

DOI: 10.3724/SP.J.1005.2012.01351

中国人类遗传多样性研究进展

杨昭庆, 褚嘉祐

中国医学科学院&北京协和医学院医学生物学研究所, 医学遗传学研究室, 昆明 650118

摘要: 人类遗传多样性表现为世界各种族、民族和个体间存在的基因组差异, 是探讨人类进化与迁徙、环境与遗传背景相互作用、疾病与健康影响因素的主要资源和工具。中国具有世界 1/5 的人口, 有 56 个民族和丰富的遗传多样性资源。经过几十年的努力, 中国人类遗传多样性研究已积累了丰富的资料, 部分成果已经达到国际先进水平。文章重点论述了近年来形态学标记、生化及免疫学标记、DNA 遗传标记在我国人类遗传多样性研究中的应用, 线粒体 DNA、Y 染色体 DNA 和 HLA 标记在中国不同民族源流和相互关系、东亚现代人起源和迁移等研究中的应用, 以及我国在中华民族遗传资源保存和利用、疾病易感基因和环境适应相关基因的鉴定及全基因组关联分析和第二代测序技术的应用、中国人基因组结构研究等方面取得的重要成果和进展。

关键词: 人类遗传多样性; 遗传标记; 基因组; 易感基因; 中华民族

The research progress of human genetic diversity in China

YANG Zhao-Qing, CHU Jia-You

Department of Medical Genetics, Institute of Medical Biology, Chinese Academy of Medical Sciences & Peking Union Medical College, Kunming 650118, China

Abstract: Human genetic diversity refers to genomic variation among races, ethnic groups, isolated populations and individuals worldwide, and is one major resource and tool on discovering human evolution and migration, interaction between genetic background and environment, and factors associated with human diseases and health. China has abundant and valuable resource of human genetic diversity due to 56 ethnic groups and a large population accounting for one fifth of the total population in the world. After decades of efforts, a large number of research data on human genetic diversity have been accumulated in China, and some of outcomes reach advanced international level. This review mainly focuses on the recent progress and outcomes achieved in applying genetic markers including morphological markers, biochemical and immunological markers and DNA markers in research of genetic diversity, and the application of mitochondrial DNA, Y chromosomal DNA, HLA and others in research of the origin and relationship of Chinese ethnic groups, and the origin and migrations of modern East Asian populations. This review also summarizes the advances in the research fields of preservation and utilization of Chinese genetic resource, identification of genes associated with disease selective and adaptive for natural pressure, application of whole genome association study and next generation sequencing, and Chinese human genome as well.

收稿日期: 2012-10-26; 修回日期: 2012-11-02

基金项目: 国家高技术研究发展计划(863)项目(编号: 2012AA021802)和国家自然科学基金面上项目(编号: 30971578)资助

作者简介: 杨昭庆, 博士, 研究员, 专业方向: 医学遗传学。Tel: 871-8334872; E-mail: zyang@imbcams.com.cn

通讯作者: 褚嘉祐, 博士, 研究员, 专业方向: 医学遗传学。Tel: 871-8334872; E-mail: chujy@imbcams.com.cn

Keywords: human genetic diversity; genetic marker; genome; susceptibility gene; Chinese ethnic groups

遗传多样性是生物多样性的的重要组成部分和核心问题,广义上指生物携带遗传信息的总和,狭义上指生物种间或种内分子、细胞和个体水平的遗传变异度,是生物物种在进化过程积累起来的遗传变异。人类遗传多样性体现在不同人类种族、民族群体及个体之间基因组和表型的差异,是研究人类起源、进化和迁移、民族源流及不同民族群体相互关系、环境与遗传背景相互作用的最好依据。更为重要的是,遗传多样性是肿瘤、心血管疾病、自身免疫系统疾病等多种疾病遗传易感性的生物基础,同时也是在疾病治疗中不同群体或个体对药物及生物制剂有不同反应的主要原因。遗传多样性的研究在疾病基因鉴定、基因诊断和基因治疗、药物筛选等疾病防治中具有重大的实用意义。

