

中国西南地区藏族人群遗传亚结构研究

王小娟^{1,2}, 钱恩芳¹, 李悦², 宋正阳^{1,2}, 赵慧², 谢何鑫², 李彩霞^{1,2},
黄江¹, 江丽²

1. 贵州医科大学法医学院, 贵阳 550004

2. 公安部物证鉴定中心, 北京市现场物证检验工程技术研究中心, 现场物证溯源技术国家工程实验室, 北京 100038

摘要: 藏族为中国西南地区典型的少数民族, 分为卫藏、康巴、安多和嘉绒等多个支系。然而, 对藏族支系人群的遗传结构, 特别是各分支人群的父亲、母系遗传结构却缺乏深度解析。本研究基于个体水平的常染色体、父系来源的 Y 染色体和母系来源的线粒体 3 个类别遗传信息, 对西藏地区卫藏藏族、四川甘孜地区康巴藏族、青海地区安多藏族和四川阿坝地区嘉绒藏族共 4 个藏族群体进行研究, 以揭示其遗传亚结构关系。采用微测序技术检测各位点分型, 利用 PowerPlex[®]Y23 和 DNATyper[™] Y26 试剂盒检测 Y-STRs 基因座分型, 通过热图和主成分分析、祖先成分分析、单倍群频率统计、网络图及多维尺度分析等探讨其遗传亚结构。结果表明, 常染色体和 Y 染色体遗传标记可将 4 个藏族人群分为 3 类: 青藏高原的卫藏藏族为一类, 高原周边地区的康巴藏族和安多藏族的遗传结构类似分为一类, “藏彝走廊”中嘉绒藏族的遗传结构与其他藏族人群差异显著而为一类。不同藏族分支人群在线粒体遗传信息方面无明显差异性。上述多类别遗传信息的分析结果为深入了解藏族不同分支人群的遗传亚结构提供了新视角。

关键词: 群体遗传学; 藏族; 单倍型; 遗传亚结构

收稿日期: 2019-12-26; 修回日期: 2020-04-14

基金项目: 国家自然科学基金项目(编号: 81772027), 国家重点研发计划项目(编号: 2017YFC0803501), 公安部物证鉴定中心基本科研业务费专项资金项目(编号: 2017JB025)和国家科技资源共享服务平台计划项目(编号: YCZYPT[2017]01-3)资助[Supported by the National Natural Science Foundation of China (No. 81772027), the National Key R&D Program of China (No. 2017YFC0803501), the Fundamental Research Funds for Institute of Forensic Science (No. 2017JB025), and the National Science and Technological Resources Platform (No. YCZYPT[2017]01-3)]

作者简介: 王小娟, 硕士研究生, 专业方向: 法医物证学。E-mail: 757164796@qq.com

通讯作者: 黄江, 博士, 教授, 研究方向: 法医物证学。E-mail: mmm_hj@gmc.edu.cn

江丽, 博士, 副主任法医师, 研究方向: 法医遗传学。E-mail: jl@mail.bnu.edu.cn

DOI: 10.16288/j.ycz.19-330

网络出版时间: 2020/4/16 14:03:17

URI: <http://kns.cnki.net/kcms/detail/11.1913.R.20200415.1535.002.html>

A genetic sub-structure study of the Tibetan population in Southwest China

Xiaojuan Wang^{1,2}, Enfang Qian¹, Yue Li², Zhengyang Song^{1,2}, Hui Zhao², Hexin Xie², Caixia Li^{1,2}, Jiang Huang¹, Li Jiang²

1. Institute of Forensic Medicine, Guizhou Medical University, Guiyang 550004, China

2. Institute of Forensic Science, Ministry of Public Security & Beijing Engineering Research Center of Crime Scene Evidence Examination & National Engineering Laboratory for Forensic Science, Beijing 100038, China

Abstract: Tibetan is a typical ethnic minority population in Southwest China, which can be divided into U-Tsang, Kham, Amdo, Jiarong and other sub-populations. However, the genetic structure of these sub-populations has not been comprehensively analyzed, especially from the perspective of paternal and maternal lineages. Based on genetic markers of autosomes, the Y chromosome and mitochondria, we studied four Tibetan populations (the U-Tsang population in Tibet Autonomous Region; the Kham population in Garze, Sichuan province; the Amdo population in Qinghai province and the Jiarong population in Aba, Sichuan province) to interpret their genetic structure. The mini-sequencing technology was used to detect the genotype of each marker. Meanwhile, the PowerPlex[®]Y23 and DNA Typer[™] Y26 kit were applied to genotype Y-STRs. Subsequently, the genetic structure was analyzed by heatmap and principal component analysis, ancestry component, haplogroup frequency, network map and multi-dimensional scaling analysis. The results showed that the four Tibetan populations could be divided into three sets based on the autosomal and Y-chromosomal genetic markers, in which set 1 was the U-Tsang population in the Tibetan Plateau, set 2 comprised of the Kham and Amdo populations in the surrounding areas of the plateau, and set 3 was the Jiarong population that resided in the Tibetan and Yi Corridor. No significant difference was observed in mitochondrial genetic markers among four Tibetan populations. In general, multi-category genetic information provides a new comprehensive insight into the Tibetan sub-population.

Keywords: population genetics; Tibetan; haplotype; genetic sub-structure

藏族是中国古老的民族之一，人数达 628.2 万 (2010 年全国第 6 次人口普查数据)，其中 43.2% 分布于西藏自治区，其余主要分布于青海和四川等地 (<http://www.stats.gov.cn>)。藏语属汉藏语系藏缅语族藏语支，分卫藏、康巴和安多 3 类方言。根据方言类型，可将藏族人群划分为卫藏藏族、康巴藏族和安多藏族^[1]。此外，嘉绒藏族等支系也是藏族人群的重要成员^[2]。

藏族人群长期居住于海拔 3000~5000 m 的青藏高原，经长期自然选择，表现出对高原低氧环境较强的适应能力^[3,4]。研究表明，藏族人群的高原适应 (high-altitude adaptation, HAA) 具有明显的遗传学基础。全基因组关联研究发现，在缺氧诱导因子 (hypoxia-inducible factor, HIF) 通路上存在 *EGLN1*、*EPAS1* 等多个候选基因与藏族人群的 HAA 相关^[5,6]。候选基因上多个单核苷酸多态性 (single nucleotide

polymorphisms, SNPs) 位点的等位基因频率分布在藏族人群与平原人群间具有显著差异，如 *EPAS1* 基因上 5 个 SNPs 位点组成的单倍型 AGGAA (rs115321619、rs73926263、rs73926264、rs73926265 和 rs55981512) 高频分布于藏族人和丹尼索瓦人^[7,8]；*EGLN1* 基因上也检测出 rs18699510 和 rs12097901 等多个分布差异位点^[9]。除常染色体外，Y 染色体的非重组区域以单倍型形式呈父系遗传，有效群体小，对漂变敏感，包含人群更多的父系历史信息^[10]。Y-SNPs 单倍群在世界人群的分布呈现明显的地理特异性，其中单倍群 D-M174 主要分布于中国西藏、日本和安达曼群岛^[11~13]。与 Y 染色体类似，线粒体 DNA 呈母系遗传，是反映人群母系遗传信息的遗传标记，其单倍群的支系分布同样具有地域差异 (<http://www.phytree.org/tree/index.htm>)。目前也有研究结合父系和母系遗传信息探索藏族人群的遗传结

构及源流^[14]。但是, 不同地理位置的藏族分支人群之间的遗传结构特点有待进一步探究。

本研究基于常染色体、Y 染色体及线粒体多类别基因组遗传标记, 对西藏地区卫藏藏族、四川甘孜地区康巴藏族、青海地区安多藏族和四川阿坝地区嘉绒藏族等 4 个藏族人群共 297 份男性样本进行检测, 分析其遗传结构, 探讨西南地区不同藏族分支人群的遗传亚结构。

