

· 综 述 ·

· 人体表征分子鉴识专题 ·

从“非洲起源说”研究到“DNA 族群地域推断”侦查应用

江丽, 刘京, 李彩霞

(公安部物证鉴定中心 现场物证溯源技术国家工程实验室 法医遗传学公安部重点实验室 北京市现场物证检验工程技术研究中心, 北京 100038)

摘 要: 伴随全基因组测序、生物信息等技术的发展, 单核苷酸多态性(single nucleotide polymorphism, SNP)、插入/缺失(insertion/deletion, InDel)等遗传标记被发现并大量应用, 从遗传学角度逐步研究获知了现代人群的起源和遗传结构。法医遗传学领域族群地域来源推断研究也应运而生并快速发展, 为案件侦办过程中确定侦查方向、缩小排查范围提供了线索和科学依据。本文对目前族群推断技术的研究进展、推断方法、发展趋势等情况进行综述。

关键词: 法医遗传学; 多态性, 单核苷酸; 非洲起源; 族群地域推断; 东亚

中图分类号: DF795.2 文献标志码: A doi: 10.12116/j.issn.1004-5619.2019.05.003

文章编号: 1004-5619(2019)05-0519-06

From the “Out of Africa Theory” Study to the Application of “DNA Biogeographic Ancestry Inference” in Criminal Investigation

JIANG Li, LIU Jing, LI Cai-xia

(National Engineering Laboratory for Crime Scene Evidence Investigation and Examination, Key Laboratory of Forensic Genetics, Beijing Engineering Research Center of Crime Scene Evidence Examination, Ministry of Public Security, Institute of Forensic Science, Beijing 100038, China)

Abstract: Genetic markers, such as single nucleotide polymorphism (SNP), insertion/deletion (InDel), were discovered and widely used with the development of whole genome sequencing and bioinformatics technology. The origin and genetic structure of the modern population had been gradually revealed from the perspective of genetics. The study on biogeographic ancestry inference in the field of forensic genetics emerged and developed rapidly, providing clues and scientific basis for the determination of investigation direction and for the narrow of the scope of investigation in the process of case investigation. This paper briefly reviews the research progress, inference methods and development trends of DNA ancestry inference technology.

Keywords: forensic genetics; polymorphism, single nucleotide; out of Africa; biogeographic ancestry inference; East Asia

基于骨骼形态等体质人类学方法可以推断样本的族群来源, 但该技术很难用于血斑、精斑等斑迹类或者毁损严重的骨骼等物证的检测。现代遗传学研究已经勾画出一幅人类起源、进化、迁移和融合图, 揭示出人群之间的遗传差异和群体遗传结构。以往相关研究^[1-4]促进了DNA族群地域推断(biogeographic ancestry inference)技术的诞生, 即检测不同人群之间具有遗传分布差异的位点判断DNA供者所属的族群地域, 这种差异位点被称为祖先信息位点(ancestry

informative marker, AIM)。

1 DNA族群推断技术研究进展

1.1 现代人群遗传结构研究

根据肤色、毛发颜色和形态、面貌等体貌特征, 人类学研究通常将人类群体大致划分为东亚黄种人(蒙古人种)、欧洲白种人(高加索人种)和非洲黑种人(尼格罗人种), 棕色人种(澳大利亚人种)通常被包括在黑色人种内^[5]。与五大洲相对应的五分法也是比较普

基金项目: 国家自然科学基金资助项目(81772027); “十三五”国家重点研发计划资助项目(2017YFC0803501); 中央级公益性科研院所基本科研业务费资助项目(2017JB025, 2017JB027); 2018首都科技领军人才培养工程资助项目(Z18110006318006)

