

DOI: 10.3724/SP.J.1005.2012.01020

福建畲族群体 17 个 Y-STR 基因座单倍型及遗传关系

百茹峰¹, 杨利海², 袁丽¹, 梁权赠¹, 鲁涤¹, 杨雪¹, 石美森¹

1. 中国政法大学证据科学教育部重点实验室, 北京 100040;
2. 北京市石景山区五里坨医院, 北京 100042

摘要: 应用Y-filer™试剂盒及基因分型技术, 检测 152 份福建畲族无关男性个体 17 个 Y-STR 基因座的多态性分布, 计算等位基因频率及单倍型多样性, 并结合已公开发表的其他 11 个群体相应基因座的单倍型资料, 分析福建畲族群体遗传距离和聚类关系。福建畲族 *DYS385a/b* 基因座检出 50 种单倍型, 其余 15 个 Y-STR 基因座分别检出 3~11 个等位基因, 基因多样性GD值在 0.4037(*DYS391*)~0.9725(*DYS385a/b*); 观察到 *DYS19* 和 *DYS390* 基因座双等位基因和 *DYS385a/b* 基因座三等位基因, 以及 *DYS448* 等部分基因座出现的“off-ladder”等位基因现象。17 个Y-STR 基因座共同构成的单倍型 144 种, 其中 138 种单倍型出现 1 次, 5 种出现 2 次, 1 种出现 4 次, 累计GD值为 0.9990。从遗传距离分析发现, 福建畲族与浙江汉族之间的遗传距离最近(0.0042), 与青海藏族(0.2378)之间的遗传距离相对较远。福建畲族最靠近由台湾群体、浙江汉族、南方汉族等典型南方汉族群体聚成的分支区域。结果表明该 17 个Y-STR 基因座在福建畲族群体中具有丰富的遗传多态性, 对建立Y染色体STR数据库, 研究群体遗传学和进行法医学应用有重要意义。

关键词: Y 染色体; 短串联重复序列; 单倍型; 遗传距离; 福建畲族

Polymorphism of 17 Y-STR loci in She ethnic population in Fujian and genetic relationship with 11 populations

BAI Ru-Feng¹, YANG Li-Hai², YUAN Li¹, LIANG Quan-Zeng¹, LU Di¹, YANG Xue¹, SHI Mei-Sen¹

1. Key Laboratory of Evidence Science(China University of Political Science and Law), Ministry of Education, Beijing 100040, China;
2. Wulituo Hospital of Shijingshan District, Beijing 100042, China

Abstract: To investigate the genetic polymorphisms of 17 Y-chromosomal short tandem repeats(Y-STR) loci in She ethnic population from Fujian province, and to evaluate their forensic application values and genetic relationship with other 11 populations, 152 unrelated male individuals of She ethnic population in Fujian were used to determine the distribution of allele frequencies and haplotypes by using Y-filer™ System. Cluster analysis and phylogenetic trees were applied to show the genetic distance among the populations. As a result, 50 haplotypes were found in *DYS385a/b* loci, and 3~11 al-

收稿日期: 2012-02-02; 修回日期: 2012-04-24

基金项目: 教育部长江学者和创新团队发展计划资助项目(证据科学研究与应用)(编号:IRT0956), 国家自然科学基金面上项目(编号:81172902), 北京市科技新星计划(编号:2010B073)和中国政法大学青年教师学术创新团队资助项目

作者简介: 百茹峰, 副教授, 研究方向: 群体遗传学, 法医损伤学。E-mail: brf1000cn@yahoo.com.cn

通讯作者: 石美森, 副教授, 研究方向: 法医物证学、群体遗传学。E-mail: shimeisen2000@yahoo.com.cn

leles were found in the rest 15 Y-STR loci. The GD value was from 0.4037(*DYS391*) to 0.9725(*DYS385a/b*). This study has also revealed “off-ladder” alleles at several Y-loci, namely *DYS448*, *DYS393*, *DYS458* and *DYS635*, and several occurrences of duplications at the *DYS385a/b*, *DYS19* and *DYS390* loci. One hundred and forty-four haplotypes were found in 17 Y-STR loci, of which 138 were unique, 5 were found in 2 individuals, 1 was found in 4 individuals, and the observed haplotypes diversity value was 0.9990. Comparing with 11 populations, the genetic distance between She ethnic and Han population in Zhejiang was the smallest (0.0042), while it was the largest between She ethnic and Tibet ethnic population (0.2380). Cluster analysis and phylogenetic tree both demonstrated that genetic distance between She ethnic and several south Han populations is closer than between She ethnic and non-Han populations. Multiplex detection of the 17 Y-STR loci revealed a highly polymorphic genetic distribution, which would be very powerful for establishing a Y-STR database, for population genetics and forensic practice.