中国不仅具有世界五分之一的人口,而且 56 个民族在遗传表型及一些遗传病的发病率上有较大的不同。中国自主的人类遗传多样性研究主要开始于 20 世纪 70~80 年代,迄今经历几十年的努力、积累和发展,由最初开展的中国人体质参数测量和皮纹学研究,到如今大规模单核苷酸多态性研究、全基因组关联分析及新一代测序技术全基因组分析的应用,在研究技术手段、理论深度及广度上都获得了长足进展和显著成就。通过中国人类基因组计划等重大项目的开展,我国在中国不同民族群体遗传资源保存及多样性分析、东亚现代人起源等方面为国际性人类遗传多样性数据库提供了中国人的资料,使中国的人类遗传多样性研究加入到国际性的人类多样性合作及成果共享体系中。目前中国在人类遗传多样性研究方面与国际水平差距不大,众多国内学者运用遗传多样性研究手段,在中华民族基因组保存、多基因疾病鉴定、东亚现代人的起源与迁徙、人群环境适应与选择的分子机制、亚洲人全基因组图谱绘制和结构分析等方面作出了独创性的成果,开辟了一些具有国际先进水平的研究领域和方向,取得了世界瞩目的成果。本文就中国人类遗传多样性研究的发展和取得的主要成果进行综合介绍,尤其是不同遗传标记的应用及中国人类遗传多样性的

应用进展。

1 中国人类遗传多样性研究中遗传标记的应用

遗传标记从广义上指可以反映受基因控制的遗传多态性的生物特征,并可用有效方法进行测量;从狭义上是指反映生物个体或种群间基因组差异的 DNA 序列。随着遗传学理论及技术的发展,人类遗传标记的发展主要经历了 4 个发展阶段,依次分为:(1)形态学标记。指具有质量性状或数量性状的、可遗传的生物外部特征,如肤色和发色等体质形态学标记;(2)细胞学标记:主要指染色体数目、形态及带型;(3)生化及免疫学标记:主要指同工酶、免疫蛋白和红、白细胞抗原等;(4)DNA 标记:是对基因核苷酸序列变异的直接反映。比起前 3 种遗传标记,DNA 标记具有数量丰富、不受环境影响、遗传稳定、多态性高等特点,是目前的主流遗传标记。在中国人类遗传多样性研究中,细胞遗传学标记因存在数目有限、分析方法相对繁琐等局限性而使用较其他类标记少。

1.1 体质形态学标记

体质遗传学标记指可遗传的体质形态学特征,是我国早期用于人类学和群体遗传学的主要形态学遗传标记。主要包括“描述性”的体质形态标记如皮肤和头发颜色、眼及面部特征等,以及“测量性”形态标记如身高、四肢长度、颅骨、皮纹等。我国的体质遗传学研究主要从 20 世纪 70 年代开始,对我国多个地区及不同民族的体质特征进行了测量和分析,积累了 20 多个省市和地区,包括 40 多个民族群体的资料,对探讨各民族起源、进化及人口迁徙具有重要作用^[1~5]。张振标^[1]从 1979 年起调查和分析了 16 个省、市、自治区的 25 个少数民族头面部的体质特征,用多元分析方法,将我国现代中国人的体质特征明显分为南部和北部两大地区类型,并认为这两大类型体征的形成起始于我国文化发达的新石器时代。胡兴宇等^[2]通过对公开发表的我国 33 个少

数民族活体测量资料的 9 项指标进行聚类分析, 结果支持现代中国人体质特征的南北两种类型的分类, 还提出有第三种类型即藏彝走廊类型的存在, 并且生活在藏彝走廊内的藏缅语族各族具有共同的体质特征, 提示他们与古羌人有共同的族源。皮纹属多基因遗传, 在不同民族和种族间差异显著, 是人类群体遗传多样性的经典指标。对皮纹参数的研究也表明, 中国各民族的皮纹具有各自特点, 群体成员间又有相对一致的体质特征, 民族之间及其与汉族之间的差异一般小于他们与白种人之间的差异, 多个民族可能有共同的起源并有民族的互相融合^[3,4]。张海国根据 11 项肤纹参数对 28 个民族的 52 个群体进行分析表明: 中国民族人群分南方群, 北方群和混合群, 南北两大聚类群以长江或北纬 $30^{\circ} \sim 33^{\circ}$ 为界带; 混合群中有各包括了南、北方民族的小聚类群, 如彝族与羌族、藏族在系统树上聚成一群。杜若甫研究组调查了中国 11 个少数民族群体 5 013 人的指纹参数, 分析表明指纹花样和指纹脊线数的分布在群体水平上存在人种、民族、性别以及指别间的差异, 各民族的指纹具有各自特点又具有蒙古人种的一般特性^[3]。除了不断补充我国各民族的皮纹数据外, 有学者认为皮纹特征的改变是先天性发育异常的一个明显标志, 并探讨了精神分裂症患者皮纹嵴线数的特征以寻找早期筛查的指标^[5]。身高、颅骨和皮纹等体质学标记研究所得的结果互相支持, 均显示中国的汉族和各少数民族在体质遗传标记上具既有异质性又有相似性, 中华民族各群体间存在基因的互相交流和融合。

