族毫无二致。现在的克里雅人也都说不清他们何时从何地来到这里。

本文利用 DYS19 和 DYS390 这两个 STR 基因座对我国新疆塔克拉玛干沙漠中克里雅河下游地区的一个封闭人群采用基因扫描技术进行基因频率调查。那里的居民过着相对与世隔绝的生活, 固守着自己的生活方式和文化传统。由于他们的地理位置以及交通的险阻, 他

收稿日期: 2003-01-07; 定稿日期: 2004-06-22

基金项目: 教育部科学技术重大项目和文物局边疆考古专项基金资助。

作者简介: 段然慧(1975—), 女, 博士研究生, 主要从事古 DNA 及群体遗传学方面的研究。

通讯作者: 周慧, E-mail: zhouhui@mail.jlu.edu.cn

们很少与外界联系,更重要的是他们从未与外界通婚,因此他们的基因多态性还未受到外部人群的干扰,这在当今世界上也是罕见的^[3]。本文旨在研究其基因座基因型及单倍型构成,并分析其遗传多态性。

1 材料和方法

1.1 样本采集和基因组 DNA 提取

来自家庭内部三代内父系无相关的男性个体共计 51 人。为了妥善保管血样,我们取 20 μ l 血直接点在 FTA 卡片(GibcoBrl 公司)上,基因组 DNA 就会固定在 FTA 卡片的基质上,在空气中自然干燥过夜。从卡片中间取出 3mm 直径的血点,先用 FTA 纯化试剂洗涤三次,接着用 TE 洗涤两次,以去除 PCR 扩增抑制剂。干燥后的血样点放入已装好扩增试剂的扩增管中,作为模板准备扩增^[4]。

1.2 引物序列

DYS19-L CTA CTG AGT TTC TGT TAT AGT(TET 荧光标记)

DYS19-R ATG GCA TGT AGT GAG GAC A

DYS390-L TAT ATT TTA CAC ATT TTT GGG CC

DYS390-R TGA CAG TAA AAT GAA CAC ATT GC(FAM 荧光标记)^[5]

1.3 PCR 扩增及片段长度分析

在 20 μ l 的扩增体系中,含有 10mM Tris-HCl(pH8.8), 2.2mM MgCl₂, 50mM KCl, 0.1% TritonX-100, 200 μ M dNTPs, 2 Unit TaqDNA 聚合酶(Promega), PCR 扩增的引物 DYS19 的浓度为 0.2 μ M。PCR 扩增反应的条件是:95 $^{\circ}$ C 变性 2 分钟,然后 94 $^{\circ}$ C 变性 1 分钟,57 $^{\circ}$ C 退火 1 分钟,72 $^{\circ}$ C 延伸 1 分钟,循环 28 次,最后 72 $^{\circ}$ C 再延伸 10 分钟。复合反应 PCR 产物通过 ABI 310 遗传分析仪进行片段长度分析,选用 POP-4 的多聚物凝胶和 47cm 的毛细管进行电泳。微卫星位点 PCR 产物 0.3 μ l 与有 Tamra 荧光标记的片段长度标记 0.5 μ l 以及 12 μ l 的甲酰胺相混合 96 $^{\circ}$ C 变性 3 分钟,而后冰浴 5 分钟进行上样。电泳约需时 29 分钟,结果通过 Genescan 软件进行片段长度分析。选定一样本,对其 DYS19 和 DYS390 基因座 PCR 产物进行测序,根据序列结果并结合片段长度的分析结果可读出重复序列的数量,以此作为内标。在对其它样本进行分析时,不但可判断其重复序列的数量,而且还可以验证实验的准确性。

2 结果

2.1 DYS19 多态性分析结果

如图 1,显示了一个样本 DYS19 基因座的片段长度分析谱图,表 1 为克里雅河下游的封闭人群以及其它相关人群 DYS19 基因座等位基因频率的分布。

2.2 DYS390 多态性分析结果

如图 2,显示了一个样本 DYS390 基因座的片段长度分析谱图,表 2 为克里雅河下游的封闭人群以及其它相关人群 DYS390 基因座等位基因频率的分布。

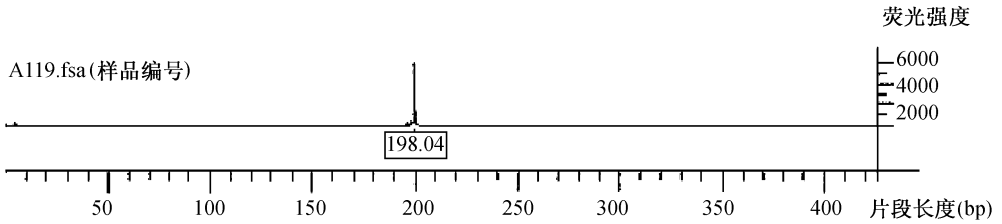


图 1 DYS19 微卫星位点片段分析图谱
Genetyper image of DYS19 microsatellite

表 1 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	n	等位基因频率 (%)							
		12	13	14	15	16	17	18	19
克里雅	51		2.0	35.3	7.8	51.0	3.9		
维吾尔族 ^[6]	56			8.9	12.5	50.0	21.4	7.1	
哈萨克族	30				10.0	23.3	66.7		
境外维吾尔 ^[7]	43		7.0	18.6	20.9	46.5	7.0		
境外哈萨克	50			8.0	8.0	84.0			
北方汉族	44	0.0	5.0	18.0	41.0	25.0	11.0		
南方汉族	50			26.0	34.0	24.0	16.0		