Keywords: Y-chromosome; short tandem repeats; haplotype; genetic distance; She ethnic population in Fujian

Y染色体短串联重复序列(Short tandem repeats, Y-STR)除拟常染色体区域外,在减数分裂过程中不发生重组,呈稳定的单倍型父系遗传特点。因此,选择多个 Y-STR 构建合适的单倍型,探讨单倍型在不同人群中的分布规律,在人类的起源进化、群体划分、种族民族差异和法医学领域具有极其重要的意义^[1-5]。畚族是福建省特有的少数民族群体,迄今有关他们 Y-STR 基因座的研究尚未见报道,本文选择包含 17 个 Y-STR 基因座的 Y-filer™ 试剂盒(*DYS19*、*DYS389I*、*DYS389II*、*DYS390*、*DYS391*、*DYS392*、*DYS393*、*DYS437*、*DYS438*、*DYS439*、*DYS385a/b*、*DYS448*、*DYS456*、*DYS458*、*DYS635*、*Y-GATA-H4*),对 152 份福建畚族无关男性个体血样进行单倍型调查,结合最近已发表的国内 11 个群体的相同基因座研究资料,运用多种统计学分析方法为该群体的法医学应用、遗传关系的分析提供基础数据。

1 材料和方法

1.1 研究对象

152 份福建省福州地区无血缘关系畚族男性个体血样,在知情同意的原则下,每个供血个体均追溯 3 代以上家族史,以确定其民族代表性。5% Chelex-100 快速提取法提取 DNA^[6]。

以获得的青海藏族、南方汉族、浙江汉族、广西瑶族、广西彝族、广西京族、广西壮族、宁夏回族、青海撒拉族、山西汉族及台湾群体等 11 个群体

17 个 Y-STR 单倍型(共 1 741 个样本)作为福建畚族群体遗传距离比较的数据^[7-14]。所有文献的样本采集遵循以下原则:3 代以内均为同一民族的无关个体,并世代居住于该地区,每个地区民族的样本采集数均大于 90 例。

1.2 方法

1.2.1 PCR 扩增及毛细管电泳检测

采用 Y-filer™ 试剂盒,扩增使用 PE9700 型扩增仪,12.5 μL 扩增体系,含 Y-filer PCR Reaction mix 4.6 μL,引物 2.5 μL,Ampli Taq Gold™ DNA (5 U/μL) 0.4 μL,模板 DNA 溶液 5 μL。PCR 热循环条件为:95 预变性 11 min,30 个循环中每循环 94 变性 1 min,61 复性 1 min,72 延伸 1 min;最后 60 延伸 80 min,4 保存。采用 3130 全自动遗传分析仪进行扩增片段的电泳分离与检测, GeneMapper IDv 3.1 软件分型。

1.2.2 数据统计分析

用直接计数法计算各基因座的等位基因与单倍型检出频率,基因多样性(Gene diversity, GD)和单倍型多样性按公式 $GD = n(1 - \sum P_i^2) / (n-1)$ (P_i 为等位基因或单倍型频率, n 为样本数)计算。采用 Arlequin v 3.1 软件进行 AMOVA 分析,计算遗传距离 R_{st} 矩阵;根据净遗传距离用 PHYLIP 3.62、Treeview 软件重建相邻连接(Neighbour-joining, NJ)系统发生树^[15],用 STATISTICA 7.0 软件进行多维尺度(Multidimensional scaling, MDS)分析构建民族群体

