



常染色体 STR 的证据支持 潮汕人的中原汉族起源

许丽娜¹, 冯国英¹, 周广彪^{1,2}, 胡盛平^{1,2}

1 汕头大学医学院分子生物学中心, 广东汕头 515031; 2 汕头大学司法鉴定中心 DNA 物证研究室, 广东汕头 515031

摘要: 潮汕人群是典型的中国南方汉族群体, 并且大规模移居海外。一般认为潮汕人群起源于河南一带。在 DNA 水平上比较潮汕和河南人群的遗传多态性, 可以深入探究潮汕人的中原汉族起源。我们采用荧光标记 PCR 和毛细管电泳相结合的技术对河南汉族 200 例健康人群的 15 个 STR 基因座遗传多态性进行研究, 结合本课题组已发表潮汕人群 STR 数据并新增样本以及搜集到的发表在国内外刊物的其它地区人群 STR 数据, 用 PhyIip 软件计算不同人群间的遗传距离, 构建高分辨率进化树, 探讨潮汕人群的中原汉族起源及与其他人群之间的关系。河南汉族人群 15 个 STR 基因座共检出 142 个等位基因 451 种基因型, 其观察值与期望值经精确检验, 均无显著性差异 (D3S1358 除外)。选取 CODIS 系统 STR 基因座的人群数据聚类分析结果显示潮汕汉族闽南汉族很近, 距河南汉族较近; 融合分析显示潮汕和闽南这两个群体具有相似的中原汉族成分且中原汉族的成分所占比例均大于 60%。聚类图上中国人群的南北分界明显。河南汉族人群 15 个 STR 基因座联合应用可满足河南人群法医学的个体识别及亲子鉴定需要; 潮汕和闽南这两个群体具有一贯的遗传相似性, 且占有较大的中原汉族起源成分。

关键词: 群体遗传学; 短串联重复序列; 潮汕人群; 河南人群; 聚合酶链式反应

Evidence from Autosomal STR Data Supports a Central China Origin of Chao-Shan Population

XU Lina¹, FENG Guoying¹, ZHOU Guangbiao¹, HU Shengping^{1,2}

1.Center for Molecular Biology, 2.Forensic Genetics Laboratory, Shantou University Medical College, Shantou, Guangdong 515031 China

ABSTRACT: Chao-Shan people are typical Southern Han Chinese, and numbers of them emigrate to many other countries. The origin of Chao-Shan people is believed to be in Central China. The comparative study of the genetic polymorphism of Henan and Chao-Shan populations can test the hypothesis of Central China origin of Chao-Shan people at the genomic level. Here we performed a combined method of fluorescence labeled PCR and capillary electrophoresis to study the genetic polymorphism of 15 STR loci in 200 unrelated healthy individuals in Henan Han population. Combining our published STR allele frequency data and the new data from additional samples in Chao-Shan population and STR data of other populations in different areas obtained from peer-reviewed journals, we calculated the Nei's genetic distance D' between the populations with the PHYLIP software. The neighbor-joining phylogenetic tree was constructed based on the obtained Nei's genetic distance to reveal population affinities and the Central China origin of Chao-Shan people. A total of 142 alleles and 451 genotypes were detected in Henan population. No deviation from Hardy-Weinberg equilibrium was observed in STR loci except for D3S1358. The phylogenetic tree based on CODIS STR allele frequencies showed that Chao-Shan Han was much closer to Minnan Han, and relatively closer to Henan Han. Admixture analysis indicated that Chao-Shan Han and Minnan Han have similar proportional contributions from the Central China Han and the proportion was greater than 60%. Distinction between north and south populations in China was noticeable in the phylogenetic tree. The 15 STR loci studied have potential usefulness for parental testing and forensic personal identification in Henan population. Our study showed the genetic similarity sustained by Chao-Shan Han and Minnan Han, and a pronounced Central China Han ancestry on the gene pools of the two groups.

Key words: Population genetics; Short tandem repeat (STR); Chao-Shan population; Henan population; Polymerase chain reaction (PCR)

收稿日期: 2008 年 6 月 23 日 **修回日期:** 2008 年 6 月 30 日 **联系人:** 胡盛平 sphu@stu.edu.cn

作者简介: 许丽娜(1982-), 女, 群体遗传学硕士; Tel: 0754-8900406; E-mail: lnxu2005@163.com.