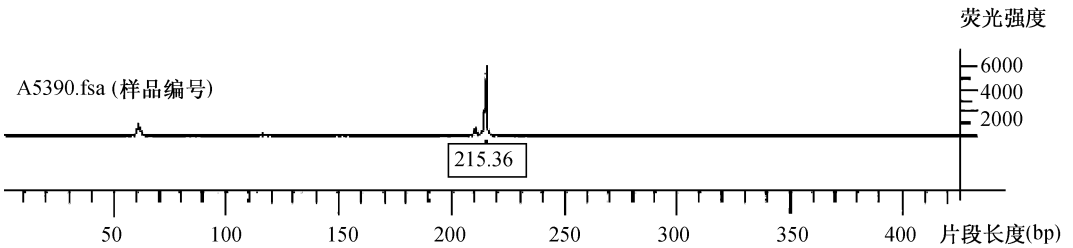


图 2 DYS390 微卫星位点片段分析图谱
Genetyper image of DYS390 microsatellite

3 讨 论

以上数据经 χ^2 检验发现: 对于 DYS19 基因座, 克里雅人群与其它所有人群都有极显著的差异 ($P < 0.01$), 而南方汉族和北方汉族之间无明显差异 ($P > 0.05$); 对于 DYS390 基因座, 同样克里雅人群与其它所有人群都有极显著的差异 ($P < 0.01$), 而东北汉族和成都汉族之间则也无明显差异 ($P > 0.05$)。

表 2 DYS390 等位基因频率在不同人群中的分布
DYS390 Allele frequencies distribution in different populations

群体	n	等位基因频率(%)								
		19	20	21	22	23	24	25	26	27
克里雅	51			23.5	9.8	5.9	43.1	17.6		
维吾尔族 ^[8]	56			1.8	5.4	19.6	42.9	28.6	1.8	
境外维吾尔	46	2.2		4.3	13	26.1	17.4	32.6	2.2	2.2
境外哈萨克	56		1.8	1.8	1.8	7.1	17.9	60.7	8.9	
蒙古族 ^[9]	30				3.3	40.0	43.3	13.3		
东北汉族	41				9.8	43.9	31.7	12.2	2.4	
成都汉族	82				3.7	48.8	32.9	11	1.8	

DYS19 微卫星位点的多态分布在不同人群中呈现出较强的人群特异性。同属中亚地区的克里雅河下游的封闭人群、我国的维吾尔族、境外的维吾尔和哈萨克都以 DYS19* 16 为主, 等位基因频率分别为 0.510、0.500、0.465 和 0.840; 而我国哈萨克族比较特殊, 以 DYS19* 17 为主, 其等位基因频率为 0.667, 而且只有三个等位基因型, 这可能与哈萨克族例数较少有关。克里雅河下游的封闭人群占第二高的基因型为 DYS19* 14, 频率达到 0.353。欧洲白人 DYS19* 14 为主, 频率高达 0.499^[10]; 非洲黑人、亚洲的东南部和东北部以及印度都以 DYS19* 15 频率最高。由于越是地处大陆边缘和岛屿以及交通险阻地区的人群越容易保持该人群的主要等位基因, 而克里雅河下游的封闭人群等位基因分布呈现“M”形分布, 以 DYS19* 14 和 DYS19* 16 最常见, 这可能是提示克里雅河下游的封闭人群的来源包含两个不同的群体分支。从历史来看, 新疆地处欧亚大陆交通之要塞, 公元前 2 世纪中国的统治者建立了从东亚到地中海的丝绸之路, 它成为东西方交流的重要通道。据体质人类学研究, 早在 3800 年以前, 分属于阿凡纳谢沃类型和安德罗诺沃类型的原始欧罗巴人种群体就已经进入罗布泊腹地, 成为目前所知最早进入新疆大地的欧罗巴人种居民^[11]。而后, 这些欧罗巴人种群体就不断的与当地的蒙古人种群体交流、同化和融合, 因此, 这也就不难解释克里雅河下游的封闭人群可能有两个不同来源祖先的原因。