1 材料与方法

1.1 样本信息

采集中国西南地区 297 份男性藏族无关个体静脉血样本, 其中西藏地区卫藏藏族 89 例、四川甘孜地区康巴藏族 88 例、青海地区安多藏族 82 例、四川阿坝地区嘉绒藏族 38 例。样本信息及人群简称见表 1。本研究已通过公安部物证鉴定中心伦理委员会审查, 上述样本供者均签署知情同意书。根据试剂盒说明书, 采用 QIAamp[®] DNA Blood Midi Kit (德国 QIAGEN 公司)进行 DNA 提取。使用 NanoDrop 2000c (美国 Thermo Fisher Scientific 公司)进行 DNA 定量并稀释至 5 ng/ μ L 待检。

1.2 位点筛选

基于多类别基因组遗传信息, 共选择 134 个位点, 位点信息见附表 1。其中包括: (1)通过文献调研^[15]选择藏族特异性常染色体位点 25 个; (2)依据 Y 染色体谱系树 2018 版(<https://isogg.org>)及文献调研, 选择在东亚人群中突变频率较高的 Y-SNPs 位点 75 个; (3)依据线粒体谱系树 17 版(<http://www.phylotree.org>)及公开数据库和文献调研, 筛选在东亚人群中突变频率较高及具有藏族特异性的 mt-SNPs 位点 34 个。

1.3 分型检测

1.3.1 筛选位点分型检测

使用 Primer Premier 5.0 软件设计 PCR 扩增及延伸引物并由生工生物工程(上海)股份有限公司合成。利用 Applied Biosystems[™] SNaPshot[™] Multiplex 试剂盒(美国 Thermo Fisher Scientific 公司)进行位点检测。应用 Mastercycler ProS 热循环仪(德国 Eppendorf 公司)进行复合扩增, 扩增产物使用 ABI 3130xl 遗传分析仪(美国 Thermo Fisher Scientific 公司)进行毛细管电泳检测, 利用 GeneMapper ID v3.2 软件分析样本分型。

1.3.2 Y-STRs 分型检测

选取 16 个单拷贝 Y-STRs 基因座(DYS19、DYS389I、DYS389II、DYS390、DYS391、DYS392、DYS393、DYS437、DYS438、DYS439、DYS448、DYS456、DYS458、DYS533、DYS635 和 GATA-H4)和 1 个多拷贝基因座 DYS385ab, 使用试剂盒检测西藏藏族(CTT)、四川甘孜藏族(CTG)、青海藏族(CTQ)和四川阿坝藏族(CTA) 4 个人群的分型数据, 其中西藏藏族(CTT)、四川甘孜藏族(CTG)和青海藏族(CTQ)人群样本使用 PowerPlex[®] Y23 荧光标记复合扩增系统(美国 Promega 公司)试剂盒检测; 四川阿坝藏族(CTA)人群样本使用 DNATyper[™] Y26 试剂盒(中国公安部物证鉴定中心)检测。扩增产物用 ABI 3500xl 遗传分析仪(美国 Thermo Fisher Scientific 公司)进行毛细管电泳检测并使用 GeneMapper[®] ID-X 软件进行 Y-STRs 分型分析。

1.4 数据分析

1.4.1 常染色体遗传标记

采用 Genepop v4.2 软件(<http://www.genepop>。

表 1 4 个藏族人群的样本信息

Table 1 Information of four Tibetan populations

| 序号 | 群体 | 所属支系 | 英文全称 | 英文简称 | 检测样本量(例) | 海拔高度(m) |
|----|--------|------|---|------|----------|---------|
| 1 | 西藏藏族 | 卫藏 | Chinese Tibetan in Tibetan Plateau | CTT | 89 | 3650 |
| 2 | 四川甘孜藏族 | 康巴 | Chinese Tibetan in Garze Sichuan, China | CTG | 88 | 3300 |
| 3 | 青海藏族 | 安多 | Chinese Tibetan in Qinghai, China | CTQ | 82 | 2420 |
| 4 | 四川阿坝藏族 | 嘉绒 | Chinese Tibetan in Aba Sichuan, China | CTA | 38 | 2130 |

curtin.edu.au) 计算 4 个藏族人群的等位基因频率, R v3.2.3 软件绘制热图并进行主成分分析。使用 DNA 族群推断系统软件(DNA Ancestry Analyzer Version 1.0, DAA v1.0)^[16] 进行人群聚类分析, 参数设置为 $K=2$, 运行 20 次。

1.4.2 Y 染色体 SNPs 及 STRs 遗传标记

基于国际 Y 染色体协会(Y chromosome consortium, YCC)的单倍群命名原则^[17], 根据 Y-SNPs 位点的分型情况划分样本所属单倍群, 采用直接计数法统计各单倍群频率。基于 16 个单拷贝 Y-STRs 基因座分型使用公式 $D = [n(1 - \sum P_i^2)] / (n-1)$ 计算每个藏族分支人群的基因/单倍型多样性, 其中 P_i 为第 i 个等位基因/单倍型频率, n 为样本数。对 Y-SNPs 单倍群相同的样本, 基于 16 个单拷贝 Y-STRs 基因座分型使用 Network v5.0 软件进行网络图分析^[18]。根据 Y-STRs 基因座的基因多样性和突变率综合评估设定权重值, 基因多样性和突变率越高, 权重值越低。其中 Y-STRs 基因座突变率参考 YHRD 数据库(https://yhrd.org/pages/resources/mutation_rates)。为评估藏族分支人群的遗传关系, 比较了其 6 个邻近人群(四川成都/泸州汉族^[19-21]、四川彝族^[22]、甘肃藏族(YA004043)、昌都康巴藏族^[23]、青海回族^[24,25]、青海撒拉族(YA003302)和南方汉族(广西汉族(YA004218))、北方汉族(北京汉族^[26-28])的 17 个 Y-filer 基因座组成的单倍型。所有参考人群的分型数据来自 YHRD 数据库。利用 YHRD 数据库在线软件 AMOVA & MDS 工具计算群体间遗传距离 R_{st} 矩阵, 并构建多维尺度(MDS)散点图进行可视化。

1.3.3 线粒体 SNPs 遗传标记

根据 PhyloTree 17 版谱系树, 结合各样本在 mt-SNPs 位点的分型判断其单倍群类型, 采用直接计数法统计各单倍群频率。

2 结果与分析

2.1 个体水平遗传结构

2.1.1 群体聚类分析

25 个位点在 4 个藏族人群的等位基因频率分布

热图如图 1A 所示。基于突变型等位基因频率可将 4 个藏族人群分为 3 类, 除 rs10206434、rs375554942、rs76347095、rs6544888 和 rs1562453 位点外, 西藏藏族(CTT)在其他各位点的突变型等位基因频率最高, 四川甘孜藏族(CTG)和青海藏族(CTQ)的突变型等位基因频率类似且次于西藏藏族(CTT), 四川阿坝藏族(CTA)的突变型等位基因频率最低, 与其他藏族分支人群具有较大差异性。根据各位点在人群的等位基因频率进行主成分分析, 结果见图 1B。PC1 和 PC2 成分的累积贡献率为 100%, 在 PC1 维度(99.7%)上, 4 个藏族人群可分为 3 类且彼此距离较远而区分开: 西藏藏族(CTT); 四川甘孜藏族(CTG)和青海藏族(CTQ); 四川阿坝藏族(CTA)。

2.1.2 祖先成分

STRUCTURE 分析可直观反映族群间的聚类关系。4 个藏族人群 STRUCTURE 分析在 $K=2$ 时的结果见图 2, 黄色和红色分别代表一种祖先成分(以下分别称为成分 1 和成分 2), 颜色比例的构成代表祖先成分比例的构成。结果显示, 基于祖先成分比例可将 4 个藏族人群分为 3 类: (1)西藏藏族(CTT), 群体水平上成分 1 和成分 2 分别为 0.719 和 0.281, 个体水平上 56.2% 个体成分 1 大于 0.75, 34.8% 个体介于 0.25~0.75 之间, 9.0% 个体小于 0.25; (2)四川甘孜藏族(CTG)和青海藏族(CTQ), 群体水平上成分 1 和成分 2 分别约为 0.45 和 0.55, 个体水平上约 25.0% 个体成分 1 大于 0.75, 45.0% 个体介于 0.25~0.75 之间, 30.0% 个体小于 0.25; (3)四川阿坝藏族(CTA), 群体水平上成分 1 和成分 2 分别为 0.175 和 0.825, 个体水平上无成分 1 大于 0.75 的个体, 36.8% 个体介于 0.25~0.75 之间, 63.2% 个体小于 0.25。