作者简介: 江丽(1989—), 女, 博士, 副主任法医师, 主要从事法医物证学研究; E-mail: jl@mail.bnu.edu.cn

通信作者: 李彩霞, 女, 博士, 主任法医师, 主任从事法医物证学研究; E-mail: licaixia@tsinghua.org.cn

遍的分类法,即蒙古人种、高加索人种、尼格罗人种、澳大利亚人种、印第安人种^[6-7]。线粒体DNA和Y染色体DNA等遗传学研究证实现代人类的祖先晚期智人在7~10万年前走出非洲,经南线为主的路线逐渐扩散和迁移到各个大陆^[6-8]。由于遗传漂变、适应性进化、迁移混合、古DNA渗入等因素,形成了不同地域人群的遗传结构和体貌特征。人类基因组中99.9%的序列相同,微小差异形成了世界各地人群的地理特异性。绝大多数的遗传差异发生在群体内部,大洲之间人群的遗传差异高于大洲内部人群之间的遗传差异^[9]。基于短串联重复(shorten tandem repeat, STR)和单核苷酸多态性(single nucleotide polymorphism, SNP)的研究表明,世界人群的遗传结构与地理分布明显相关^[10-12],随着地理距离的增加,等位基因频率的差异逐渐加大,即不同族群之间没有绝对的边界。但是这种渐变分布存在不连续性,往往与地理、语言和宗教等的隔绝具有一致性,故而形成了不同地域人群的遗传亚结构^[9]。

1.2 洲际人群推断体系研究

1973年,NEEL首次提出人群“私有”遗传变异(“private” genetic variants)的概念^[13],STR、SNP、插入/缺失(insertion/deletion, InDel)和微单倍型(microhaplotype)等具有人群特异性的遗传标记陆续被报道^[14-16]。由于STR突变率较高,不是最理想的祖先信息标记,SNP、InDel是目前选择AIM的主要遗传标记。检测少量的AIM即可实现洲际人群的遗传推断,洲际人群内部的进一步区分则需更多的AIM。但是位点数目并非越多越好,遗传距离不同的族群,位点筛选标准和参考数据集不同,需采用不同位点组合从洲际到亚人群逐层推断。目前,国内外文献报道了大量族群推断体系,区分度为3~5个洲际人群(如东亚、欧洲、非洲等)^[17],其中,34-SNP^[18]、27-SNP^[19]、31-SNP^[20]等基于聚合酶链反应-毛细管电泳(polymerase chain reaction-capillary electrophoresis, PCR-CE)平台(以下简称为“PCR-CE平台”)建立了适合法医学应用的检测体系,灵敏度达到皮克级,适于各种现场生物检材,检测时间约7 h。InDel是一种二态的长度多态性遗传标记(如46-InDel^[21]、21-InDel^[22]等),约3 h可以完成PCR-CE平台检测。耶鲁大学KIDD教授实验室基于KOSOY等^[23]研究的128-SNP等位点筛选出55-SNP^[24],实现了撒哈拉以南非洲、北非、西南亚、欧洲、南亚、东亚、大洋洲、美洲印第安人种的区分。美国Thermo Fisher Scientific、美国Illumina等公司^[25-26]采纳Seldin 128-SNP^[23]、Kidd 55-SNP等位点构建了下一代测序(next generation sequencing, NGS)检测体系或者芯片检测体系。

1.3 东亚人群推断相关研究

1.3.1 东亚人群遗传结构研究

东亚人口占全球人口的22%,是研究人类源流历史及民族演化的重要地区之一,“非洲起源说”认为现代人到达东亚的时间为5~6万年前^[27-28]。Y-SNP、mtDNA和常染色体SNP等研究均表明东亚人群存在明显的南北分化,北方人群由于受到来自中亚和欧洲遗传成分的影响,呈现东西走向的变化趋势,南北方人群遗传的差异以长江为地理分界^[29-32]。汉族具有混合特征,呈现明显的南北分化,汉族人群与当地少数民族之间的遗传差异小于南北方汉族之间的遗传差异。南北方人群的遗传成分对当前南方汉族人群基因库的贡献具有性别偏向性,北方人群的父亲遗传成分和南方人群的母亲遗传成分分别构成了现代南方汉族人群基因库的主体^[33]。东亚人群分属汉藏语系、苗瑶语系、侗台语系、南亚语系、南岛语系及阿尔泰语系等^[34],东亚人群的遗传结构与族源历史和语言结构具有对应关系,同一语系人群有聚类倾向^[29,35]。