1.2 生化及免疫学标记

在 DNA 序列变异检测技术出现之前, 多应用生化及免疫学的表现型与基因型间的对应关系来了解等位基因的分布及其频率, 这些质量性状表型的遗传性状标记也被称为经典遗传标记。我国从 1980 年起开始发表较大规模的基因频率群体调查, 调查较多的经典遗传标记有红细胞血型如 ABO、MNS、Rh 和 P 血型, 红细胞酶如腺苷酸激酶 1(AK1)、磷酸葡萄糖变位酶(PGM1)及 6-磷酸葡萄糖脱氢酶(G6PD)等, 血清血浆蛋白如 α 球蛋白、Gm 和 Km, 以及形态生理性状如眈眈的干湿和色盲等。迄今, 我国学者已对 56 个民族群体、74 个基因座位的基因频率进行了调查, 以 1 个人群在 1 个基因座位的基因频率为 1 套, 全

国已累积的基因频率数据达 2 000 多套^[6-9]。其中被调查民族数较多的标记是 ABO 血型系统及免疫球蛋白 Gm 和 Km 的基因频率。利用这些数据, 我国学者研究了各基因座位的各等位基因频率在我国的地理分布特点, 以了解我国各地区人群的遗传结构特点, 探讨中华各民族的起源和进化。陈稚勇等^[7]分析了 291 份发表的中国人 ABO 型分布调查资料的分析, 发现中国人可分为北方和南方两大类, 分别以较高的 B 基因频率和较高的 O 基因频率为其特征。刘祖洞等^[10]对中国人 14 个群体中 Gm 和 Km 因子的分析结果提示, 我国古代华夏族同许多其他民族存在有基因交流, 汉族群体是一个具有南北群体间多样性的、不均一的群体, 古代不同群体的迁移和混杂作用受地理位置的限制。赵桐茂等^[9]对中国 24 个民族、74 个群体的人免疫球蛋白同种异型的分布进行了调查, 研究结果支持中华民族起源于古代南北两大群体的假说, 这两个群体大致以北纬 30° 为界, 分别居栖在黄河和长江流域, 群体之间的遗传距离和他们所居住的地理位置之间, 大致上呈平行关系。结果还显示汉族具有显著的异质性, 蒙古人种明显地分为南、北两大类型, 中国西北地区的少数民族中混有高加索人种血缘, 不同种族间不但 Gm 单体型频率有明显差异, 而且还存在一些具有“人种特异性”的 Gm 单体型。杜若甫等用 ABO、MNSs、G6PD、免疫球蛋白 Gm 与 Km 因子等 38 个基因座位的频率对中国 30 个省区的汉族和 37 个少数民族进行了遗传距离计算和聚类分析, 结果表明中国汉族与少数民族都分为南方蒙古人种与北方蒙古人种两大类型, 并以长江为界, 确凿地从遗传学角度证明各地汉族与当地少数民族血缘相近, 汉族与少数民族血缘有互相融合^[8]。总的来说, 经典遗传标记表明汉族群体具有异质性, 支持汉族群体粗分为南北类型的学说。

1.3 DNA 标记

形态标记、细胞学标记和生化免疫标记都是受基因表达和环境影响, 是对基因的间接反映, 而 DNA 标记是则可直接反映核酸水平的遗传变异。由于早期应用的 DNA 标记多以电泳图谱的形式表现, 因此又称 DNA 指纹图谱。根据 DNA 标记的发展, 通常将限制性片段长度多态性(Restriction fragment length polymorphisms, RFLP)称为第一代 DNA 标记, 以核酸序列变异导致限制性内切酶切片长度变化