2.2 父系遗传结构

2.2.1 Y-SNPs 单倍群在藏族人群中的分布

藏族人群 Y 染色体单倍群类型及分布频率见表 2。各单倍群频率在藏族人群之间存在显著差异, 其中以单倍群 D-M174 (34.21%~71.91%) 和 O-M175 (16.85%~60.53%) 为主要类型。在下游分支单倍群中, 以单倍群 D1a2-P99 (44.44%) 和 O2-M122 (22.90%) 为主。基于单倍群 D-M174 类型, 可将 4 个藏族人

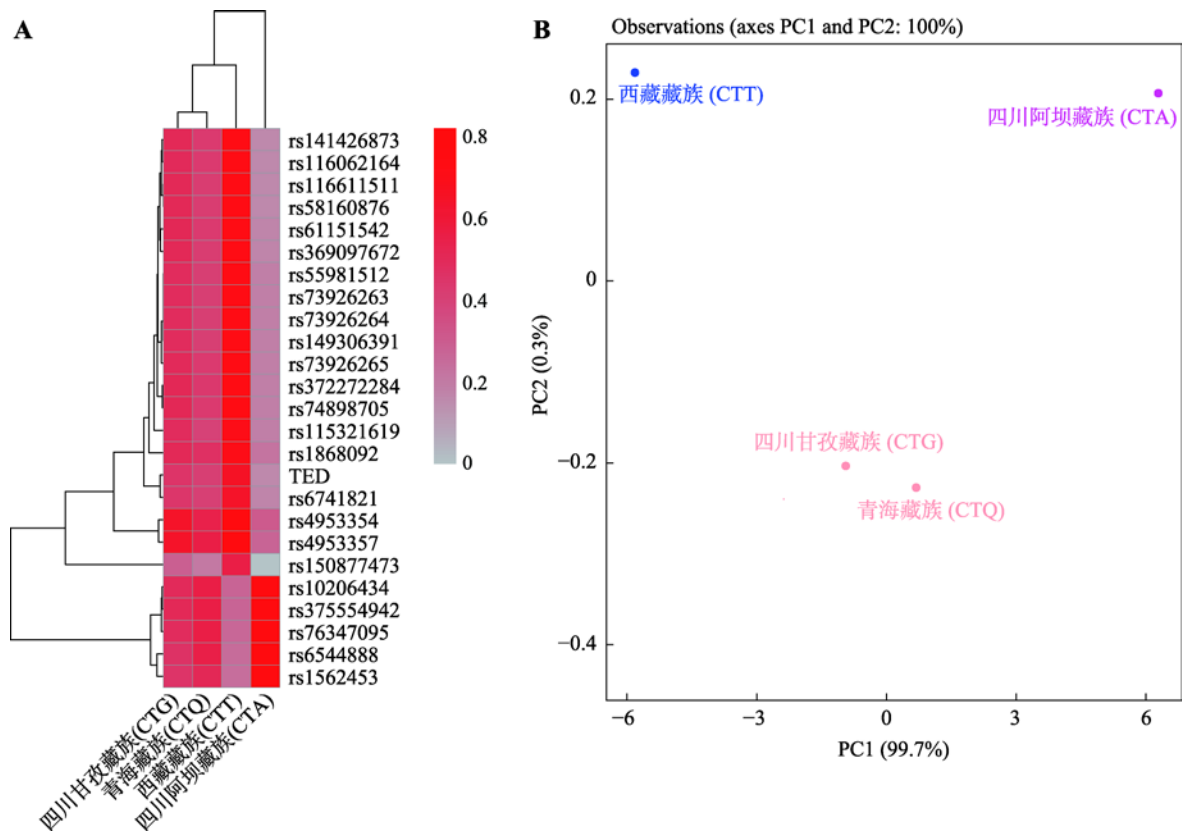


图 1 4 个藏族人群聚类分析
Fig. 1 Cluster analysis of four Tibetan populations

A: 基于 25 个常染色体位点等位基因频率的 4 个藏族人群热图分析。颜色深浅表示各位点在人群中的突变型等位基因频率大小, 树状分枝表示人群和位点的聚类关系。B: 基于 25 个常染色体位点等位基因频率的 4 个藏族人群主成分分析。

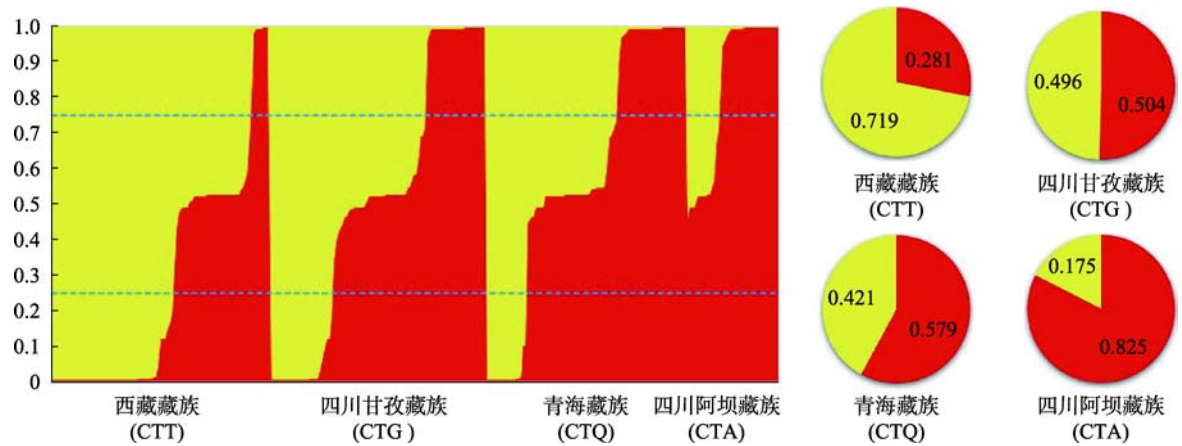


图 2 基于 25 个常染色体位点分型的 4 个藏族人群 STRUCTURE 分析
Fig. 2 STRUCTURE analysis of four Tibetan populations based on the genotype of 25 autosomal markers

个体水平祖先成分见柱状图, 其中每个柱形代表一个个体, 每种颜色代表一种祖先成分, 颜色比例代表该个体祖先成分比例; 群体水平祖先成分见饼图, 其中每种颜色代表的祖先成分同柱状图, 颜色比例代表该群体祖先成分比例。 $K=2$, $Ln=-5201.9$ 。

群分为 3 类: (1)西藏藏族(CTT), 71.91%个体属于单倍群 D-M174, 下游分支以 D1a2-P99 为主, 分布频率为 53.93%; (2)四川甘孜藏族(CTG)和青海藏族(CTQ), 约 60%个体属于单倍群 D-M174, 下游分支

为进一步观察不同藏族分支人群的遗传关系,选择在 4 个藏族人群中分布频率较高的 3 个单倍群 D1a1-M15、D1a2-P99 和 O-M175,根据 16 个单拷贝 Y-STRs 基因座组成的单倍型绘制网络图,权重值设置见附表 2,网络图如图 4 所示。结果显示,各单倍群中每一分支人群均未发现明显聚类。西藏藏