南亚语系、侗台语系和苗瑶语系人群分别对应于中国历史记载的南方的百濮、百越和南蛮人群^[36-37]。汉藏语系中的汉语族人群以黄河中下游的古代华夏族为主体,逐渐融合周围其他民族形成^[36]。藏缅语族起源于甘肃、宁夏和青海等中国西北的氏羌人群,沿藏彝走廊大规模迁移至西藏、云南等中国西南地区,经历了与南方人群的基因融合^[38]。其中,藏族人群的基因库中保留了新石器时代中国北方人群的遗传成分(氏羌人群)和旧石器时代定居青藏高原的人群的遗传组分(Y-SNP的D-M174和mtDNA的M16、A10等)^[39-41]。在藏族人群中,发现EPAS1、EGLN1等高原适应基因^[42-43]。EPAS1基因高原适应单倍型在藏族人群的频率为72.32%,在平原地区人群的频率小于2.5%^[44]。阿尔泰语系属于北方人群,蒙古语族和满通古斯语族起源于中国东北古代少数民族^[45]。突厥语族人群主要分布在中国西北地区,南方起源的单倍群O、C、D、N的频率占Y染色体所有单倍群频率的64.36%,表明在中国西北人群中东亚的Y染色体谱系占主导地位^[46]。日本、韩国人群的语系归属存在争议,日本人群的二元遗传结构模式被广泛接受,即绳文人和弥生人的遗传混合^[47-48]。汉族、日本、朝鲜人群的遗传结构存在差异,也存在基因流^[49]。

1.3.2 东亚人群的DNA族群推断研究

目前,针对法医学应用的东亚人群推断体系研究报道较少。使用前述报道的洲际人群推断体系,东亚人群的北亚类型(亚洲北部)、南亚类型(长江以南至东南亚)和东亚类型(蒙古高原至长江以北)往往表现

为一种遗传主成分。LI等^[50]针对法医学应用筛选了74个位点,并分别构建了基于微流控芯片和CE平台的检测体系,DNA模板用量为纳克级,该体系可实现北非、西南亚与欧洲人群的区分以及北亚、东南亚与东亚人群的区分,其中北方汉族和南方汉族分别表现出北亚和东南亚人群的遗传混合。WANG等^[49,51]筛选出南北方汉族人群以及汉族、朝鲜和日本人群相关的祖先信息位点。YUASA等^[52]研究发现了67个可能源自绳文人的日本人群特异SNP位点,其中rs3778922(*GALNT11*),rs76162918(*H19*)和rs2285715(*PLA2G12A*)在日本人群中的特异性最高。

2 DNA族群地域推断基本方法

常用流程包括AIM的筛选与评估、复合检测体系构建、参考人群分型库建立、推断算法和软件设计、体系和算法的验证评估等。

2.1 AIM的筛选

常用统计学指标如下:

δ 值是两个群体之间等位基因频率的差值, δ 值越大代表该位点在两个群体之间的频率差异越大,通用公式如下^[14]:

$$\delta = \frac{\sum_{x=1}^k |p_x - q_x|}{2}, \quad (1)$$

$\delta \geq 0$, k 为该位点的等位基因数, p_x 和 q_x 分别代表等位基因 x 在群体 p 和群体 q 中的频率。如果遗传标记位点为双等位基因, δ 值计算的简化公式为:

$$\delta = |p_x - p_y| \text{ 或 } |q_x - q_y|, \quad (2)$$

其中 p_x 和 p_y 是群体 x 和 y 中的一个等位基因 p 的频率, q_x 和 q_y 是群体 x 和 y 中的另一个等位基因 q 的频率。

Wright's F_{st} 是群体遗传学中衡量群体间分化程度的一个重要指标^[53], 也叫做固定指数(fixation index)。当一个大的群体分化成相互隔离的数个亚群以后, 与未分化之前相比, 总体杂合度会降低。 F_{st} 值的大小反应了每个位点的等位基因频率在不同群体间的变化程度, F_{st} 值越大, 该位点在不同人群间的等位基因频率差别越大。取值范围为0~1, 0表示没有群体分化, 1表示完全隔离, 而实际观察到的值往往远小于1。在Hardy-Weinberg平衡的前提下:

$$F_{st} = \frac{H_t - H_s}{H_t}, \quad (3)$$

式中, H_t 为总群体的杂合度(total heterozygosity), H_s 为亚群体的平均杂合度(average subpopulations heterozygosity)。

I_n 值(informativeness for assignment)^[54]也是常用的衡量AIM位点信息量的指标, 信息量与 F_{st} 值类似^[20]。

计算公式如下:

$$I_n(Q; J) = \sum_{j=1}^N (-p_j \log p_j + \sum_{i=1}^K \frac{p_{ij}}{K} \log p_{ij}), \quad (4)$$

式中, Q 为人群, 取值 $i=1-K$; J 为等位基因, 取值 $j=1-N$ 。

SNP为双等位基因 $j=1-2$, 公式变形为:

$$I_n(Q; J) = (-p_A \log p_A + \sum_{i=1}^K \frac{p_{iA}}{K} \log p_{iA}) + (-p_a \log p_a + \sum_{i=1}^K \frac{p_{ia}}{K} \log p_{ia}) \quad (5)$$

2.2 位点的评估与优选

2.2.1 族群聚类分析

基于贝叶斯的model-based聚类方法, 通过SNP、STR等遗传标记的分型数据来推测群体的聚类群组, 并把每个个体分配到这些群组中, 如果某个个体有混合遗传成分, 则被分配到两个或更多群组中。通过分析可以获知每个人群和个体的群组成分构成或者祖先成分(ancestry component)。使用Structure软件可以评估AIM位点达到的人群区分度, 并确定最稳定的群组数目, 即 K 值^[55]。DISTRUCT或CLUMPAK(<http://clumpak.tau.ac.il/index.html>)可将结果绘制成图。

2.2.2 主成分分析

主成分分析(principal component analysis, PCA)是从多个指标之间的相互关系入手, 利用降维思想通过少数几个主成分来揭示多个变量间的内部结构, 从原始变量中导出少数几个主成分, 使他们尽可能多地保留原始变量的信息。每个主成分原始数据的线性组合, 仅代表一部分变量, 第一主成分代表了最多的信息量, 其次是第二主成分、第三主成分等。在人群遗传结构分析中, 一般使用多个遗传标记的等位基因频率进行分析, 每个遗传标记作为一种指标, 把所有遗传位点揭示出的主要人群结构反映出来, 通常选择前三个主成分, 以PC1-PC2、PC1-PC3二维的形式展现。

此外, 反映人群间遗传距离的系统进化树分析也可评估AIM位点对人群的区分度。位点的筛选与评估是一个逐步减少位点和评估的过程。

2.3 复合检测体系构建

筛选到AIM后需构建复合检测体系, 才可用于法医现场生物物证的检测。检测体系需满足微量DNA检测需求。Sanger测序、NGS等技术是序列检测的金标准, 但对于法医学应用而言, 仍需构建便捷灵敏的复合检测体系, 目前常用技术包括单碱基延伸结合毛细管电泳检测技术、单碱基延伸结合质谱检测技术、等位基因特异PCR、基因芯片等。

2.4 未知个体的族群推断

决定DNA族群推断体系的区分度和准确性的关

键因素之一是参考人群库的选择和使用。筛选到AIM位点后,评估确定人群区分度,然后选择每个群体成分较高的个体作为参考人群库样本。对于未知个体,检测获得AIM位点分型之后,通过参考人群库的分型和频率数据计算样品的群体匹配概率(matching probability,MP)、群体似然比(likelihood ratio,LR)、多元逻辑回归(multivariate logistic regression,MLR)、祖先成分(ancestry component),绘制个体的族群归类图(classification of unknown individual),即可推断样品的族群来源^[56-57]。