潮汕僻处我国的东南隅,广东省的最东端,与福建省毗邻,其地域包括今天的汕头、潮州、揭阳三个地级市。潮汕人指居住于该区域并有自己独特方言和生活习惯的汉族民系。据记载,现代潮汕人源自古中原汉族[1],主要是自秦汉以来大量的中原汉族移民与南方土著居民杂居后形成的[2],其中百越族是东南沿海人口最多,分布最广的土著民族[3]。我们[4]前期应用序列特异引物-聚合酶链反应(PCR-SSP)对潮汕汉族进行 HLA-A、B 基因分型,计算等位基因和单倍型频率,并与包括大陆闽南、客家和中部(长江流域)群体在内的其他 9 个汉族人群相应位点的遗传多态性分布进行比较。结果显示潮汕人与大陆闽南人亲缘关系最近,中部群体大致位于南方相关汉族群体和北方群体的中间位置。然而,所得结果尚不能为潮汕人的中原汉族起源提供遗传学上的肯定证据。我们认为采用更为敏感且能反映群体亲缘关系、分化、进化以及人群特征的遗传标记可能有助于研究的深入。

在众多遗传标记中,短串联重复(STR)以其高度多态性、分布广泛等优点被广泛应用于人类群体遗传学研究中。Nei M 等[5]用实验的方法表明,相比对遗传关系较远的群体之间的关系分析,对于关系较近的群体,STR 基因座的分析能提供更可靠的遗传关系。一般认为,STR 是研究较近群体的理想遗传标记。通过研究 STR 多态性、变异以及比较序列间差异、人群间差异,分析不同人群间的遗传距离,构建高分辨率进化树,就可从分子生物学角度揭示人类的起源、迁移、进化等历史进程。这些研究在国内外均有广泛开展。我国学者褚嘉佑等[6]利用 30 个常染色体 STR 基因座分析中国 28 个人群遗传结构并进行进化树聚类分析,证实了中国南北方汉族人群间的遗传差异;李生斌等[7]利用 6 个常染色体 STR 基因座分析中国广西地区 15 个不同民族人群,结果显示广西地区 15 个不同民族人群之间存在一定的基因交流,个别民族相对独立;李辉等[8]对百越族群历史上重要的一支闽越族群进行研究,利用 Y 染色体 STR 对现代福建和其他闽语人群进行遗传学分析,研究表明现代闽语人群基本都是来源于北方的汉族移民,历史上的闽越族在福建地区基本上已经消失。

一、材料和方法

1. 研究对象

根据知情同意原则,从健康体检人群中

随机采集出生于河南省遂平县并 3 代以上直系亲属居住于当地的 201 名无关个体外周血,做成血纱布,阴干,室温保存。

2. 人群数据搜集

为了分析潮汕人群与其他人群的关系,通过公用数据库或同行评议期刊等途径获取其他人群的 STR 基因座频率数据,人群数据的筛选标准是:1,截至 2007 年 8 月 30 日以前发表在同行评议期刊上的人群数据;2,华人(大陆中国人和海外华人)汉族;3,STR 频率数据且涵盖的 STR 基因座个数达到或总和达到 13 个(D8S1179, D21S11, D7S820, CSF1PO, D3S1358, THO1, D13S317, D16S539, VWA, TPOX, D18S51, D5S818 和 FGA);4,来自人群的 STR 数据既包括以行政区划(省、直辖市、自治区、特别行政区、省辖市)进行界定的数据,也包括以民系进行界定(如潮汕人、闽南人)的数据;5,无样本数限制;6,排除无基因频率或频率数据有误、不完全的人群数据或某一个 STR 基因座等位基因之和不等于 1 的人群数据。按照此筛选标准我们共选择了 38 个汉族人群数据作为比较数据,其中有 13 个汉族人群有不止一套的 STR 频率数据,针对这一情况,首先对这 38 个汉族人群进行初筛,构建遗传发育树(见统计分析部分)。对于有两套或两套以上基因频率数据的相同人群,在树状图上,若这些人群聚在同一小支上彼此非常接近且 bootstrap 值大于 80%,则对这些人群的基因频率数据用加权平均法进行合并作为该地区的人群的最终数据,合并公式为 $p_i = \sum p_{ij}n_j / \sum n_j$,其中, p_i 为合并后人群的第 i 个等位基因的频率, p_{ij} 为第 j 个相同人群的第 i 个等位基因的频率, n_j 为第 j 个相同人群的样本数。若这些人群聚在同一小支上但 bootstrap 值小于 80%,或者人群聚在不同的分支上且彼此相距很远则选择其中一个最佳的人群数据作为该地区人群的最终数据,即:首先考虑文献中材料和方法部分对样本搜集的描述,包括样本来源地、是否三代以上都在本地区居住、是否按照知情同意的原则;其次考虑选择样本例数最大的人群的频率数据。需要指出的是,用于分析比较的潮汕人群频率数据是在我们原来发表的 15 个 STR 频率数据[9]基础上增加了标本量(数据尚未发表)后统计的数据。河南汉族人群和闽南汉族人群频率数据选自本室尚未发表和已经发表的研究数据。最终,共用了 25 个汉族群体的基因频率数据作为比较数据。所用的汉族 25 个人群及地理分布详见表 1 及图 1。