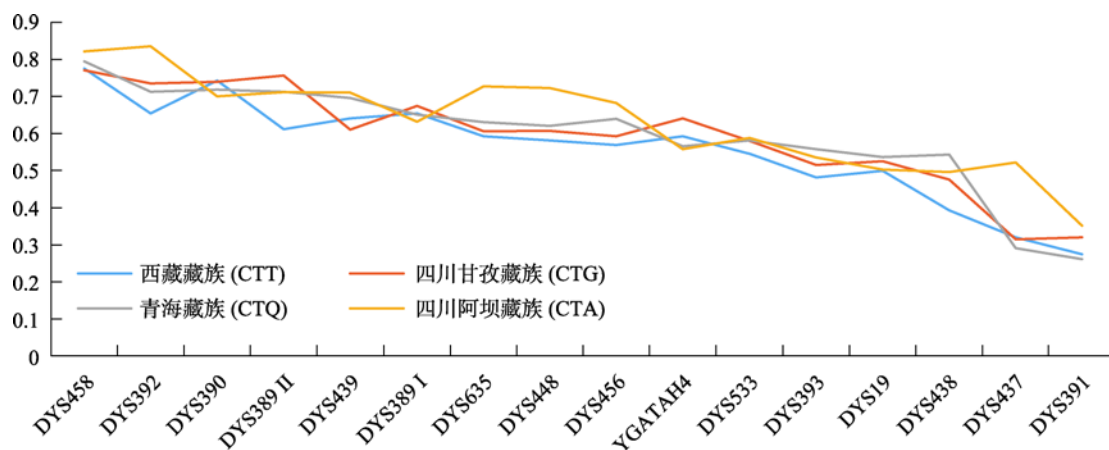


图 3 16 个 Y-STRs 基因座在 4 个藏族人群中的基因多样性

Fig. 3 Genetic diversity of four Tibetan populations at 16 Y-STRs loci

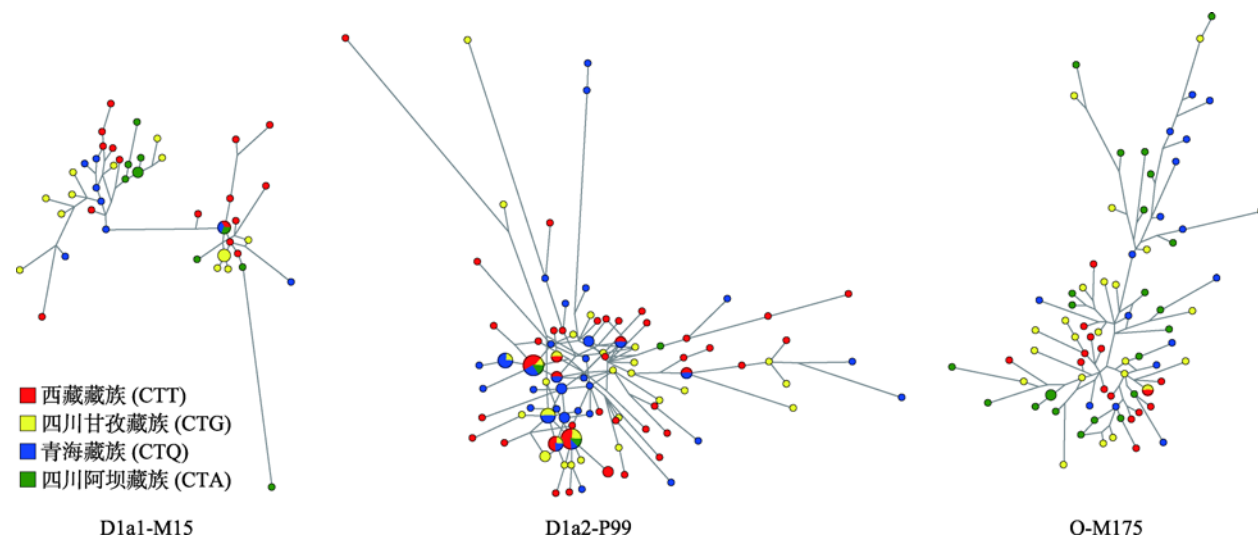


图 4 基于 16 个 Y-STRs 基因座分型的 4 个藏族人群中 D1a1-M15、D1a2-P99 和 O-M175 单倍群网络图分析结果

Fig. 4 The network analysis of four Tibetan populations with D1a1-M15, D1a2-P99 and O-M175 haplogroups based on 16 Y-STRs genotypes

族(CTT)、四川甘孜藏族(CTG)和青海藏族(CTQ)在单倍群 D1a2-P99 中的分布较多;四川阿坝藏族(CTA)中单倍群 O-M175 分布较多。各单倍群均存在不同人群共享 Y-STRs 单倍型。

2.2.4 藏族分支人群与其他群体间遗传关系

遗传距离是反映群体间遗传差异的重要指标。基于遗传距离的 4 个藏族人群与其他 8 个人群的 MDS 散点图如图 5 所示。结果表明,西藏藏族(CTT)、四川甘孜藏族(CTG)和青海藏族(CTQ)聚为一簇,与昌都康巴藏族距离较近,与其他群体相距较远;四

川阿坝藏族(CTA)与四川彝族距离最近,与其他藏族人群相距较远。

2.3 母系遗传结构

4 个藏族人群线粒体单倍群类型及分布频率见表 3, 各单倍群的频率分布在藏族人群之间无显著差异性。各藏族人群均检出 M、N 两支单倍群, 其中单倍群 M 的频率(56.76%~67.05%)高于单倍群 N (32.95~43.24%)。单倍群 M 以 M8、M9、G 和 D 单倍群 4 个下游分支为主, 其中单倍群 D 的分布频率约为 25%。单倍群 N 以 A 和 R 两类分支单倍群为主,

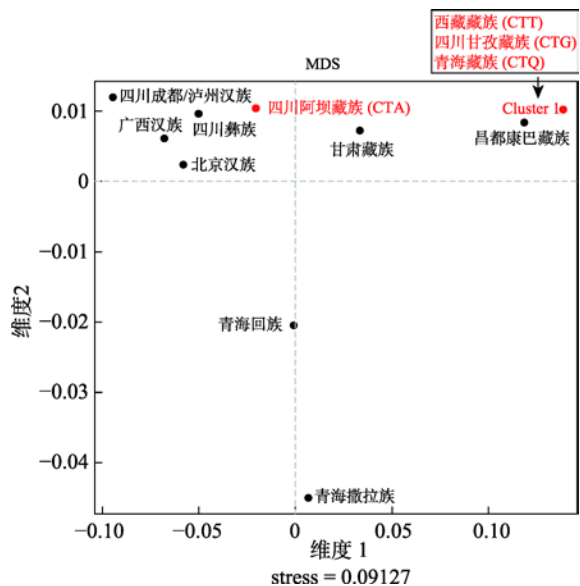


图 5 基于遗传距离的人群多维尺度分析
Fig. 5 Multi-dimensional scaling (MDS) plot for populations based on Rst values

红色字体为本研究的 4 个藏族人群，其他人群来源于 YHRD 数据库；Cluster 1 表示西藏藏族(CTT)、四川甘孜藏族(CTG)和青海藏族(CTQ)因遗传距离近而聚为一簇。