对于 DYS390 基因座, 克里雅河下游地区的封闭人群、我国的维吾尔族和蒙古族均以 DYS390* 24 最多见, 频率分别为 0.431、0.429 和 0.433, 欧洲白人也以此基因型基因频率最高, 达到 0.377(12)。境外的维吾尔和哈萨克都以 DYS390* 25 为主, 我国东北汉族和成都汉族都以 DYS390* 23 为最多见。同样, 克里雅河下游的封闭人群等位基因分布也并非呈现“钟形”分布, 而是以 DYS390* 21 和 DYS390* 24 两种基因型基因频率最高, 并且 DYS390* 21 为此人群等位基因重复次数最少的基因型。此种基因分布看作如下解释: (1) 此封闭人群可能包含两个不同的群体分支; (2) 基因漂变的影响, 改变了 DYS390 这个微卫星位点在此人群中的基因频率分布; (3) 此封闭人群在形成的过程中发生了“瓶颈”现象。

个人识别力(discrimination power, DP)是遗传标记识别无关个体的能力, 也是遗传标记对于个人识别效能的定量评估, 对于一个遗传标记而言, DP 值越高, 它从同一群体中随机抽取两个个体表型不相同的概率越高, 其识别无关个体的能力就越强。DYS19 基因座在克里雅人群中的 DP 值为 0.7591, DYS390 基因座在克里雅人群中的 DP 值为 0.8936, 都属高鉴别能力的遗传标记, 因此它们对于研究克里雅人群的遗传结构以及个体识别都是非常好的遗

传标记。对克里雅河下游地区封闭人群的 DYS19 和 DYS390 两微卫星位点多态性的研究丰富了我国民族 STR 遗传资料数据库, 重构其父系进化史, 结合母系遗传的 mtDNA 分析为追溯此人群的起源和迁移提供了非常有价值的信息。

参考文献:

- [1] Mark AJ, Chris TS. New uses for new haplotypes the human Y chromosome, disease and selection[J]. TIG, 2000, 16: 356—362.
- [2] Gene M, Pique E, Borrego N *et al.* DYS19, DYS385, DYS389(I and II), DYS390, DYS391, DYS392, DYS393 haplotypes in a population sample from Catalonia (north east Spain)[J]. J Forensic Sci, 2002, 47(5): 1156—1159.
- [3] 胡文康. 走进塔克拉玛干[M]. 新疆: 新疆人民出版社, 2000, 33—48.
- [4] Kalavathy S, Marlene D, Barry W. Amplification of large DNA from blood stored at room temperature[J]. Focus, 1999, 21: 10—12.
- [5] Mark GT, Neil B, Helen MF. High throughput analysis of 10 microsatellite and 11 diallelic polymorphisms on the human Y-chromosome[J]. Hum Genet, 1999, 105: 577—581.
- [6] 包维东, 徐玖瑾, 朱苏玲. 微卫星位点 DYS19 在中国人群中的多态研究[J]. 人类学学报, 1997, 16(4): 255—261.
- [7] Anna EL, Francisc C, David C *et al.* Sex specific migration patterns in Central Asian populations, revealed by analysis of Y-Chromosome short tandem repeats and mtDNA[J]. Am J Hum Genet, 1999, 65: 208—219.
- [8] 卢岩, 王保捷, 丁梅. 中国新疆维吾尔族群体 DYS389II, DYS390 基因座及单倍型遗传多态性分析[J]. 中国法医学杂志, 2000, 15(3): 151—152.
- [9] 滕百军, 陈峰, 傅松滨. 我国东北地区 3 个群体 DYS390 多态位点的遗传学研究[J]. 人类学学报, 2000, 19(4): 318—323.
- [10] Knijff P, Kayser M, Caglia A *et al.* Chromosome Y microsatellites: population genetic and evolutionary aspects[J]. Int J Legal Med, 1997, 110(3): 134—149.
- [11] 余太山. 西域通史[M]. 郑州: 中国古籍出版社, 1996.

GENETIC POLYMORPHISMS OF DYS19 AND DYS390 IN AN “INSULATED” POPULATION IN THE LOWER REACHES OF KERIYAN RIVER

DUAN Ran hui¹, LIU Wei qiang¹, ZHOU Hu¹, ZHU Hong¹

(1. Laboratory of ancient DNA, College of life science, Jilin University, Changchun 130012)

Abstract: Two Y-linked short-tandem repeat polymorphisms (DYS19 and DYS390) were analyzed by Genescan in an “insulated” population in the center of the Taklamakan desert of Xinjiang province. And 51 unrelated male individuals were also studied. The allelic distribution of DYS19 in this Keriyan population shows “M” shape. DYS19* 14 and DYS19* 16 are the most common allele, frequencies of which are 0.353 and 0.510, respectively. As DYS390* 21 owns the least repeat number and DYS390* 24 are the most common allele. This study suggests that the ancestor of Keriyan population is derived from two distinct origins.

Key words: Y-Chromosome ; “Insulated” population; Short-tandem repeat; Polymorphism