表 1 本研究所用汉族人群

Tab.1 Han Chinese populations analyzed in the current study

地区	语支	方言	参考文献
北京	秦	北京官话	[5]
河南	秦	中原官话	本研究
安徽	秦	江淮官话	[6]
甘肃	秦	兰银官话	[7]
陕西西安	秦	中原官话	[8,9]
吉林	秦	东北官话	[10]
辽宁辽阳	秦	东北官话	[11]
陕西	秦	中原官话	[12]
陕西延安	秦	中原官话	[13]
山东青岛	秦	胶辽官话	[14]
香港	秦	粤语	[15,16]
新加坡	齐	闽南话	[17]
四川成都	秦	西南官话	[18]
四川	秦	西南官话	[19]
广东	秦	粤语	[20]
台湾	齐	闽南话	[21]
广东潮汕	齐	闽南话	[4]*
福建闽南	齐	闽南话	[22]
浙江	楚	吴语	[23]
江苏盐城	秦	江淮官话	[24]
江苏	秦	江淮官话	[25]
湖南	楚	湘语	[26]
江西	楚	赣语	[27]
广东深圳	秦	粤语	[28]
马来西亚	齐	闽南话	[29,30]

* 该频率数据是在我们原来发表的 15 STR 频率数据基础上增加了标本量（数据尚未发表）后统计的数据。



图 1 汉族 25 个人群的地理分布图

Fig.1 Geographic distribution of 25 Han Chinese populations

3. STR 基因分型

Chelex-100 法[36]提取血样 DNA，溴化乙锭荧光法测 DNA 含量。使用 AmpFLSTR Identifiler 16 STR 扩增试剂盒（ABI 公司），总反应体系改为 8μL[9,27]，其中 PCR Reaction Mix 3μL, Primer Set 1.8μL, Gold Taq DNA polymerase 0.25μL (5U/μL)，DNA 1-2μL，补水至 8μL，扩增条件与试剂盒说明书保持一致。

取 Hi-Di 甲酰胺 8.7μL, GS-500 LIZ 分子内标 0.3μL, PCR 产物 1μL，混匀，用 ABI Prism 3100 遗传分析仪电泳，GeneScan 和 Genotyper 软件进行基因分型。

4. 统计分析

采用 PowerStatesV12 软件 (www.promega.com)对数据进行分析处理，获得各基因座的等位基因频率、观察值的杂合度 (H_o)、个体识别率 (DP)、非父排除率 (PE) 和多态信息含量 (PIC)；期望杂合度 (H_e) 按 $H_e = (n/(n-1))(1-\sum(p_i^2))$ 计算，其中 n 是样本数， p_i 为第 i 个等位基因频率；用 PyPop 软件精确检验法[37]进行 Hardy-Weinberg 平衡检验；用 PHYLIP 3.65 软件[38]计算群体间遗传距离 D 值[39]并构建比较人群的相邻连接 (Neighbor-joining, NJ)遗传树[40]，bootstrap 法[41]评估树状图的可靠性，用 LEADMIX 软件[42]计算 M_{RH} 统计值[43]，进行融合分析。

二、结果

1. 不同地区汉族人群比较

基于 25 个人群 CODIS STR 等位基因频率数据计算 D 值并构建 NJ 遗传树 (图 2)，在树状图上，汉族基本上聚为南北两大类，南北分界明显，这与褚嘉佑[6]、Shaw[44]、杜若甫[45]和肖春杰[46]等的研究报道一致。深入分析树状图的拓扑结构可以看出河南汉族大致位于南北汉族的中间位置。来自南方的汉族人群（广东汉族、香港汉族、江西汉族、新加坡汉族、马来汉族、台湾汉族、成都汉族、四川汉族、潮汕汉族、闽南汉族、湖南汉族、浙江汉族）彼此分散聚在同一支。他们中间，潮汕汉族与闽南汉族很接近，尽管潮汕地区位于广东省境内，潮汕汉族与广东汉族的亲缘程度还是比较远，相反，相对于广东汉族，潮汕汉族和闽南汉族与河南汉族的亲缘关系要近一些。北方汉族族群彼此聚在一起，可以看出有两个大的分支，青岛汉族与江苏汉族首先聚合再与吉林汉族聚类组成第一个大分支，陕西汉族与延安汉族聚合再与盐城汉族聚到一起形成另一大分