3 应用情况和发展趋势

案件侦查前期推断涉案人员的族群地域,有助于确定侦查方向,缩小嫌疑人排查范围,协助案件定性,降低工作量,提高效率,已经逐步在案件中应用。SUN等^[58]利用27-plex SNP族群推断技术对1例骨骼样本进行检测,并推断出该样本来自欧洲族群的可能性最大。由于该技术方法较新,实验技术没有在法医学实验室普及,实验结果仍需要法医遗传学专业人士提供解读,因此应用范围仍然局限于部分地区及机构。另外,对于美国等移民国家,族群来源推断技术有助于了解人员的家族来源,23andme、Ancestry、Parabon等商业化公司为社会大众提供相关的收费服务。国内也有类似商业公司提供消费检测服务,提供祖源推断结果。

但法医族群推断研究技术体系目前仍存在如下不足:(1)大量的洲际人群区分AIM组合包含的位点和检测的人群各有差异,缺乏统一的位点组合和全球普适的参考人群库,也缺乏国际公认的评价标准。(2)东亚人群尤其是我国人群精细区分体系急需研究。(3)DNA族群推断反映的是样本的遗传结构特点,与省份、民族等民众熟知的信息并不完全对应,而侦查办案人员大多不了解群体遗传学,结果解析是目前面临的问题之一。未来DNA族群地域推断技术将向技术的标准化、区分的精细化等方向发展,伴随宏基因组、表观基因组等技术的发展,将会筛选出大量地域、饮食等相关的遗传标记,多遗传标记综合应用将会更全面精细地刻画个体的族群和地域来源。

总之,法医族群推断研究技术体系是法医遗传学重要的研究方向之一,是从传统的“比对识别”向新型“主动侦查”模式迈出的重要一步。未来结合DNA表型特征刻画和家族搜索等技术形成新一代法医DNA技术体系,将全面提升法医DNA在“刻画搜索和提供线索”方面的能力,为涉外反恐、跨区域犯罪、冷案积案等疑难案件的侦破提供科技支撑。

参考文献:

- [1] BROWN E M, BARRATT B J. Chapter 3: The HapMap - A haplotype map of the human genome[M]// BARNES M R. Bioinformatics for geneticists: A bioinformatics primer for the analysis of genetic data. 2nd ed. West Sussex: John Wiley & Sons, Ltd, 2007.
- [2] AUTON A, BROOKS L D, DURBIN R M, et al. A global reference for human genetic variation[J]. Nature, 2015, 526(7571): 68-74.
- [3] International HapMap Consortium. The international HapMap project[J]. Nature, 2003, 426(6968): 789-796.
- [4] SHERRY S T, WARD M H, KHOLODOV M, et al. dbSNP: The NCBI database of genetic variation[J]. Nucleic Acids Res, 2001, 29(1): 308-311.
- [5] 吴新智, 崔娅铭. 人种及其演变[J]. 科学通报, 2016, 61(34): 3630-3637.
- [6] UNDERHILL P A, SHEN P, LIN A A, et al. Y chromosome sequence variation and the history of human populations[J]. Nat Genet, 2000, 26(3): 358-361.
- [7] 李辉. 从基因中重新认识人类进化历程[J]. 现代人类学通讯, 2013, 7(e1): 1-11.
- [8] 田娇阳, 李玉春, 孔庆鹏, 等. 遗传学视角下东亚人群的起源和演化[J]. 遗传, 2018, 40(10): 814-824.
- [9] LI J Z, ABSHER D M, TANG H, et al. World-wide human relationships inferred from genome-wide patterns of variation[J]. Science, 2008, 319(5866): 1100-1104.
- [10] ROSENBERG N A, PRITCHARD J K, WEBER J L, et al. Genetic structure of human populations[J]. Science, 2002, 298(5602): 2381-2385.
- [11] ROSENBERG N A, MAHAJAN S, RAMACHANDRAN S, et al. Clines, clusters, and the effect of study design on the inference of human population structure[J]. PLoS Genet, 2005, 1(6): e70.
- [12] SHRIVER M D, MEI R, PARRA E J, et al. Large-scale SNP analysis reveals clustered and continuous patterns of human genetic variation[J]. Hum Genomics, 2005, 2(2): 81-89.
- [13] NEEL J V. "Private" genetic variants and the frequency of mutation among South American Indians[J]. Proc Natl Acad Sci USA, 1973, 70(12): 3311-3315.
- [14] SHRIVER M D, SMITH M W, JIN L, et al. Ethnic-affiliation estimation by use of population-specific DNA markers[J]. Am J Hum Genet, 1997, 60(4): 957-964.
- [15] FRUDAKIS T, VENKATESWARLU K, THOMAS M J, et al. A classifier for the SNP-based inference of ancestry[J]. J Forensic Sci, 2003, 48(4): 771-782.
- [16] KIDD K K, SPEED W C, PAKSTIS A J, et al. Evaluating 130 microhaplotypes across a global set of 83 populations[J]. Forensic Sci Int Genet, 2017,