表 3 藏族人群 mt-SNPs 单倍群谱系树及频率分布
Table 3 Mitochondrial phylogenetic tree and haplogroup frequencies of each Tibetan population

| 单倍群谱系树 | | | 西藏藏族 (CTT) | | 四川甘孜藏族 (CTG) | | 青海藏族 (CTQ) | | 四川阿坝藏族 (CTA) | |
|---------|---------|-----------|------------|--------|--------------|--------|------------|--------|--------------|--------|
| | | | <i>n</i> | % | <i>n</i> | % | <i>n</i> | % | <i>n</i> | % |
| A1018G | L3 | | 88 | 100.00 | 87 | 100.00 | 79 | 100.00 | 37 | 100.00 |
| G10400A | M | | 59 | 67.05 | 54 | 62.07 | 48 | 60.76 | 21 | 56.76 |
| | T4715C | M8 | 5 | 5.68 | 9 | 10.34 | 3 | 3.80 | 6 | 16.22 |
| | A13263G | C | 3 | 3.41 | 8 | 9.20 | 2 | 2.53 | 3 | 8.11 |
| | G4491A | M9 | 19 | 21.59 | 13 | 14.94 | 11 | 13.92 | 1 | 2.70 |
| | A3394G | M9a'b | 16 | 18.18 | 9 | 10.34 | 8 | 10.13 | 1 | 2.70 |
| | C7697T | M9a1a1c1b | 11 | 12.50 | 5 | 5.75 | 2 | 2.53 | | |
| | G14569A | M12'G | 13 | 14.77 | 10 | 11.49 | 7 | 8.86 | 4 | 10.81 |
| | A4833G | G | 13 | 14.77 | 10 | 11.49 | 7 | 8.86 | 4 | 10.81 |
| | T4561C | M62'68 | | | | | 1 | 1.27 | | |
| | G2735A | M62 | | | | | 1 | 1.27 | | |
| | C4883T | M80'D | 15 | 17.05 | 19 | 21.84 | 24 | 30.38 | 9 | 24.32 |
| | C5178A | D | 15 | 17.05 | 19 | 21.84 | 24 | 30.38 | 9 | 24.32 |
| | C3010T | D4 | 10 | 11.36 | 14 | 16.09 | 16 | 20.25 | 7 | 18.92 |
| | C5262T | D4j1 | 4 | 4.55 | 3 | 3.45 | 2 | 2.53 | | |
| | C10873T | N | 29 | 32.95 | 33 | 37.93 | 31 | 39.24 | 16 | 43.24 |
| | A663G | A | 10 | 11.36 | 5 | 5.75 | 10 | 12.66 | 3 | 8.11 |
| | T12705 | R | 19 | 21.59 | 26 | 29.89 | 21 | 26.58 | 12 | 32.43 |
| | C10310T | F | 11 | 12.50 | 14 | 16.09 | 10 | 12.66 | 3 | 8.11 |

其中单倍群 R 的分布频率约为 30%。

3 讨论

本研究结合个体水平的常染色体、父系 Y 染色体和母系线粒体多类别基因组信息，对中国西南地区 4 个藏族人群的遗传亚结构进行了研究，常染色体位点和 Y 染色体 SNPs 遗传结构分析结果表明 4 个藏族人群可分为 3 类：海拔最高的西藏藏族(CTT)为一类；海拔次之的四川甘孜藏族(CTG)和青海藏族(CTQ)为一类；海拔最低的四川阿坝藏族(CTA)为一类。线粒体 SNPs 遗传标记角度的分析结果表明不同海拔的藏族人群的遗传结构无显著差异。

常染色体遗传标记可从个体水平了解群体遗传结构。藏族人群在长期与环境相适应的进化过程中形成 HAA 机制，其 HAA 的生物学基础中，HIF 通路作为机体调控氧平衡的主要信号通路，是藏族人群 HAA 的关键，其分布的 HAA 相关基因(如 *EGLN1*、*EPAS1*、*PPARA*、*EDNRA*、*ATP6V1E2*、*HMOX2* 和

HBG 等)中多个位点在高原藏族和平原人群的频率具有差异性^[7,9,29]。使用一组较少的在人群中位基因频率差异显著的位点可实现人群遗传结构的初步解析。因此,本研究基于文献调研后最终筛选出在藏族人群与平原人群间具有显著差异的 25 个藏族特异性位点,分别位于 2 号染色体 HIF 通路的 *EPAS1* 和 *ATP6V1E2* 等 HAA 相关基因。各位点等位基因在不同藏族分支人群间的频率分布也存在地域特异性,与前人的相关研究符合^[5,8,30]。同时,各位点的等位基因频率分布差异与海拔呈正相关,同柯金坤^[31]的研究结果类似,进一步证明了高海拔环境对 HAA 相关基因存在适应性选择。当然,藏族人群的 HAA 机制研究不局限于热点研究的基因,如 *EPAS1* 下游的 *TMEM247* 基因中 rs116983452 的突变型等位基因在藏族人群和其他人群中表现出显著的频率差异^[32]; *MKL1* 基因中 rs59098711 的等位基因 C 高频分布于藏族人群^[33],此外该基因上 163 bp 片段缺失和 *SCUBE2* 基因中 662 bp 片段插入也在藏族人群中富集^[34]。为增加信息覆盖度,未来可进一步筛选藏族特异性位点对藏族不同分支人群进行研究,探索其遗传亚结构。

Y 染色体遗传标记是研究群体父系遗传信息的重要手段。根据 Y-SNPs 的发生时间可将男性人群分为不同单倍群,同一父系祖先的男性后代享有相同单倍群。东亚人群以 O-M175、C-M130、D-M174 和 N-M231 为主要单倍群,其中 O-M175 为优势单倍群,约 75% 的中国人群归于该单倍群; D-M174 在国内主要分布于藏族人群^[35]。本研究的 4 个藏族人群主要为 D-M174 和 O-M175 两种核心单倍群(79.80%),结果与文献^[36]报道一致。藏族人群特异性单倍群 D-M174 及其下游分支在不同藏族分支人群中的分布频率具有差异性。其中 D1a2-P99 在西藏藏族(CTT)、四川甘孜藏族(CTG)和青海藏族(CTQ)的频率明显高于 D1a1-M15,印证了前人研究结果^[35-37]。而四川阿坝藏族(CTA)以下游分支 D1a1-M15 为主,区别于其他藏族人群。此外,青海藏族(CTQ)中 J-M304 和 R-M207 的分布频率显著高于其他藏族人群,该两类单倍群在国内主要分布于回族人群(分别为 9.69% 和 11.85%)^[25],提示两人群间存在历史上的基因交流或部分共同父系祖源。Y-STRs 突变率较高,可反映人群近期父系基因交流。

4 个藏族人群的 HD 均高于 0.99,表明藏族人群的 Y-STRs 遗传多态性较高,其多态性分布的地区差异性较小。在网络图和 MDS 中,西藏藏族(CTT)、四川甘孜藏族(CTG)和青海藏族(CTQ)紧密聚集,也提示 3 者的父系遗传差异较小。值得注意的是,四川阿坝藏族(CTA)在 MDS 中与其他藏族人群相距较远,这与其常染色体遗传结构差异一致,不过也不排除其样本量较少导致的偏差。

线粒体 SNPs 是解析群体母系遗传结构的理想标记,其单倍群的支系分布具有族群地域差异^[38-40]。本研究的不同藏族分支人群中以主干单倍群 M9a'b、G、D、F 的分布频率较高,均为东亚起源支系且在藏、汉群体中均广泛分布^[41-44]。其中高频分布于中国西南地区的单倍群 D 是 4 个藏族人群最普遍的单倍群类型(17.05%~30.38%),其次为 F。而单倍群 M9a'b 广泛分布于东亚大陆^[45],国内人群中其频率及多样性在西藏藏族人群中达到最高^[36,42]。本研究中除四川阿坝藏族(CTA)外, M9a'b 在其他藏族人群中均大于 10%。单倍群 G 在 4 个藏族人群的分布频率约为 10%,类似于前人研究的那曲、拉萨、甘孜等地藏族人群的频率分布^[36]。然而藏族人群特异单倍群 M62 仅在青海藏族(CTQ)检出 1 例,低于文献报道的其在藏族人群中的分布频率(14/680=2.1%)^[42],可能为样本采集范围局限和样本量较少所致。

不同藏族分支人群的常染色体位点和 Y 染色体 SNPs 遗传信息具有显著差异性,且与海拔具有相关性。从地理位置分析,西南地区汇聚濮系、氏系、羌系及夷系民族群体^[46]。西藏藏族(CTT)居住于以藏族为主体民族(90.5%)的高海拔青藏高原地区,逶迤的山脉(如喜马拉雅山脉)和陡峭的峡谷(如雅鲁藏布大峡谷)等复杂多样的地貌形成地理隔离^[47],可能阻碍了西藏藏族(CTT)与其他人群的基因交流,导致其遗传结构更具有高原代表性。四川甘孜藏族(CTG)和青海藏族(CTQ)居于青藏高原周边地区,所在的四川甘孜处于青藏高原东南源,包含藏族(78.3%)和汉族(18.2%)等民族以大范围聚居小范围杂居形式分布^[48];青海位于青藏高原东北部,容纳汉族(53.0%)、藏族(24.4%)、回族(14.8%)等民族聚居(<http://www.stats.gov.cn>)。多民族混居的生活环境使得四川甘孜藏族(CTG)和青海藏族(CTQ)的遗传结构与西藏藏族(CTT)的差异显著,提示其与周围民族之间存在基