及以 Southern 杂交为特点;将以短串联重复序列(Short tandem repeat, STR)为代表的基于 PCR 技术的可变串联重复序列标记称为第二代 DNA 标记,包括数目可变串联重复序列(Variable number of tandem repeats, VNTR)、短串联重复序列(Sort tandem repeat, STR)和特定序列位点(Sequencetagged site, STS)等;将基于单核苷酸多态性(Single nucleotide polymorphism, SNP)的标记称为第三代 DNA 分子标记。近年来,拷贝数变异(Copy number variation, CNV)作为一类新型的 DNA 标记逐渐得以推广应用。

1.3.1 限制性片段长度多态性(RFLP)标记

我国从 20 世纪 90 年代起开始应用 RFLP 标记进行民族群体遗传学和疾病易感基因多态分析等。1994 年陈仁彪研究组用对我国汉族和彝族 rRNA 基因 3' 端不转录间隔区进行基因多态分析,提示上海汉族与云南彝族有着广泛的族源关系^[11]。除了常染色体外,线粒体和 Y 染色体的 RFLP 多态性研究为比较我国各民族之间遗传相似性及探讨本民族的遗传和变异机制提供了有力的证据。由于 RFLP 多态性标记的等位基因具有共显性的特点,并且结果稳定可靠,因而在我国不同民族和不同人群中心血管疾病、肿瘤、感染性疾病的易感基因位点筛选中得到广泛应用。如顾东风等^[12]通过分析 PON 基因簇潜在功能位点的 RFLP 多态性与冠心病的关联分析,提示 *PON1* 基因可能是中国北方汉族人群中的冠心病易感基因。同时 RFLP 也在我国人群的脊髓性肌萎缩症、软骨发育不全、聋病基因、血友病等疾病的基因诊断中有较多应用。但由于 RFLP 较难以自动化,在全基因组上的分布密度较低,多态性频率也较低,在人群多样性研究中的应用受到一定限制。

1.3.2 微卫星(STR)标记

STR 又称微卫星 DNA 标记(Microsatellite DNA),由于具有在基因组中密度高、分布广、多态信息含量(PIC)较高、易扩增的特点,是遗传作图、基因定位、群体分化及历史事件估计等人类遗传多样性研究及个体鉴定和亲子鉴定等法医学研究中最常用的常染色体 DNA 遗传标记。由于毛细管电泳技术的发展,显著提高了检测的自动化程度,从 20 世纪 90 年代起我国人群的遗传多样性研究就广泛应用了微卫星标记,并在检测方法的建立和优化、遗传数

据统计分析、民族和种族特异性微卫星标记发现、中华民族源流等方面作出了独特的贡献,并在国内外知名杂志上发表了一批具有国际影响力的研究成果。况少青等建立了一种以 Taq Start™ 抗体为 Taq 酶保护剂的多重 PCR 方法,在国内率先提供了一种简便、快速的半自动基因组扫描方法^[13],袁文涛等应用多重 PCR 方法对 72 名汉族人群中的 139 个微卫星位点进行分析,结果发现其中 88% 位点的杂合度 > 0.6, 94% 位点的 $PIC > 0.5$, 标记具有高度的可提供信息。通过与法国人类多态研究中心(CEPH)提供的这些位点的杂合度进行比较,证明在不同种族、人群之间微卫星位点的杂合度和等位片段频率有差异^[14]。褚嘉祐等^[15]分析了中国 28 个民族群体样本的 30 个常染色体微卫星数据,并结合国外 15 个群体的数据,构建了 32 个东亚人群的系统树,结果表明中国南北人群存在差异,民族群体间存在基因交流。中国现代人来自非洲的现代直立人,经过中亚后由东南亚进入中国并取代了中国原有的人类群体。东亚人群的基因池有多个来源,这是国际上首次报道东亚(包括中国)现代人的非洲起源的研究。在此之后,褚嘉祐的研究组还用 30 个微卫星位点对我国的白族、纳西族、土族等 6 个民族群体进行了基因扫描和分型,用 MICA 基因微卫星位点对中国的拉祜族、畲族、藏族等 13 个民族群体进行了基因分型,群体系统树构建和遗传距离计算,为揭示中国多个民族的源流、相互关系及遗传多样性提供了大量的参考数据^[16, 17]。目前对中国不同民族群体仍然没有一个利用微卫星位点较为完整而系统的研究,仍然需要增加所选取的民族数量、每个民族中的样本量及微卫星标记的数量及分布,以较为全面完整地反应中国不同人群的遗传结构及其差异。