支。这两大分支最后与甘肃汉族、安徽汉族和辽阳汉族聚类。

2. 潮汕汉族和闽南汉族的融合分析

聚类图上(图 2)显示河南汉族与潮汕汉族和闽南汉族比较接近,为了进一步探究潮汕人的中原汉族起源,我们假设潮汕汉族是一个融合民族,对它进行融合分析,用中原汉族和南方百越民族作为其亲本群体的代表民族。两亲本融合分析结果显示潮汕汉族的大部分血缘来自中原汉族,只有小部分来自南方百越民族。为了验证潮汕和闽南两个群体的亲缘关系,我们同时也对闽南汉族做了融合分析,结果表明潮汕和闽南汉族具有相似的中原汉族起源成分(表 2)。同时,我们把畲族也作为潮汕和闽南这两个融合民族的亲本对照,进行三亲本的融合分析。结果表明,潮汕和闽南汉族的大部分血缘仍然是来自中原汉族,但南方土著民族对潮汕汉族的混血影响要比闽南汉族大得多(表 3)。

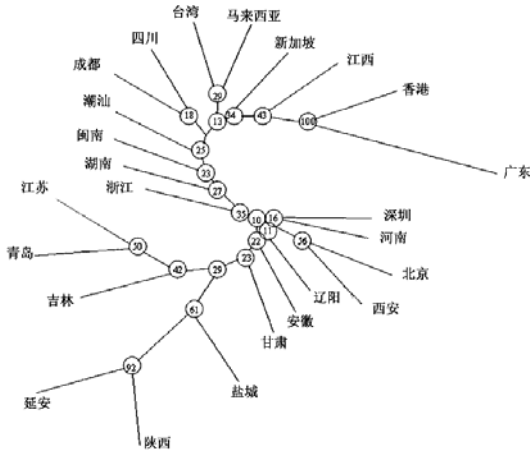


图 2 汉族 25 个人群的遗传树状图 树状图节点内的数值是经过 1000 次迭代之后的 bootstrap 值(百分比)
Fig.2 Neighbor-joining (NJ) tree based on the allelic frequencies of CODIS STR loci in 25 Han Chinese populations The numbers in the nodes are percentage bootstrap values estimated from 1,000 reiterations.

表 2 潮汕汉族闽南汉族的二亲本融合分析
Tab.2 Admixture analysis with two parental populations for Chao-Shan Han and Minnan Han

混合样本	二亲本群体的 M _{RH}	
	中原汉族	百越土著民族
潮汕汉族	0. 7348	0. 2652
闽南汉族	0. 8467	0. 1533

表 3 潮汕汉族闽南汉族的三亲本融合分析
Tab.3 Admixture analysis with three parental populations for Chao-Shan Han and Minnan Han

混合样本	三亲本群体的 M _{RH}		
	中原汉族	百越土著民族	畲族
潮汕汉族	0. 6400	0. 2406	0. 1194
闽南汉族	0. 8007	0. 1414	0. 0579

三、讨论

1. 潮汕人的中原汉族起源

我们前期的关于 HLA 的研究未能为潮汕人的中原汉族起源提供遗传学上肯定的证据[4],此外,中部群体样本采自比较广阔的地理区域[44],可能不能很好的代表该人群,从代表性区域的单一地点收集有代表性的样本可能有助于研究的深入。这种考虑基于中国历史,历史上汉族的经济和文化中心——古中原地区,北抵黄河,南临长江,西接汉水,东达淮水,范围覆盖了今河南大部[47]。来自有代表性的局限性区域——遂平县的无关个体,均是三代以上居住在当地的比较纯的当地人。因此来自此区域的样本应该能够代表中原人群的特征,之后的融合分析也是选择河南汉族做为中原汉族的代表。