因交流。四川阿坝藏族(CTA)虽同样居于高原周边地区,但其所隶属的嘉绒藏族是夷系民族在唐代吐蕃东进后融合羌人和吐蕃人基因而形成的^[2,49],其所在的“藏彝走廊”位于横断山脉-河谷交错地区,是西北、西南各民族流动迁徙的重要场所。特殊的民族交流和地理位置,使四川阿坝藏族(CTA)呈现明显混融特征而区别于其他藏族人群。然而,线粒体 SNPs 遗传信息在藏族分支人群间无明显差异性,从社会学角度分析可能与藏族人群的婚配制度有关。藏族作为中国典型的少数民族之一,是氏羌族群、游牧民族和西藏土著居民融合的民族^[50],可与其他民族通婚。在公元 6 世纪,西藏松赞干布与唐朝文成公主的联姻打开了西部与中原地带的经济文化等方面的交流新格局^[51]。此后中原地区王朝对藏区设置行政管理机构使民族融合趋势加强,推进了藏汉通婚的进程^[50]。民族间基因交流尤其是母系基因的渗入使藏族人群的母系遗传多态性增加。

综上所述,本研究基于多类别基因组信息探讨了 4 个不同藏族分支人群的遗传亚结构,结果表明 4 个藏族人群间具有遗传差异性,获得的数据可为群体遗传学和法医学实践研究提供数据支持。但涉及的样本量少,且对甘肃、云南等地的藏族人群遗传亚结构尚未探索,未来可增加样本量及更多藏族分支人群开展研究,以进一步探索藏族不同分支的遗传相似性和差异性。

附录:

附表详见文章电子版 www.chinagene.cn。

参考文献(References):

- [1] Shi S. Tibetan regional characteristics and the related problems-and the Kham characteristic. *J Qinghai Natl Inst(Soc Sci)*, 2015, 41(1): 207–208.
石硕. 藏族的地域特点及相关问题——兼论康区之特点. 青海民族大学学报(社会科学版), 2015, 41(1): 207–208. [DOI]
- [2] Dejizhuoga, The origin of Jirarong Tibetan. *Tibet Stud*, 2004, (2): 51–56.
德吉卓嘎. 试论嘉绒藏族的族源. 西藏研究, 2004, (2): 51–56. [DOI]
- [3] Majmundar AJ, Wong WJ, Simon MC. Hypoxia-inducible factors and the response to hypoxic stress. *Mol Cell*, 2010, 40(2): 294–309. [DOI]
- [4] Beall CM. Tibetan and Andean Patterns of adaptation to high-altitude hypoxia. *Hum Biol*, 2000, 72(1): 201–228. [DOI]
- [5] Yi X, Liang Y, Huerta-Sanchez E, Jin X, Cuo ZXP, Pool JE, Xu X, Jiang H, Vinckenbosch N, Korneliussen TS, Zheng HC, Liu T, He WM, Li K, Luo RB, Nie XF, Wu HL, Zhao MR, Cao HZ, Zou J, Shan Y, Li SZ, Yang Q, Asan, Ni PX, Tian G, Xu JM, Liu X, Jiang T, Wu RH, Zhou GY, Tang MF, Qin JJ, Wang T, Feng SJ, Li GH, Huasang, Luosang J, Wang W, Chen F, Wang YD, Zheng XG, Li Z, Bianba Z, Yang G, Wang XP, Tang SH, Gao GY, Chen Y, Luo Z, Gusang L, Cao Z, Zhang QH, Ouyang WH, Ren XL, Liang HQ, Zheng HS, Huang YB, Li JX, Bolund L, Kristiansen K, Li YR, Zhang Y, Zhang XQ, Li RQ, Li SG, Yang HM, Nielsen R, Wang J, Wang J. Sequencing of 50 human exomes reveals adaptation to high altitude. *Science*, 2010, 329(5987): 75–78. [DOI]
- [6] Simonson TS, Yang Y, Huff CD, Yun H, Qin G, Witherspoon DJ, Bai Z, Lorenzo FR, Xing J, Jorde LB, Prchal JT, Ge R. Genetic evidence for high-altitude adaptation in Tibet. *Science*, 2010, 329(5987): 72–75. [DOI]
- [7] Huerta-Sánchez E, Jin X, Asan, Bianba Z, Peter BM, Vinckenbosch N, Liang Y, Yi X, He M, Somel M, Ni P, Wang B, Ou X, Huasang, Luosang J, Cuo ZX, Li K, Gao G, Yin Y, Wang W, Zhang X, Xu X, Yang H, Li Y, Wang J, Wang J, Nielsen R. Altitude adaptation in Tibetans caused by introgression of Denisovan-like DNA. *Nature*, 2014, 512(7513): 194–197. [DOI]
- [8] Jiang L, Peng JX, Huang MS, Liu J, Wang L, Ma Q, Zhao H, Yang X, Ji AQ, Li CX. Differentiation analysis for estimating individual ancestry from the Tibetan Plateau by an archaic altitude adaptation EPAS1 haplotype among East Asian populations. *Int J Legal Med*, 2018, 132(50): 1527–1535. [DOI]
- [9] Xiang K, Ouzhuluobu, Peng Y, Yang Z, Zhang X, Cui C, Zhang H, Li M, Zhang Y, Bianba, Gonggalanzi, Basang, Ciwangsangbu, Wu T, Chen H, Shi H, Qi X, Su B. Identification of a Tibetan-specific mutation in the hypoxic gene EGLN1 and its contribution to high-altitude adaptation. *Mol Biol Evol*, 2013, 30(8): 1889–1898. [DOI]
- [10] Wen B. Y chromosome, mtDNA polymorphism and genetic structure of East Asian population [Dissertation]. Shanghai: Fudan Univ, 2004.
文波. Y 染色体、mtDNA 多态性与东亚人群的遗传结构 [学位论文]. 复旦大学, 2004. [DOI]
- [11] Hammer MF, Karafet TM, Park H, Omoto K, Harihara S, Stoneking M, Horai S. Dual origins of the Japanese: common ground for hunter-gatherer and farmer Y chromosomes. *J Hum Genet*, 2006, 51(1): 47–58. [DOI]
- [12] Thangaraj K, Singh L, Reddy AG, Rao VR, Sehgal SC,