微卫星标记还在我国的法医学研究中得以广泛应用,由于不同种族和不同民族间的差异,使国外的法医学鉴定研究成果不能直接应用于中国人群。我国已经收集和分析了多个地区和不同民族群体中,包含性染色体、线粒体及常染色体 DNA 上的部分法医学常见微卫星位点的等位基因频率、杂合度、个体识别率和非父排除率等群体遗传学指标,为法医学的个人识别、亲权鉴定及司法精神病鉴定等提供参考依据。侯一平等^[6]共完成了 45 个常染色体 STR 基因座位在自然群体中的等位基因频率,还有学者

调查了中国南方汉族人群中 9 个 STR 基因座位的多态性^[18], 这些数据成为法医 DNA 分型的概率计算依据。此外, 我国还研制出 X-STR 基因座的多重 PCR 等多种试剂盒^[19], 为法医学应用提供了更多有效手段。

1.3.3 单核苷酸多态性(SNP)标记

单核苷酸多态性(SNP)在全基因组范围内广泛分布, 在人群中的分布频率大于 1%。SNP 是一种具有高度稳定性的遗传学标记, 代表了不同个体之间最大的遗传差异, 是研究基因多态性、人类进化、复杂疾病病因及个性化医疗等的主要工具。约从 2000 年起, 我国开始了 SNP 在人类遗传多样性研究中的应用, 在我国民族群体多样性研究中较多应用 Y 染色体或线粒体的 SNP 多态性调查分析, 而常染色体 SNP 的研究较多集中在对某些疾病相关基因的多态性分析。目前已积累了布依族、白族、佤族、朝鲜族等 20 多个民族群体中 Y 染色体 SNP、糖尿病和高血压等多种疾病相关基因 SNP 的多态性资料。由于单个 SNP 所提供的遗传信息有限, 在分析民族间的遗传关系或疾病易感基因时时, 常采用多个 SNP 位点共同分析的策略。单可人等调查了贵州布依族、仡佬族、毛南族和壮族 Y-SNP 的多态性及单倍型并分析其父系遗传结构^[20]。有学者对我国汉族、蒙古族和壮族的 13 个 SNPs 位点多态性进行了调查, 结果显示多个 SNP 的多态性具有明显的民族差异, 地理同处于北方的汉族与蒙古族亲缘关系较近^[21]。从 2007 年起, 随着基因芯片等高通量 SNP 检测技术的不断发展, 高密度 SNP 扫描乃至全基因组关联分析(GWAS)成为我国分析群体遗传关系和寻找疾病易感基因的最主要方法之一。我国学者将佤族、维吾尔族等多个民族群体样本参与到与人类基因组泛亚 SNP 共同体(The HUGO Pan-Asian SNP Consortium)的合作研究中, 分析了 54 794 个常染色体 SNP, 对亚洲人群进行遗传多样性图谱绘制, 结果表明亚洲各群体的遗传源流呈现出与各群体的语系和地理关系的一致性, 东南亚群体是东亚群体的主要起源^[22]。几乎我国所有的 GWAS 都在汉族人群中进行, 但是越来越多的证据表明汉族群体间也存在遗传异质性, 对汉族群体的选样不当可能会导致 GWAS 分析结果的误差。金力等^[23]采用 160K 的 SNP 芯片对分

布于中国 26 个地区的 1 700 例汉族样品进行了分析, 结果发现汉族群体可粗略划分成北方、中部和南方 3 个遗传亚群, 尽管在这 3 个亚群间的遗传差异是非常小的, 但即便是在样品量适度的情况下也已经足够引起 GWAS 的“病例-对照”分析假阳性结果的产生。研究结果还发现 *FADS2* 和 *HCP5* 基因多态性在南、北汉族人群中具有显著差异, 在 3 个汉族亚群之间, 涉及心脏动脉病变通路的基因多态性也有显著差异。我国学者与新加坡研究人员共同对中国 10 个省份的 6 000 多份汉族样品进行了 35 万个 SNP 的分析, 首次绘制出了中国汉族人人群的遗传进化史图谱, 结果表明汉族群体存在“南-北”群体遗传结构, 这种遗传差异与中国的地理有密切的关联, 显示了不同地区汉族群体中存在的细微差别和多样性, 结果同样提示在 GWAS 研究设计中必须考虑汉族群体样品的来源^[24]。