种系发生树(图 2)显示,河南汉族大致位于南北汉族的中间位置,且比较接近潮汕汉和闽南汉。这与历史记载中原地区是汉族的中心地带相一致,同时与我们先前的研究结果,中部人群位于南方相关汉族和北方汉族的中间位置相一致[4]。据记载,黄河流域的中原地区是汉族的发祥地,并由此形成了所谓的中原文化,居住在公元前 4000 多年前的华夏族,其首领是炎黄二帝,被认为是汉人的始祖。大约公元前 221-206 年的秦朝时期,秦始皇派兵南下统一中国,为汉民族的扩张和融合奠定了基础。公元前 206 至公元 220 年的西汉和东汉年间,经过汉朝 400 年的统一相处,各个民族之间的差异越来越小,逐渐由多个民族融合为一个崭新的民族——汉族,中原是其文化和经济中心[47]。此后的数百上千年中,汉族逐渐向外扩展,最后遍布全国各地。汉族在华夏族的基础上同化吸收了部分匈奴、鲜卑、突厥、羌等少数民族,属于北方族群,而南方汉族主要是自秦汉以来大量的汉族移民与当地百越、百濮等南方土著居民杂居后形成的,其中比较重大的迁移是中原汉族分批迁入南方。为了逃避战乱和饥荒,中原汉族的几次南迁渐渐形成了现代潮汕人的主体。然而,中原汉人并不是直接移入潮汕地区,而是先移居福建,然后迁居入潮,因此,潮汕和闽南这两个群体享有共同的祖先[1]。本研究结果(图 2)显示这两个群体紧密聚在一起,与先前的潮汕汉和闽南汉一贯的遗传相似性研究结果[4]和历史记载保持一致。进一步的融合分析(表 2,表 3)显示这两个群体具有相似的中原汉族成分且中原汉族的成分所占比例均大于 60%,中原汉族应该是潮汕和闽南这两个群

体的主体民族。

非常有趣的是，闽南、潮汕、台湾[48]、马来西亚、新加坡的汉族主要都是闽南语人群[49]，我们的聚类分析结果（图2）均显示这五个闽语人群彼此的亲缘关系很近，与史料记载和语言学研究保持一致[1]。同时也证明台湾汉族并没有太多的台湾土著成分，而是和其他闽南语人群相似。

2. 南方土著民族对潮汕人群基因库的影响

中原起源的潮汕人在南迁的过程中可能会受到南方土著民族的影响，而历史上百越民族是中国南方中人口最大，覆盖范围最广的土著居民，两广地区历来都是百越民族分布的中心地带[3]，所以选择百越民族作为潮汕汉族的一个亲本对照。侗台语系的各民族被认为是百越族的直接后裔，我国境内的较大民族有壮、傣、侗、黎等[50]。黎族相对于其他民族，由于世居于海岛上，大海阻隔了岛内外的交流，使黎族长期处于相对隔离状态，因此受到汉族的混血影响最小[51]。同时在地理分布上，黎族也是最接近潮汕的百越民族。而壮族等民族受到汉族混血影响太大[52]，因此作为融合分析的亲本对照可能会造成很大的偏差。综合考虑，所以我们选择黎族作为百越民族的代表进行融合分析。结果表明，潮汕和闽南这两个群体有小部分的南方百越民族的血缘。生活在南方的当地土著人（百越人）种植水稻，能供养大量人口，因此中原汉族的南迁人数数量上少于百越人。而南迁的汉人也并非一下涌来，而是一波一波，这样每一次进入南方的数量更少。汉族的扩张同时带来了文化的传播，南迁的汉族虽然人数较少，却掌握了先进的技术(铁器)和文化，在文化上处于强势地位，这样互有所长的两个种群相遇，影响是互相的。土著群体文化上被汉族同化，从此成为汉族，血统上和语言上保留了一部分百越人的特征。结果，在不同的百越族群地区，发育出了不同的南方汉族和对应方言。

畬族也是南方古老的荆蛮族系民族之一，现分布在我国东南部闽、粤、浙、赣等省的山区。对畬族渊源目前学术界尚无定论，目前有土著论和非土著论两种[53]。目前较为公认的是闽、粤、赣三省交界地为畬族先民的主要活动区域。陈耿之等[54,55]认为潮州凤凰山是畬族始祖的聚居地，郭启熹[56]认为畬族曾是闽西当地的土著主居，但畬族并不是土著的百越民族，而属于荆蛮族系的苗瑶系统[53,57]。南迁的中原汉族是先移居福建然后迁居入潮的[1]，因此潮汕和闽南都有可能受到当地土著畬族的混血影响。对潮汕

和闽南汉族这两个混合民族的三亲本融合分析表明，畬族对潮汕的混血影响要比闽南大得多，这可能提示畬族在粤东一带曾大量聚居。总之，汉族与当地土著居民的融合和通婚使得今天的潮汕人与中原人群有了一定的差异[1,58]。