- Underhill PA, Pierson M, Frame IG, Hagelberg E. Genetic affinities of the Andaman Islanders, a vanishing human population. *Curr Biol*, 2003, 13(2): 86–93. [DOI]
- [13] Karafet T, Xu L, Du R, Wang W, Feng S, Wells RS, Redd AJ, Zegura SL, Hammer MF. Paternal population history of East Asia: sources, patterns, and microevolutionary processes. *Am J Hum Genet*, 2001, 69(3): 615–628. [DOI]
- [14] Bhandari S, Zhang XM, Cui CY, Bianba, Liao SY, Peng Y, Zhang H, Xiang K, Shi H, Ouzhuluobu, Baimakongzhu, Gonggalanzi, Liu SM, Gengdeng, Wu TY, Qi XB, Su B. Genetic evidence of a recent Tibetan ancestry to Sherpas in the Himalayan region. *Sci Rep*, 2015, 5: 16249. [DOI]
- [15] Huang MS, Ma Q, Wang L, Ma X, Li CX, Jiang L. The study of Tibetan ancestry informative SNPs on high-altitude adaptive genes. *Chin J Foren Med*, 2017, 32(6): 588–599. 黄美莎, 马泉, 王玲, 马新, 李彩霞, 江丽. 高原适应基因中藏族祖先信息位点的研究. 中国法医学杂志, 2017, 32(6): 588–599. [DOI]
- [16] Liu J, Li S, Jiang L, Zhao L, Zhao WT, Feng L, Liu HB, Ji AQ, Li CX. DNA ancestry analyzer: an automatic program for ancestry inference of unknown individuals. *Life Sci Res*, 2018, 22(1): 3–7. 刘京, 李盛, 江丽, 赵蕾, 赵雯婷, 丰蕾, 刘海渤, 季安全, 李彩霞. 对于未知来源个体进行族群推断的自动分析系统. 生命科学研究, 2018, 22(1): 3–7. [DOI]
- [17] Karafet TM, Mendez FL, Meilerman MB, Underhill PA, Zegura SL, Hammer MF. New binary polymorphisms reshape and increase resolution of the human Y chromosomal haplogroup tree. *Genome Res*, 2008, 18(5): 830–838. [DOI]
- [18] Bandelt HJ, Forster P, Röhl A. Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol*, 1999, 16(1): 37–48. [DOI]
- [19] Wang H, Mao J, Xia Y, Bai XG, Zhu WQ, Peng D, Liang WB. Genetic polymorphisms of 17 Y-chromosomal STRs in the Chengdu Han population of China. *Int J Legal Med*, 2017, 131(4): 967–968. [DOI]
- [20] Hidding M, Schmitt C. Haplotype frequencies and population data of nine Y-chromosomal STR polymorphisms in a German and a Chinese population. *Forensic Sci Int*, 2000, 113(1–3): 47–53. [DOI]
- [21] Shi MS, Li YB, Tang JP, Zhang HJ, Hou YP. Southwest China Han Population data for nine Y-STR loci by multiplex polymerase chain reaction. *J Forensic Sci*, 2007, 52(1): 228–230. [DOI]
- [22] Fan GY, An YR, Peng CX, Deng JL, Pan LP, Ye Y. Forensic and phylogenetic analyses among three Yi populations in Southwest China with 27 Y chromosomal STR loci. *Int J Legal Med*, 2018, 133(3): 795–797. [DOI]
- [23] He GL, Wang Z, Su YD, Zou X, Wang MG, Chen X, Gao B, Liu J, Wang SY, Hou YP. Genetic structure and forensic characteristics of Tibeto-Burman-speaking Ü-Tsang and Kham Tibetan Highlanders revealed by 27 Y-chromosomal STRs. *Sci Rep*, 2019, 9(1): 7739. [DOI]
- [24] Zhao Q, Bian Y, Zhang S, Zhu R, Zhou W, Gao Y, Li C. Population genetics study using 26 Y-chromosomal STR loci in the Hui ethnic group in China. *Forensic Sci Int Genet*, 2017, 28: e26–e27. [DOI]
- [25] Xie MK, Song F, Li JN, Lang M, Luo HB, Wang Z, Wu J, Li CZ, Tian CC, Wang WZ, Ma H, Song Z, Fan YJ, Hou YP. Genetic substructure and forensic characteristics of Chinese Hui populations using 157 Y-SNPs and 27 Y-STRs. *Forensic Sci Int Genet*, 2019, 41: 11–18. [DOI]
- [26] Nothnagel M, Fan GY, Guo F, He YF, Hou YP, Hu SP, Huang J, Jiang XH, Kim W, Kim K, Li CT, Li H, Li LM, Li SL, Li Z, Liang WB, Liu C, Lu D, Luo HB, Nie SJ, Shi MS, Sun HY, Tang JP, Wang L, Wang CC, Wang D, Wen SQ, Wu HY, Wu WW, Xing JX, Yan JW, Yan S, Yao HB, Ye Y, Yun LB, Zeng ZS, Zha L, Zhang SH, Zheng XF, Willuweit S, Roewer L. Revisiting the male genetic landscape of China: a multi-center study of almost 38,000 Y-STR haplotypes. *Hum Genet*, 2017, 136(5): 485–497. [DOI]
- [27] Lang M, Liu H, Song F, Qiao XH, Ye Y, Ren H, Li JN, Huang J, Xie MK, Chen SJ, Song MY, Zhang YF, Qian XQ, Yuan TX, Wang Z, Liu YM, Wang MG, Liu YC, Liu J, Hou YP. Forensic characteristics and genetic analysis of both 27 Y-STRs and 143 Y-SNPs in Eastern Han Chinese population. *Forensic Sci Int Genet*, 2019, 42: e13–e20. [DOI]
- [28] Kwak KD, Jin HJ, Shin DJ, Kim JM, Roewer L, Krawczak M, Tyler-Smith C, Kim W. Y-chromosomal STR haplotypes and their applications to forensic and population studies in east Asia. *Int J Legal Med*, 2005, 119(4): 195–201. [DOI]
- [29] Bigham AW, Lee FS. Human high-altitude adaptation: forward genetics meets the HIF pathway. *Genes Dev*, 2014, 28(20): 2189–2204. [DOI]
- [30] Peng Y, Yang ZH, Zhang H, Cui CY, Qi XB, Luo XJ, Tao X, Wu TY, Ouzhuluobu, Basang, Ciwangsangbu, Danzengduojie, Chen H, Shi H, Su B. Genetic variations in Tibetan populations and high-altitude adaptation at the Himalayas. *Mol Biol Evol*, 2011, 28(2): 1075–1081. [DOI]
- [31] Ke JK. A study of polymorphism sites of HIF-1 α and HIF-2 α gene in three Tibetan groups of different altitude [Dissertation]. *Peking Union Med Coll*, 2010. 柯金坤. 不同海拔三个藏族人群 HIF-1 α 基因与 HIF-2 α 基因多态位点的比较研究[学位论文]. 北京协和医学院, 2010. [DOI]
- [32] Deng L, Zhang C, Yuan K, Gao Y, Pan YW, Ge XL, He YX, Yuan Y, Lu Y, Zhang XX, Chen H, Lou HY, Wang XJ, Lu DS, Liu JJ, Tian L, Feng QD, Khan A, Yang YJ, Jin ZB, Yang J, Lu F, Qu J, Kang LL, Su B, Xu SH. Prioritizing natural-selection signals from the deep-sequencing genomic data suggests multi-variant adaptation in Tibetan highlanders. *Natl Sci Rev*, 2019, 6(6): 1201–1222. [DOI]