二维分布图。

2 结果与分析

2.1 福建畬族男性群体 17 个 Y-STR 基因座遗传多态性

在 152 份福建畬族无关男性个体血样中, Y-filer™ 试剂盒中的各基因座都得到了有效扩增, 且各基因座间扩增产物平衡, 扩增片段长度均在 100~340 bp 之间。除 *DYS385a/b* 基因座检出 50 种单倍型外, 其余 15 个基因座分别检出 3~11 个等位基因, 各等位基因频率分布在 0.0066~0.7500 之间, 基因多样性 GD 值在 0.4037(*DYS391*)~0.9725(*DYS385a/b*) 之间。由 17 个 Y-STR 基因座组成的 YH 系统单倍型共检出 144 种(表 1), 其中 138 种单倍型均出现 1 次, 5 种单倍型出现 2 次, 1 种单倍型出现 4 次, 累计 GD 值为 0.9990。

2.2 福建畬族与国内 11 个群体间的遗传距离及聚类分析

应用本文获得的福建畬族群体 *DYS19*、*DYS389I*、*DYS389II*、*DYS390*、*DYS391*、*DYS392*、*DYS393*、*DYS437*、*DYS438*、*DYS439*、*DYS385a/b*、*DYS448*、*DYS456*、*DYS458*、*DYS635*、*Y-GATA-H4* 单倍型数据, 计算与国内 11 个群体的遗传距离(表 2)。结果表明, 福建畬族与浙江汉族之间的遗传距离最小(0.0042), 与青海藏族(0.2378)之间的遗传距离相对最大。福建畬族除与浙江汉族不具有显著的遗传学差异外, 与其他 10 个群体之间均具有显著的遗传学差异($P < 0.05$)。基于遗传距离所构建的 MDS 和系统发生树基本相符, 见图 1、图 2。福建畬族和台湾群体、浙江汉族和南方汉族聚为一大类; 广西壮族和广西京族、广西瑶族聚为一大类; 宁夏回族、青海撒拉族、广西彝族以及青海藏族聚为一大类。

3 讨论

3.1 福建畬族群体 17 个 Y-STR 基因座的遗传多态性

遗传多态性结果显示, *DYS456*、*DYS389I*、*DYS390*、*DYS389II*、*DYS 458*、*DYS19*、*DYS393*、*DYS391*、*DYS439*、*DYS635*、*DYS392*、*Y-GATA-H4*、*DYS437*、*DYS438*、*DYS448* 基因座分别检出 6、4、

6、9、9、6、5、4、7、9、10、4、3、5、9 个等位基因; *DYS385a/b* 基因座检出 50 种单倍型; 等位基因频率分布在 0.0066~0.7500 之间, 基因多样性在 0.4037(*DYS391*)~0.9725(*DYS385a/b*) 之间, 除了 *DYS391*、*DYS437* 和 *DYS438*, 其余基因座的 GD 值均大于 0.5。17 个 Y-STR 基因座共同构成 144 种单倍型, 单倍型多样性为 0.9990, 系统鉴定能力达到 94.74%, 在福建畬族群体中具有较好的遗传多态性, 适合法医学应用。

按照福建畬族群体 *DYS19*、*DYS389I*、*DYS389II*、*DYS390*、*DYS391*、*DYS392*、*DYS393*、*DYS385a/b*、*DYS438*、*DYS439*、*DYS437*、*DYS448*、*DYS456*、*DYS458*、*DYS635*、*Y-GATA-H4* 单倍型数据, 在包含 766 个群体的 101 055 个单倍型数据库(www.yhrd.org)(最新更新日期 02/17/2012)中检索。8.33%(12 种)的福建畬族单倍型全部和亚洲群体发生共享; 而 91.67%(132 种)的单倍型为零共享, 是具有福建畬族特征性的单倍型群。

在本研究过程中还观察到 2 例双等位基因: 单倍型 H41 的 *DYS19* 基因座(16/18), 单倍型 H51 的 *DYS390* 基因座(25/26); 1 例三等位基因: 单倍型 H130 的 *DYS385a/b* 基因座(12/13/17); 此外, 有部分样本在个别基因座上出现超出标准等位基因阶梯现象, 如 *DYS458* 分别检出等位基因 13 和 21, *DYS635* 分别检出等位基因 18 和 19, *DYS393* 检出等位基因 17, *DYS448* 检出等位基因 16, 对此吴微微等也有报道^[9,16~19]。结果显示 Y-STR 基因座的多态性分布具有显著的种族、地域差异, 建立本地区本民族的等位基因频率分布资料是法医学应用的前提与基础。