在聚类图上（图2），较大分支上的bootstrap 值都比较小，这可能表明在很强的基因流动背景下，STR 提供的信息量不足以解决较近人群间的亲缘关系[5]。另外，标本来源和标本均一性也会对人群的聚类造成一定影响。江苏汉族和盐城汉族主要是针对苏北地区人群的调查，这也就解释了这两个汉族与北方人群聚类的原因（图2）。深圳是个移民城市，其人群来自四面八方，比较杂，可能是深圳汉族最终没有和南方人群聚在一起的原因之一（图2）。

总之，我们此次关于潮汕人群中原汉族起源的STR 的研究结果肯定了前期HLA 的研究结果即潮汕人和闽南人的亲缘关系最近，进一步验证了这两个群体的遗传相似性。同时潮汕人群的遗传组成主要来源于中原地区的汉族移民，中国东南地区的百越民族原住群体以及晚期移民畬族等南方土著成分在现代潮汕群体中比重较小。除遗传因素之外，环境和生活习惯等因素也会对中原起源的潮汕人群产生一定的影响。北方来的中原汉族人群在迁到南方定居以后，在人群基因库对于环境的巨大差异的适应过程中，会产生各种代谢的不协调。在潮汕地区某些疾病的高发可能就是这种不协调的结果。此外，综合常染色体STR 以及能反映母系进化历史和父系进化历史的遗传标记 mtDNA 和 Y-STR、Y-SNP 等[59]方面的研究以及环境调查等流行病学方面的研究，可以为后期环境基因组学的研究做铺垫。

致谢：我们衷心的感谢为本课题的研究工作默默奉献的血样捐献者。衷心感谢李辉博士在数据分析和论文修改上给予的指导和帮助。本项目得到一下基金支持：香港李嘉诚基金会剑桥大学国际合作项目基金；英国剑桥大学基金；广东省自然科学基金（000819）；汕头大学研究与发展基金（L00007）；汕头市重点科技计划项目（汕府科[2005]116号）。

参考文献

1. 黄挺著(1997)潮汕文化源流(第1版). 广州:广东高等教育出版社. 1-74.
2. 葛剑雄, 吴松弟, 曹树基 (1997) 中国移民史. 福州: 福建人民出版社.
3. Li H, Cai XY, Elizabeth R WC, Wen B, Cheng X, Qin ZD, Liu WH, Liu YF, Pan SL, Qian J, Tan CC, Jin L (2007) Mitochondrial DNA diversity and population differentiation in Southern East Asia. *Am J Phys Anthropol* 184:481-488.
4. Hu SP, Luan JA, Li B, Chen JX, Cai KL, Huang LQ, Xu XY (2007) Genetic link between Chaoshan and other Chinese Han populations: evidence from HLA-A and HLA-B allele frequency distribution. *Am J Phys Anthropol* 132:140-150.
5. Nei M, Takezaki N (1996) The root of the phylogenetic tree of Human populations. *Mol Biol Evol* 13:170-171.
6. Chu JY, Huang W, Kuang SQ, Wang JM, Xu JJ, Chu ZT, Yang