- 29:29-37.
- [17] SOUNDARARAJAN U, YUN L, SHI M, et al. Minimal SNP overlap among multiple panels of ancestry informative markers argues for more international collaboration[J]. *Forensic Sci Int Genet*, 2016, 23:25-32.
 - [18] PHILLIPS C, SALAS A, SANCHEZ J J, et al. Inferring ancestral origin using a single multiplex assay of ancestry-informative marker SNPs[J]. *Forensic Sci Int Genet*, 2007, 1(3/4):273-280.
 - [19] WEI Y L, WEI L, ZHAO L, et al. A single-tube 27-plex SNP assay for estimating individual ancestry and admixture from three continents[J]. *Int J Legal Med*, 2016, 130(1):27-37.
 - [20] de la PUENTE M, SANTOS C, FONDEVILA M, et al. The Global AIMs Nano set: A 31-plex SNaP-shot assay of ancestry-informative SNPs[J]. *Forensic Sci Int Genet*, 2016, 22:81-88.
 - [21] SANTOS C, PHILLIPS C, OLDONI F, et al. Completion of a worldwide reference panel of samples for an ancestry informative InDel assay[J]. *Forensic Sci Int Genet*, 2015, 17:75-80.
 - [22] ZAUMSEGEL D, ROTHSCCHILD M A, SCHNEIDER P M. A 21 marker insertion deletion polymorphism panel to study biogeographic ancestry[J]. *Forensic Sci Int Genet*, 2013, 7(2):305-312.
 - [23] KOSOY R, NASSIR R, TIAN C, et al. Ancestry informative marker sets for determining continental origin and admixture proportions in common populations in America[J]. *Hum Mutat*, 2009, 30(1):69-78.
 - [24] KIDD K K, SPEED W C, PAKSTIS A J, et al. Progress toward an efficient panel of SNPs for ancestry inference[J]. *Forensic Sci Int Genet*, 2014, 10:23-32.
 - [25] KEATING B, BANSAL A T, WALSH S, et al. First all-in-one diagnostic tool for DNA intelligence: genome-wide inference of biogeographic ancestry, appearance, relatedness, and sex with the Identitas v1 Forensic Chip[J]. *Int J Legal Med*, 2013, 127(3):559-572.
 - [26] WARSHAUER D H, DAVIS C P, HOLT C, et al. Massively parallel sequencing of forensically relevant single nucleotide polymorphisms using TruSeq forensic amplicon[J]. *Int J Legal Med*, 2015, 129(1):31-36.
 - [27] SU B, XIAO J, UNDERHILL P, et al. Y-Chromosome evidence for a northward migration of modern humans into Eastern Asia during the last Ice Age[J]. *Am J Hum Genet*, 1999, 65(6):1718-1724.
 - [28] WANG C C, LI H. Inferring human history in East Asia from Y chromosomes[J]. *Investig Genet*, 2013, 4(1):11.
 - [29] ZHANG F, SU B, ZHANG Y P, et al. Genetic studies of human diversity in East Asia[J]. *Philos Trans R Soc Lond B Biol Sci*, 2007, 362(1482):987-995.
 - [30] 文波. Y染色体、mtDNA多态性与东亚人群的遗传结构[D]. 上海:复旦大学, 2003.
 - [31] 杜若甫. 我国的人类群体遗传学研究[J]. *生物学通报*, 1997(7):9-12.
 - [32] XU S, YIN X, LI S, et al. Genomic dissection of population substructure of Han Chinese and its implication in association studies[J]. *Am J Hum Genet*, 2009, 85(6):762-774.
 - [33] WEN B, LI H, LU D, et al. Genetic evidence supports demic diffusion of Han culture[J]. *Nature*, 2004, 431(7006):302-305.
 - [34] 蔡晓云. Y染色体揭示的早期人类进入东亚和东亚人群特征形成过程[D]. 上海:复旦大学, 2009.
 - [35] ABDULLA M A, AHMED I, ASSAWAMAKIN A, et al. Mapping human genetic diversity in Asia[J]. *Science*, 2009, 326(5959):1541-1545.
 - [36] 杜若甫, 葉傳昇. 中国的民族[M]. 北京:科学出版社, 1994.
 - [37] 李辉. 百越遗传结构的一元二分迹象[J]. *广西民族研究*, 2002(4):26-31.
 - [38] WEN B, XIE X, GAO S, et al. Analyses of genetic structure of Tibeto-Burman populations reveals sex-biased admixture in southern Tibeto-Burmans[J]. *Am J Hum Genet*, 2004, 74(5):856-865.
 - [39] SU B, XIAO C, DEKA R, et al. Y chromosome haplotypes reveal prehistorical migrations to the Himalayas[J]. *Hum Genet*, 2000, 107(6):582-590.
 - [40] ZHAO M, KONG Q P, WANG H W, et al. Mitochondrial genome evidence reveals successful Late Paleolithic settlement on the Tibetan Plateau[J]. *Proc Natl Acad Sci USA*, 2009, 106(50):21230-21235.
 - [41] QIN Z, YANG Y, KANG L, et al. A mitochondrial revelation of early human migrations to the Tibetan Plateau before and after the last glacial maximum[J]. *Am J Phys Anthropol*, 2010, 143(4):555-569.
 - [42] BEALL C M, CAVALLERI G L, DENG L, et al. Natural selection on *EPAS1* (*HIF2α*) associated with low hemoglobin concentration in Tibetan highlanders[J]. *Proc Natl Acad Sci USA*, 2010, 107(25):11459-11464.
 - [43] YANG J, JIN Z B, CHEN J, et al. Genetic signatures of high-altitude adaptation in Tibetans[J]. *Proc Natl Acad Sci USA*, 2017, 114(16):4189-4194.
 - [44] JIANG L, PENG J, HUANG M, et al. Differentiation analysis for estimating individual ancestry from the Tibetan Plateau by an archaic altitude adaptation