1.3.4 拷贝数变异(CNV)标记

拷贝数变异(CNV)主要是指基因组上发生涉及大片段 DNA 序列的变异, 如亚显微重复和微缺失从而导致 DNA 片段拷贝数量的改变。CNV 是一种在人类基因组中广泛分布的遗传多态性, 近年来逐渐广泛用于研究世界范围内多个群体之间的相互遗传关系^[25]。CNV 与出生缺陷等重大疾病密切相关, 全球有多个研究中心正在开展 CNV 研究项目, 规模较大的有人类基因组结构变异研究组及千人基因组计划, 后者已发现约 2 万个 CNV 位点^[26]。我国与其他国家同步, 也是近 3、4 年开始应用 CNV 标记进行群体遗传多样性及疾病关联分析。2011 年金力的研究组利用基因芯片技术, 对代表中国不同语系的汉族、藏族、侗族、瑶族、黎族和维吾尔族群体样本中检测了全基因组范围的 CNV, 构建了首张包括中国汉族和少数民族在内的 CNV 图谱, 并与其他国家和地区人群的数据进行比较分析。研究发现在少数民族群体中约有 35% 的 CNV 区域与汉族人群不同, 提示中国人群的遗传多样性大部分存在于少数民族人群中。同时发现一些明显有民族群体差异的 CNV, 如频率在侗族和壮族中分别高达 44.4% 和 50% 的一个 DNA 片段缺失在汉族中却未被检出, 群体特异性 CNV 可能与人群对其特定生存环境的长期适应有关。此外还发现在中国人中 SNP 与标签 CNV(Tagged-

CNV)的关联性比在欧洲人群中的低^[27]。有报道 *CYP2D6* 基因的CNV分布特点在我国东部不同地区汉族人群中存在差异^[28]。我国学者还发现在染色体 16p12.3 区段的CNV在中国人与欧洲人之间有显著差异并与肥胖表型相关^[29]。CNV对基因功能和表型的影响效应比SNP要大得多,但是其频率较低,往往需要参考数据库以进行全基因组分析来寻找与特定疾病或表型相关的CNV,这些均提示全面研究中国不同民族人群的CNV具有必要性。

2 应用 HLA、线粒体和 Y 染色体 DNA 标记研究中国人群基因组多态性

2.1 HLA 标记的应用

HLA是人类的主要组织相容复合体,是目前已知的多态性最为丰富的基因系统,其等位基因分布、单倍型组成及频率在不同人种和地区有明显不同。自 20 世纪 80 年代开始,我国应用国际上统一的血清学方法进行了两次大规模的群体调查,第一次是 1984~1996 年进行的HLA-A、-B、-Cw、-DR和DQ分型,调查涉及 20 个实验室和我国 10 个城市的汉族,2 个实验室分别调查了维吾尔族和朝鲜族^[30]。第二次是 1991 年对我国主要少数民族和南北汉族样本HLA-I、II 类抗原多态性的分析^[31]。结果提示中国汉族属蒙古人种,中华民族包含南北两大群体,南北人群中有高度显著连锁不平衡的双位点单倍型。苗、布依与南方汉族聚类,蒙古、满、回、藏与北方汉族集群。高加索人种起源的HLA抗原(A3、B8 等)频率由西向东又由北向南递减。而东南亚蒙古人种起源的HLA抗原(B46)基因频率由南向北递减,反应了中华民族祖先自远古史前时期以来的不断迁移和民族相互融合的过程。吴国光等采用DNA测序等以DNA为基础的HLA基因分型技术鉴定了 6 965 份无关汉族的HLA-A、B、DRB1 座位上的等位基因,结果发现北方汉族人群带有较高频率的A1、A3、A30 等抗原,而南方汉族以高频率的A11、B46、B58 等抗原为特征,南北汉族人群HLA基因频率分布存在明显差异,支持了之前血清学方法调查所得的结论,而且提示在计算基因频率时,应该按南北群体分开统计^[32]。随着DNA分析技术的应用与发展,我国学者建立和应用了序列特异性寡