- ZQ, Lin KQ, Li P, Wu M, Geng ZC, Tan CC, Du RF, Jin L (1998) Genetic relationship of populations in China. *Proc Natl Acad Sci USA* 95:11763-11768.
7. 金天博, 高雅, 陈腾, 阎春霞, 李生斌 (2004) 广西地区15个不同民族人群的群体遗传学关系. *西安交通大学学报(医学版)* 25:422-424.
 8. 李辉 (2007) 分子人类学所见历史上闽越族群的消失. *广西民族大学学报(哲学社会科学版)* 29(2):42-47.
 9. Hu SP, Yu XJ, Liu JW, Cai KL (2005) Analysis of STR polymorphisms in the Chaoshan population in South China. *Forensic Sci Int* 147:93-95.
 10. Liu YC, Huo ZY, Tang H, Wang J, Jiao ZP, Ma WS, Yan JW, Yang J, Jia SQ, Zhu XJ (2003) Frequency data for 15 STR loci and forensic use in a Beijing-Han population. *International Congress Series* 1239:267-270.
 11. 杨玉玲, 王冬花, 陈玲, 侯庆唐, 刘忠, 方勤, 吕蓉, 郭晓婕 (2007) 安徽汉族人群15个STR基因座遗传多态性. *法医学杂志* 23:36-38.
 12. 陶晓岚, 聂笑联, 张杰 (2006) 甘肃地区汉族人群15个STR基因座遗传多态性. *中国法医学杂志* 21:306-307.
 13. 王振原, 高树辉, 王新淮, 于荣军 (2003) 西安地区强奸犯(汉族)13个STR基因座遗传多态性. *公安大学学报(自然科学版)* 5:48-51.
 14. 王振原, 马骏, 徐永城, 樊栓良, 方俊邦 (2005) 西安地区汉族人群13个STR基因座多态性分析. *中南大学学报(医学版)* 30:236-240.
 15. Yang BQ, Wang G, Liu YC, Yang WS (2005) Population data for the AmpFLSTR IdentifierTM PCR Amplification Kit in China Han in Jilin Province, China. *Forensic Sci Int* 151:293-297.
 16. Wang W, Jia HL, Wang Q, Cai ZM, Wei LH, Wang DB, Bittles AH (2003) STR polymorphisms of "forensic loci" in the northern Han Chinese population. *J Hum Genet* 48:337-341.
 17. Wang ZY, Yu RJ, Wang F, Li XS, Gin TB (2005) Genetic polymorphisms of 15 STR loci in Han population from Shaanxi (NW China). *Forensic Sci Int* 147:89-91.
 18. 王振原, 王芳, 余荣军, 李涛, 贺光社, 方俊邦 (2004) 延安地区汉族人群15个STR基因座多态性分析. *中国医学科学院学报* 26:549-553.
 19. 李志强, 李永红, 李震海, 赵彩霞, 纪尚起, 接贵祥, 于丰钦 (2002) 青岛地区汉族人群13个STR基因座的频率分布及法医学应用. *法律与医学杂志* 9:219-222.
 20. Chan KM, Chiu CT, Tsui P, Wong DM, Fung WK (2005) Population data for the IdentifierTM 15 STR loci in Hong Kong Chinese. *Forensic Sci Int* 152:307-309.
 21. Law MY, To KY, Ho SH, Pang BCM, Wong LM, Wun HL, Yau SK, Chan KL (2002) STR data for the PowerPlexTM 16 loci for the Chinese population in Hong Kong. *Forensic Sci Int* 129:64-67.
 22. Yong RY, Aw LT, Yap EP (2004) Allele frequencies of 15 STR loci of three main ethnic populations in Singapore using an in-house marker panel. *Forensic Sci Int* 141:175-183.
 23. 李英碧, 吴谨, 侯一平, 张霁, 廖森, 张林, 陈国弟 (2005) 成都汉族群体15个短串联重复序列基因座遗传多态性. *中华医学遗传学杂志* 22:169-173.
 24. Zhang HJ, Li YB, Jiang JP, Zhang J, Wu J, Du H, Yan J, Shen YH, Hou YP (2007) Analysis of 15 STR loci in Chinese population from Sichuan in West China. *Forensic Sci Int* 171:222-225.
 25. 李海燕, 台运春, 陆惠玲, 刘超, 陈先勤 (2004) 中国汉族人群15个STR基因座德等位基因频率调查. *中国法医学杂志* 19:330-333.
 26. Wang CW, Chen DP, Chen CY, Lu SC, Sun CF (2003) STR data for the AmpF/STR SGM Plus and Profiler loci from Taiwan. *Forensic Sci Int* 138:119-122.
 27. Hu SP, Wu DQ, Xu XH, Liu JW, Li B (2005) Genetic profile of 15 STR loci in the Min Nan population in Southeast China. *Forensic Sci Int* 152:263-265.
 28. 郝宏蕾, 吴微微, 潘立鹏, 徐向春, 郑小婷 (2006) 浙江汉族人群15个STR基因座遗传多态性. *中国法医学杂志* 21:303-304.
 29. 孙耀东, 李健, 陈月勇, 任宏 (2006) 盐城地区汉族人群15个STR基因座的遗传多态性. *中国法医学杂志* 21:235-236.
 30. 俞卫东, 闵涯邻, 张斌, 冯燕 (2003) 江苏地区汉族人群15个STR基因座频率调查. *中国法医学杂志* 18:245-246.
 31. 刘秋玲, 吕德坚, 陈丽娟, 陆惠玲, 李健金 (2004) 湖南汉族人群15个STR基因座德遗传多态性. *中国法医学杂志* 19:98-99.
 32. 赵会安, 刘小莹, 杨泉根, 黎光明 (2003) 江西汉族人群13个STR基因座的遗传多态性调查. *江西科学* 21:21-25.