- [33] Ouzhuluobu, He YX, Lou HY, Cui CY, Deng L, Gao Y, Zheng WS, Guo YB, Wang XJ, Ning ZL, Li J, Li B, Bai CJ, Baimakangzhuo, Gonggalanzi, Dejiquzong, Bianba, Duoizhuoma, Liu SM, Wu TY, Xu SH, Qi XB, Su B. De novo assembly of a Tibetan genome and identification of novel structural variants associated with high-altitude adaptation. *Natl Sci Rev*, 2019, 7(2): 391–402. [DOI]
- [34] Zhang Q, Ping J, Zhang HX, Kang B, Li YF, Zhou GQ. Genetic association of MKL1 gene polymorphisms with the high-altitude adaptation. *Hereditas(Beijing)*, 2019, 41(7): 634–643.
张晴, 平杰, 张昊翔, 康波, 李元丰, 周钢桥. MKL1 基因多态性与高原环境适应性的遗传关联研究. *遗传*, 2019, 41(7): 634–643. [DOI]
- [35] Wang CC, Li H. Inferring human history in East Asia from Y chromosomes. *Investig Genet*, 2013, 4(1): 11. [DOI]
- [36] Qi XB, Cui CY, Peng Y, Zhang XM, Yang ZH, Zhong H, Zhang H, Xiang K, Cao XY, Wang Y, Ouzhuluobu, Basang, Ciwangsangbu, Bianba, Gonggalanzi, Wu TY, Chen H, Shi H, Su B. Genetic evidence of paleolithic colonization and neolithic expansion of modern humans on the Tibetan Plateau. *Mol Biol Evol*, 2013, 30(8): 1761–1778. [DOI]
- [37] Wen B, Xie X, Gao S, Li H, Shi H, Song X, Qian T, Xiao C, Jin J, Su B, Lu D, Chakraborty R, Jin L. Analyses of genetic structure of Tibeto-Burman populations reveals sex-biased admixture in southern Tibeto-Burmans. *Am J Hum Genet*, 2004, 74(5): 856–865. [DOI]
- [38] Cavalli-Sforza LL, Feldman MW. The application of molecular genetic approaches to the study of human evolution. *Nat Genet*, 2003, 33(Suppl.): 266–275. [DOI]
- [39] Chu X, Shan KR, Wen B, Qi XL, Li Y, Wu CX, Liu X, Zhao Y, Ren XL, Jin L. Analysis of polymorphisms in Y-DNA haplotypes and mtDNA haplogroups in Yao ethnic group from Guizhou. *Hereditas(Beijing)*, 2006, 28(2): 153–158.
褚迅, 单可人, 文波, 齐晓岚, 李毅, 吴昌学, 刘烜, 赵艳, 任锡麟, 金力. 贵州瑶族 3 支系 Y-DNA 及线粒体 DNA 序列多态性分析. *遗传*, 2006, 28(2): 153–158. [DOI]
- [40] Wang XQ, Wang CC, Deng QY, Li H. Genetic analysis of Y chromosome and mitochondrial DNA polymorphism of Mulam ethnic group in Guangxi, China. *Hereditas (Beijing)*, 2013, 35(2): 168–174.
王晓庆, 王传超, 邓琼英, 李辉. 广西仫佬族 Y 染色体和 mtDNA 的遗传结构分析. *遗传*, 2013, 35(2): 168–174. [DOI]
- [41] Gu ML, Wang YJ, Shi L, Jiang F, Qiu MJ, Lin KQ, Tao YF, Shi L, Huang XQ, Liu B, Chu JY. Comparative analysis of the complete mitochondrial genome between Tibetan and Han population. *Chin J Med Genet*, 2008, 25(4): 382–386.
顾明亮, 汪业军, 史磊, 姜枫, 邱梦洁, 林克勤, 陶玉芬, 史荔, 黄小琴, 刘斌, 褚嘉佑. 藏汉民族线粒体基因组全序列的比较研究. *中华医学遗传学杂志*, 2008, 25(4): 382–386. [DOI]
- [42] Zhao M, Kong QP, Wang HW, Peng MS, Xie XD, Wang WZ, Jiayang, Duan JG, Cai MC, Zhao SN, Cidanpingcuo, Tu YQ, Wu SF, Yao YG, Bandelt HJ, Zhang YP. Mitochondrial genome evidence reveals successful Late Paleolithic settlement on the Tibetan Plateau. *Proc Natl Acad Sci USA*, 2009, 106(50): 21230–21235. [DOI]
- [43] Qin ZD, Yang YJ, Kang LL, Yan S, Cho K, Cai XY, Lu Y, Zheng HX, Zhu DC, Fei DM, Li SL, Jin L, Li H. A mitochondrial revelation of early human migrations to the Tibetan Plateau before and after the last glacial maximum. *Am J Phys Anthropol*, 2010, 143(4): 555–569. [DOI]
- [44] Peng MS, Palanichamy MG, Yao YG, Mitra B, Cheng YT, Zhao M, Liu J, Wang HW, Pan H, Wang WZ, Zhang AM, Zhang W, Wang D, Zou Y, Yang Y, Chaudhuri TK, Kong QP, Zhang YP. Inland post-glacial dispersal in East Asia revealed by mitochondrial haplogroup M9a'b. *BMC Biol*, 2011, 9: 2. [DOI]
- [45] Soares P, Trejaut JA, Loo JH, Hill C, Mormina M, Lee CL, Chen YM, Hudjashov G, Forster P, Macaulay V, Bulbeck D, Oppenheimer S, Lin M, Richards MB. Climate change and postglacial human dispersals in southeast Asia. *Mol Biol Evol*, 2008, 25(6): 1209–1218. [DOI]
- [46] Meng M. Exploration of the relationship of Yi and Qiang in the Southwest ethnic groups of Han dynasty. *Hist Res*, 1985, 1: 11–32.
蒙默. 试论汉代西南民族中的“夷”与“羌”. *历史研究*, 1985, 1: 11–32. [DOI]
- [47] Zheng D, Zhang QS, Wu SH. Mountain Geocology and Sustainable Development of the Tibetan Plateau. *Geojournal Library*, 2000, 57(2): 203–204. [DOI]
- [48] Wang CM. The changes of population size and structure in Garze prefecture based on the census data. *J Sichuan Univ Natl*, 2013, 22(3): 58–67.
王长明. 基于人口普查数据的甘孜州人口规模与结构变化. *四川民族学院学报*, 2013, 22(3): 58–67. [DOI]
- [49] Li Q. The relationship between Jiarong, Jialiangyi, Ranmang and Ge populations – and explore the origin of Jiarong Tibetan. *J Sichuan Univ Natl*, 2010, 19(4): 1–6.
李青. 试论嘉绒、嘉良夷、冉驢与戈人的关系——兼论嘉绒藏族的族源. *四川民族学院学报*, 2010, 19(4): 1–6. [DOI]
- [50] Xie L. On the cultural reform of the Han-Tibetan intermarriage and its evolution. *J Neijiang Norm Univ*, 2006, 21(1): 147–149.
谢蕾. 藏汉通婚的文化整合及演变. *内江师范学院学报*, 2006, 21(1): 147–149. [DOI]
- [51] Yan MS. History of Marriage in Ancient China. Guizhou: Guizhou Ethnic Publishing House, 2003.
阎明恕. 中国古代和亲史. 贵州: 贵州民族出版社, 2003. [DOI]

Supplementary Table 1 The information of selected makers

附表 2 网络图分析中各 Y-STR 基因座的权重值信息

Supplementary Table 2 The weighting information of each Y-STR loci in Network analysis

| 基因座 | 权重 | 突变率 | CTT | CTG | CTQ | CTA |
|-----------|----|-------|--------|--------|--------|--------|
| DYS458 | 1 | 0.006 | 0.7750 | 0.7704 | 0.7943 | 0.8208 |
| DYS392 | 1 | 0.001 | 0.6545 | 0.7348 | 0.7124 | 0.8350 |
| DYS390 | 1 | 0.002 | 0.7421 | 0.7396 | 0.7185 | 0.6999 |
| DYS389 II | 2 | 0.004 | 0.6117 | 0.7565 | 0.7127 | 0.7112 |
| DYS439 | 2 | 0.005 | 0.6411 | 0.6108 | 0.6960 | 0.7098 |
| DYS389 I | 2 | 0.003 | 0.6540 | 0.6742 | 0.6501 | 0.6316 |
| DYS635 | 2 | 0.004 | 0.5924 | 0.6063 | 0.6308 | 0.7269 |
| DYS448 | 2 | 0.002 | 0.5812 | 0.6074 | 0.6208 | 0.7226 |
| DYS456 | 2 | 0.004 | 0.5687 | 0.5927 | 0.6399 | 0.6821 |
| GATA-H4 | 3 | 0.003 | 0.5927 | 0.6408 | 0.5652 | 0.5576 |
| DYS533 | 3 | 0.004 | 0.5457 | 0.5798 | 0.5815 | 0.5875 |
| DYS393 | 4 | 0.001 | 0.4814 | 0.5154 | 0.5580 | 0.5349 |
| DYS19 | 4 | 0.002 | 0.4990 | 0.5256 | 0.5363 | 0.5021 |
| DYS438 | 5 | 0.001 | 0.3933 | 0.4762 | 0.5434 | 0.4964 |
| DYS437 | 6 | 0.001 | 0.3202 | 0.3145 | 0.2906 | 0.5220 |
| DYS391 | 6 | 0.002 | 0.2743 | 0.3195 | 0.2611 | 0.3514 |

注: 突变率来源 YHRD 数据库(https://yhrd.org/pages/resources/mutation_rates)。