核苷酸探针反向斑点杂交(ASO)PCR-序列特异性引物(PCR-SSP)和液相芯片等检测技术,获得了傣族、基诺族、彝族、壮族、怒族、土族、裕固族、维吾尔族等 10 多个中国不同民族的HLA等位基因型数据^[33-35]。近年来,我国每年有近千篇关于HLA的研论文或报告发表,为进一步揭示各民族间的源流和关系,分析感染性疾病、过敏性疾病、肿瘤易感性、自身免疫性疾病易感性、免疫耐受性、器官移植排斥、疫苗接种弱反应性等的分子机制研究提供了更为详细的遗传数据。

2.2 线粒体 DNA 标记的应用

线粒体DNA(mtDNA)由于具有母系遗传、缺乏重组和高进化速率的遗传特性而被广泛应用于研究现代人类起源、人群迁徙和多样性。我国在上世纪 70 年代起就有学者陆续对我国民族人群的mtDNA多态性进行了研究,获得了云南傣族、白族等、贵州的侗族、仡佬族、土家族、彝族、新疆的维吾尔族、以及广西壮族等 20 多个民族的mtDNA多态性数据,发现了一些新的基因型,并对各民族的源流和迁移进行了分析^[36-38]。1978 年,贺林等^[39]用 11 种限制性内切酶对中国汉族的mtDNA进行了RFLP的初步研究,随后俞民澍等在我国汉族、维吾尔族、哈萨克族和回族的mtDNA限制酶切片多态类型研究中发现了一批新的类型,其中包括一些只存在于某一种族和民族中的突变类型,表明mtDNA限制酶切片多态性类型的分布在不同民族和种族中有显著差别。到 2000 年前后,我国对中国人群mtDNA遗传变异的研究出现了一个高峰时期,比较有影响力的工作主要由中国科学院昆明动物研究所、中国科学院遗传与发育研究所、中国医学科学院医学生物学研究所及复旦大学等多家单位完成,在这个时期获得了我国西南、西北等地多个少数民族群体、以及我国多个地区汉族群体的单倍型多样度和核苷酸多样性数据^[6]。姚永刚等^[40]对 372 份中国汉族的线粒体DNA第一高变区(HSVI)序列进行检测并与之前所报道的中国其他民族群体数据相对比分析,初步获得了我国汉族地理人群中的mtDNA单倍型和我国人群中各单倍型类群的初步地理分布特点。从多个地区民族群体的数据来看,各民族间在等位基因及单倍型分布频率上存在有较明显差异,同一民族不

同群体间的单倍群频率分布存在差异,但差异不很大,一般小于不同族源民族间的差异。各地汉族人群的单倍型多样性普遍较高且相近,而少数民族群体中遗传多样性整体上变化较大。从聚类结果上来看,汉族地理群体没有完全聚在一起,各地的汉族群体和当地的少数民族群体并没有表现出因地域分布接近而呈现聚类关系相近的格局,用语系和地理学对民族的分类与用mtDNA分类所得结果不完全相同,这些结果与用其他遗传标记所得的研究结果也有所不同。提示在进行中华民族群体间遗传亲缘关系形成研究中,要充分考虑样本的选择和遗传标记选择的局限性,简单地以地理标志来划分或合并民族群体的做法难以反映民族人群遗传结构的全貌。

在对各民族源流的分析研究中,有学者通过对我国及东南亚地区苗瑶语系、壮侗语系及欧亚语系民族群体HVS-1区的多态性分析,推断了各人群特异性单倍群形成的时间并认为中国南部很早就有现代人居住,随后才发生了人群间民族群体的分化及民族间的基因混合^[41]。对海南黎族mtDNA HVS-I和-II的多态性分析提示黎族在海南岛的居住史可追溯到7千多年前^[42]。在西藏人中发现了新的M16单倍群,并确定藏族人群中有大量的母系遗传,其成分可追溯到石器时代并在中世纪由中国北部迁入西藏平原^[43]。线粒体多态性为我国民族群体的考古学和历史学提供了生物学的证据,但中国人群存在的一些未定的mtDNA世系的系统发育关系如何,族源、民族文化和群体的母系遗传结构之间是否有关联性等问题依然需要深入探讨^[6]。