3.2 福建畬族与其他群体遗传关系的分析

畬族是我国具有悠久历史的少数民族, 主要分布在福建、浙江、广东、江西、安徽等省丘陵山区。2000 年全国人口普查时, 全国畬族有 73 万人。畬族自称“山哈”。对于畬族族源, 学术界尚未定论, 但普遍认为唐代闽、粤、赣三省交界地为畬族先民的主要活动区域。畬族与汉族人民由于长期交错杂居, 在政治、经济、文化上都有密切联系, 畬族的生产生活水平与当地汉族日渐接近。只是在语言、生活习俗和宗教信仰上, 还保留有本民族的特点。

表 2 12 个群体 Rst 值遗传距离矩阵(对称轴上是对应的 P 值, 对称轴下是 Rst 值)

	福建畚族	青海藏族	南方汉族	浙江汉族	广西瑶族	广西彝族	广西京族	广西壮族	宁夏回族	青海撒拉族	台湾群体	山西汉族
福建畚族	0.0000	+	+	-	+	+	+	+	+	+	+	+
青海藏族	0.2378	0.0000	+	+	+	+	+	+	+	+	+	+
南方汉族	0.0172	0.2813	0.0000	+	+	+	+	+	+	+	+	+
浙江汉族	0.0042	0.2587	0.0106	0.0000	+	+	+	+	+	+	+	+
广西瑶族	0.1593	0.4296	0.1924	0.2085	0.0000	+	+	+	+	+	+	+
广西彝族	0.1567	0.1644	0.2034	0.1901	0.3815	0.0000	+	+	+	+	+	+
广西京族	0.0501	0.3147	0.0226	0.0540	0.1467	0.2091	0.0000	+	+	+	+	+
广西壮族	0.0560	0.2934	0.0459	0.0715	0.1968	0.1385	0.0304	0.0000	+	+	+	+
宁夏回族	0.0644	0.1042	0.0843	0.0701	0.2639	0.0938	0.1343	0.1130	0.0000	+	+	+
青海撒拉族	0.0906	0.1134	0.1121	0.1041	0.2721	0.0719	0.1457	0.1142	0.0093	0.0000	+	+
台湾群体	0.0108	0.2523	0.0149	0.0045	0.2430	0.1671	0.0626	0.0645	0.0655	0.0979	0.0000	+
山西汉族	0.0192	0.2289	0.0137	0.0152	0.1275	0.1364	0.0324	0.0475	0.0545	0.0714	0.0306	0.0000

注：“+”：P < 0.05；“-”：P > 0.05。

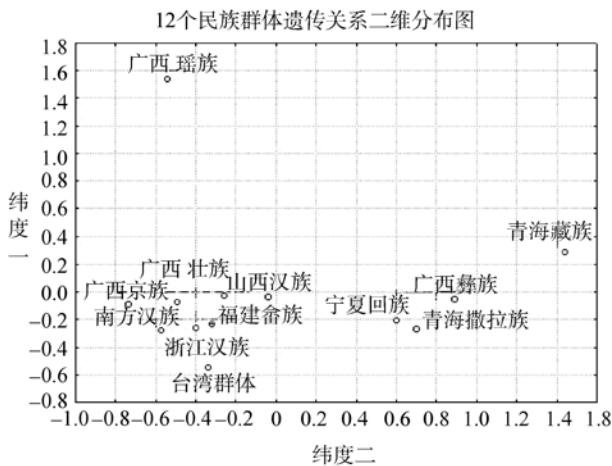


图 1 12 个群体 MDS 分析图

遗传距离为表示群体间或物种间遗传差异或遗传分化的最重要参数。本研究构建的 *Rst* 遗传距离矩阵显示, 福建畚族与浙江汉族之间的遗传距离最近(0.0042), 其次与台湾群体、南方汉族之间的遗传距离较近(0.0108、0.0172), 与青海藏族之间的遗传距离相对最远(0.2378)。基于遗传距离所构建的 *MDS* 和系统发生树也说明了相同的结果, 与其他 7 个少数民族比较, 福建畚族最靠近由台湾群体、浙江汉族、南方汉族等典型南方汉族群体聚成的分支区域。其原因可能是在明清时期, 畚族大量迁往汉族地区并形成了目前的“大分散, 小聚居”分布格局, 这个时期畚族地区的社会经济生活相对比较安定, 畚族的迁移基本上不带军事性质, 而主要是基于畚