 33. 李晓斌, 赖跃, 徐振波, 胡海洋 (2002) 深圳地区汉族人群13个STR基因座的遗传多态性. *中国法医学杂志* 17:150-151.
 34. Izuam M, Seah LH, Panneerchelvam S, Nor NM (2005) STR data for the PowerPlex 16 System loci for the Malays, Chinese and Indians groups of the Malaysian population. *J Forensic Sci* 50:1225-1228.
 35. Seah LH, Jeevan NH, Othman MI, Jaya P, Ooi YS, Wong PC, Kee SS (2003) STR Data for the AmpFLSTR Identifier loci in three ethnic groups (Malay, Chinese, Indian) of the Malaysian population. *Forensic Sci Int* 138:134-137.
 36. Walsh PS, Metzger DA, Higuchi R (1991) Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. *Biotechniques* 10:506-513.
 37. Guo SW, Thompson EA (1992) Performing the exact test of Hardy-Weinberg proportion for multiple alleles. *Biometrics* 48:361-372.
 38. Felsenstein J (2004) PHYLIP (Phylogeny Inference Package) version 3.6. Seattle: Department of Genome Sciences, University of Washington. (<http://evolution.genetics.washington.edu/phylip/phylipweb.html>)
 39. Nei M (1987) Genetic variation within species. In: *Molecular evolutionary genetics*. New York: Columbia University Press, 220-221.
 40. Saitou N, Nei M (1987) The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4:406-425.
 41. Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39:783-791.
 42. Wang J (2003) Maximum-likelihood estimation of admixture proportions from genetic data. *Genetics* 164:747-765.
 43. Roberts DF, Hiorns RW (1965) Methods of analysis of the genetic composition of a hybrid population. *Hum Biol* 37:38-43.
 44. Shaw CK, Chen LL, Lee A, Lee TD (1999) Distribution of HLA gene and haplotype frequencies in Taiwan: A comparative study among Minnan, Hakka, Aborigines and Mainland Chinese. *Tissue Antigens* 53:51-64.
 45. 杜若甫, 肖春杰, Cavalli-Sforza LL (1998) 用38个基因座的基因频率计算中国人间遗传距离. *中国科学(C辑)* 28:83-89.
 46. 肖春杰, 杜若甫, Cavalli-Sforza LL (2000) 中国人基因频率的主成分分析. *中国科学(C辑)* 30:434-442.
 47. 李立新 (2006) 论河南的根文化与根文化研究. *黄河科技大学学报* 8(3):42-44.
 48. Lin M, Chu CC, Chang SL, Lee HL, Loo JH, Akaza T, Fuji T, Ohashi J, Tokunaga K (2001) The origin of Minnan and Hakka, the so-called "Taiwanese", inferred by HLA study. *Tissue Antigens* 57:192-199.
 49. Gan RJ, Pan SL, Mustavich JF, Qin ZD, Cai XY, Qian J, Liu CW, Peng JH, Li SL, Xu JS, Jin L, Li H (2008) Pinghua population as an exception of Han Chinese's coherent genetic structure. *J Hum Genet* 53:303-313.
 50. 李辉 (2002) 百越族群遗传结构分析的初步思路. *现代人类学国际研讨会论文集*. 上海, 2002年4月12日, 89-94.
 51. Li D, Li H, Ou C, Lu Y, Sun Y, Yang B, Qin Z, Zhou Z, Li S, Jin L (2008) Paternal Genetic Structure of Hainan Aborigines Isolated at the Entrance to East Asia. *PLoS ONE* 3(5):e2168.
 52. Chen J, Li H, Qin ZD, Liu WH, Lin WX, Yin RX, Jin L, Pan SL (2006) Y-chromosome genotyping and genetic structure of Zhuang populations. *Acta Genetica Sinica* 33:1060-1072.
 53. 施联朱 (1983) 关于畲族来源与迁徙. 见: *施联朱 畲族研究论文集*. 北京: 民族出版社, 34-52.
 54. 李筠文 (2006) 广东畲族与畲族研究. *广东技术师范学院学报* 2:1-6.
 55. 陈歌之 (2004) 畲族的发源地与畲族的文化影响. *学术研究* 10:112-119.
 56. 郭启熹 (2002) 闽西畲族源流初探. *闽西职业大学学报* 2:14-17.
 57. 练铭志 (1998) 试论粤东历史上的畲族和瑶族. *民族研究* 5:97-106.
 58. 杜松年 (1994) *潮汕大文化(第一版)*. 北京: 中国科学技术出版社, 1-104.
 59. Thangaraj K, Singh L, Reddy AG, Rao VR, Sehgal SC, Underhill PA, Pierson M, Frame IG, Hagelberg E (2003) Genetic affinities of the Andaman islanders, a vanishing human population. *Curr Biol* 13:86-93.