- EPAS1* haplotype among East Asian populations[J]. *Int J Legal Med*, 2018, 132(6):1527-1535.
- [45] KONG Q P, YAO Y G, LIU M, et al. Mitochondrial DNA sequence polymorphisms of five ethnic populations from northern China[J]. *Hum Genet*, 2003, 113(5):391-405.
- [46] 陆艳. 中国西部人群的遗传混合[D]. 上海:复旦大学, 2011.
- [47] NAKAGOME S, SATO T, ISHIDA H, et al. Model-based verification of hypotheses on the origin of modern Japanese revisited by Bayesian inference based on genome-wide SNP data[J]. *Mol Biol Evol*, 2015, 32(6):1533-1543.
- [48] JINAM T A, KANZAWA-KIRIYAMA H, INOUE I, et al. Unique characteristics of the Ainu population in northern Japan[J]. *J Hum Genet*, 2015, 60(10):565-571.
- [49] WANG Y, LU D, CHUNG Y J, et al. Genetic structure, divergence and admixture of Han Chinese, Japanese and Korean populations[J]. *Hereditas*, 2018, 155:19.
- [50] LI C X, PAKSTIS A J, JIANG L, et al. A panel of 74 AISNPs: Improved ancestry inference within eastern Asia[J]. *Forensic Sci Int Genet*, 2016, 23:101-110.
- [51] QIN P, LI Z, JIN W, et al. A panel of ancestry informative markers to estimate and correct potential effects of population stratification in Han Chinese[J]. *Eur J Hum Genet*, 2014, 22(2):248-253.
- [52] YUASA I, UMETSU K, ADACHI N, et al. Investigation of Japanese-specific alleles: Most are of Jomon lineage[J]. *Leg Med (Tokyo)*, 2015, 17(1):52-55.
- [53] WRIGHT S. The genetical structure of populations[J]. *Ann Eugen*, 1951, 15(4):323-354.
- [54] ROSENBERG N A, LI L M, WARD R, et al. Informativeness of genetic markers for inference of ancestry[J]. *Am J Hum Genet*, 2003, 73(6):1402-1422.
- [55] PRITCHARD J K, STEPHENS M, DONNELLY P. Inference of population structure using multilocus genotype data[J]. *Genetics*, 2000, 155(2):945-959.
- [56] 刘京, 李盛, 江丽, 等. 对于未知来源个体进行族群推断的自动分析系统[J]. *生命科学研究*, 2018, 22(1):3-7, 41.
- [57] RAJEEVAN H, SOUNDARARAJAN U, PAKSTIS A J, et al. Introducing the forensic research/reference on genetics knowledge base, FROG-kb[J]. *Investig Genet*, 2012, 3(1):18.
- [58] SUN Q, JIANG L, ZHANG G, et al. Twenty-seven continental ancestry-informative SNP analysis of bone remains to resolve a forensic case[J]. *Forensic Sci Res*, 2017. doi:10.1080/20961790.2017.1306431.
- (收稿日期:2019-04-04)
(本文编辑:边英男)