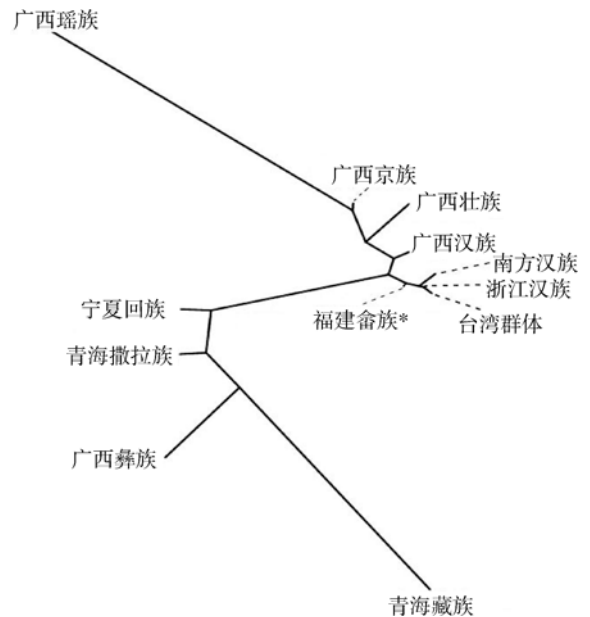


图 2 12 个群体 NJ 系统发生树

族山地游耕的传统; 福建畚族多迁至广东, 浙江、江西畚族多迁至福建, 安徽畚族多迁至浙江, 湖南、贵州畚族多迁至江西; 本研究中的福建畚族群体与这些地方的汉族人群在大规模、频繁的人口迁移过程中相互间基因交流较多所致。而其他少数民族与汉族群体及其他群体间遗传距离差异明显, 可能与其民族特有的封闭环境、高度聚居、族内通婚及长期相同的生活环境形成了自己特有的体质特征和遗传基因。总体上看, 17 个 Y-STR 基因座组成的单倍

型计算的遗传距离结果与福建畲族群体形成历史较一致, 说明了 Y 染色体遗传标记在民族群体间遗传距离和基因漂流的评价中起到一定作用。

本研究获得了福建畲族群体 *DYS19*、*DYS389I*、*DYS389II*、*DYS390*、*DYS391*、*DYS392*、*DYS393*、*DYS437*、*DYS438*、*DYS439*、*DYS385a/b*、*DYS448*、*DYS456*、*DYS458*、*DYS635*、*Y-GATA-H4* 等 17 个基因座等位基因及单倍型频率数据, 为建立福建畲族群体分布数据库, 为该地区法医学应用及遗传关系的分析提供了良好的遗传背景数据。

参考文献(References):

- [1] Jobling MA, Pandya A, Tyler-Smith C. The Y chromosome in forensic analysis and paternity testing. *Int J Legal Med*, 1997, 110(3): 118–124. [DOI](#)
- [2] de Knijff P, Kayser M, Caglia A, Corach D, Fretwell N, Gehrig C, Graziosi G, Heidorn F, Herrmann S, Herzog B, Hidding M, Honda K, Jobling M, Krawczak M, Leim K, Meuser S, Meyer E, Oesterreich W, Pandya A, Parson W, Penacino G, Perez-Lezaun A, Piccinini A, Prinz M, Roewer L. Chromosome Y microsatellites: population genetic and evolutionary aspects. *Int J Legal Med*, 1997, 110(3): 134–149. [DOI](#)
- [3] Stumpf MP, Goldstein DB. Genealogical and evolutionary inference with the human Y chromosome. *Science*, 2001, 291(5509): 1738–1742. [DOI](#)
- [4] Parson W, Niederstätter H, Köchl S, Steinlechner M, Berger B. When autosomal tandem repeats fail: optimized primer and reaction design for Y-chromosome short tandem repeat analysis in forensic casework. *Croat Med J*, 2001, 42(3): 285–287. [DOI](#)
- [5] Sibille I, Duverneuil C, Lorin de la Grandmaison G, Guerrouache K, Teissière F, Durigon M, de Mazancourt P. Y-STR DNA amplification as biological evidence in sexually assaulted female victims with no cytological detection of spermatozoa. *Forensic Sci Int*, 2002, 125(2–3): 212–216. [DOI](#)
- [6] Walsh BS, Petzger DA, Higuchi R. Chelex-100 as medium for simple extraction of DNA for PCR-based typing from forensic materia. *Biotechniques*, 1991, 10(4): 506–513. [DOI](#)
- [7] Zhu BF, Wu YM, Shen CM, Yang TH, Deng YJ, Xun X, Tian YF, Yan JW, Li T. Genetic analysis of 17 Y-chromosomal STRs haplotypes of Chinese Tibetan ethnic group residing in Qinghai province of China. *Forensic Sci Int*, 2008, 175(2–3): 238–243. [DOI](#)
- [8] 陈悦康, 李茜, 李大成, 邓志辉. 南方汉族无关男性群体 17 个 Y-STR 基因座单倍型遗传多态性的研究. *实验与检验医学*, 2008, 26(4): 351–354, 386. [DOI](#)
- [9] 吴微微, 郑小婷, 潘立鹏, 郝宏蕾, 傅汀. 浙江汉族人群 16 个 Y-STR 基因座遗传多态性调查. *刑事技术*, 2005, 179(5): 11–17. [DOI](#)
- [10] 冯冬亮, 刘长晖, 梁祚仁, 刘超. 广西 4 个少数民族 17 个 Y-STR 基因座的多态性分析. *遗传*, 2009, 31(9): 921–935. [DOI](#)
- [11] Guo H, Yan JW, Jiao ZP, Tang H, Zhang QX, Zhao L, Hu N, Li HF, Liu YC. Genetic polymorphisms for 17 Y-chromosomal STRs haplotypes in Chinese Hui population. *Leg Med*, 2008, 10(3): 163–169. [DOI](#)
- [12] Zhu BF, Shen CM, Xun X, Yan JW, Deng YJ, Zhu J, Liu Y. Population genetic polymorphisms for 17 Y-chromosomal STRs haplotypes of Chinese Salar ethnic minority group. *Leg Med*, 2007, 9(4): 203–209. [DOI](#)
- [13] 石美森, 百茹峰, 傅博. 山西汉族 17 个 Y-STR 基因座遗传多态性及遗传关系. *遗传*, 2011, 33(3): 228–238. [DOI](#)
- [14] Huang TY, Hsu YT, Li JM, Chung JH, Shun CT. Polymorphism of 17 Y-STR loci in Taiwan population. *Forensic Sci Int*, 2008, 174(2–3): 249–254. [DOI](#)
- [15] Felsenstein J. PHYLIP: phylogeny inference package (version 3. 2). *Cladistics*, 1989, 5: 164–166. [DOI](#)
- [16] 鲁滌, 袁丽, 杨雪. 北京汉族群体 17 个 Y-STR 基因座遗传多态性. *中国法医学杂志*, 2010, 25(6): 432–433. [DOI](#)
- [17] 王永在, 王勇, 黄太宇, 冬日布, 胡日查, 徐艳丽. 内蒙古汉族人群 17 个 Y-STR 基因座遗传多态性. *中国法医学杂志*, 2009, 24(2): 117–118. [DOI](#)
- [18] 李斌, 吕政, 王澍, 黄建春, 卢水发. 福建汉族人群 17 个 Y-STR 基因座遗传多态性调查. *刑事技术*, 2008, (2): 60–62. [DOI](#)
- [19] 郭利红, 田曾元, 刘亚举, 刘海, 李效阳, 张毅. 河南汉族人群 16 个 Y-STR 基因座遗传多态性. *法医学杂志*, 2010, 26(1): 66–67. [DOI](#)