(上接第518页)

- [57] JACOBS L C, HAMER M A, GUNN D A, et al. A genome-wide association study identifies the skin color genes *IRF4*, *MC1R*, *ASIP*, and *BNC2* influencing facial pigmented spots[J]. *J Invest Dermatol*, 2015, 135(7):1735-1742.
- [58] HERNANDO B, IBANEZ M V, DESERIO-CUESTA J A, et al. Genetic determinants of freckle occurrence in the Spanish population: Towards ephelides prediction from human DNA samples[J]. *Forensic Sci Int Genet*, 2018, 33:38-47.
- [59] ENDO C, JOHNSON T A, MORINO R, et al. Genome-wide association study in Japanese females identifies fifteen novel skin-related trait associations[J]. *Sci Rep*, 2018, 8(1):8974.
- [60] PENG Q, LI J, TAN J, et al. *EDARV370A* associated facial characteristics in Uyghur population revealing further pleiotropic effects[J]. *Hum Genet*, 2016, 135(1):99-108.
- [61] SHAFFER J R, LI J, LEE M K, et al. Multiethnic GWAS reveals polygenic architecture of earlobe attachment[J]. *Am J Hum Genet*, 2017, 101(6):913-924.
- [62] KEATING B, BANSAL A T, WALSH S, et al. First all-in-one diagnostic tool for DNA intelligence: Genome-wide inference of biogeographic ancestry, appearance, relatedness, and sex with the Identitas v1 Forensic Chip[J]. *Int J Legal Med*, 2013, 127(3):559-572.
- (收稿日期:2019-05-22)
(本文编辑:张素华)