2.3 Y 染色体 DNA 标记的应用

Y 染色体标记代表了父系遗传,与线粒体 DNA 标记同为单倍体存在,由于不受重组的干扰,更能忠实地记录人类的进化事件并能从较低水平的人群多样性中有效地捕捉到人群分化、迁徙的信息。从20世纪90年代起我国学者就开始了较为系统性的民族群体中Y染色体多态性分析,主要对我国人群的源流和相互关系,以及东亚现代人的起源进行了探讨。杜若甫研究组调查了中国10个人群中Y染色体,发现YAP+的基因频率北方群体高于南方群体,彝族和藏族一样,有较高的YAP+等位基因,均源于古代我国西北地区的羌人^[44]。复旦大学、中国科学

院昆明动物研究所、国家人类基因组南方中心、中国医学科学院医学生物学研究所等单位的学者联合研究了19个SNP组成的Y染色体单倍型在全国22个省市汉族人群中的分布,结果表明南方人群的多态性明显高于北方人群,提示现代人类自南方进入中国,随后由南向北迁移。同时估算了现代人类进入中国的时间大致在1.8~6万年前^[45]。傅松滨研究组对赫哲、满、锡伯等8个我国北方民族群体的17个Y染色体STR位点多态性进行分析,结果显示这些群体均具有较小的族内差异而有较大的民族间差异,但与日本、台湾等地人群相比,这些民族在遗传距离上仍然比较相近^[46]。

除了研究我国民族人群间的相互关系及历史迁移外,结合我国积累的多民族遗传数据,我国学者还利用Y染色体多态性探讨了东亚现代人的起源与史前迁徙。宿兵、金力和肖春杰的研究组开展合作,通过Y染色体遗传标记单倍型组O3-M122,对40个东亚南北方代表群体,共2322个男性个体进行了系统的SNP和STR比较分析,结果表明非洲起源的东亚现代人最早到达东亚的南部,东南亚南方群体是祖先群体,北方群体约在2.5~3万年以前从南方迁到北方^[47]。该研究结果是迄今为止最系统的有关东亚现代人最早史前迁徙的遗传学证据。该合作研究组还发现D-M174单倍型(YAP单倍型)起源于东亚的南部,而且其由南向北的迁徙大约发生在距今6万年前,可能是发生在东亚大陆最早的大规模现代人群的迁徙,现在的东亚人群中只有少数群体保存了那次史前迁徙的遗传印记^[48]。另一项合作研究通过高密度Y-染色体单倍群及STR基因分型,检测了遍布中国境内116个地区和朝鲜地区的共3826名男性的DNA样本,构建了精细的Y染色体单倍群进化地理分布图谱并测定了有关单倍型的STR年代^[49]。这项研究结果清楚回答了前期研究中难以解释的东亚人群中存在的中亚及欧洲遗传成分问题,证明了东亚人群史前迁徙的“北线”假说。

3 中国人类遗传多样性研究的应用

3.1 中国人类遗传多样性资源的保存

中国人口数量约占世界的22%,具有56个民族及众多的遗传隔离群体,这些群体在民族源流、遗

传标记、遗传病种类及疾病的易感性方面呈现出较大的差异,是研究人类群体生物学特征及疾病基因的重要研究材料。随着社会经济发展、民族间的交流及通婚日趋增加,原有的相对隔离状况逐渐打破,各民族独特的基因组面临着消失的危险,建立中国人不同民族永生细胞库,对长期保存不同民族基因组以供永久性研究具有重要意义。自 1993 年起,作为中国人类基因组计划的重要组成部分,由中国医学科学院医学生物学研究所、中科院遗传与发育研究所和哈尔滨医科大学共同合作,采用 EB 病毒转化外周血 B 淋巴细胞的方法建立中华民族永生细胞库,同时对病毒转化细胞、支原体检测及遗传稳定性分析等多项技术进行了质量控制优化^[50],并探讨和建立从事我国人类遗传资源采集和利用所需遵循的“知情同意”等多项伦理原则。到 2008 年已建立和储存了 47 个民族 70 个群体(含民族支系)的 3 982 株永生细胞株,保存了 7 210 份各民族 DNA 样本^[51],是目前规模最大的较为完整的中国国家级各民族永生细胞库。其收集保存的民族群体数量还在逐年增加之中,样本收集范围也扩大到隔离人群或疾病相关的家系永生细胞建立。永生细胞库不仅可以提供取之不竭的 DNA、RNA 和蛋白样品,而且由于是活细胞,还可提供从基因型到表型的一系列基因功能研究及药物筛选测试。同时,此项工作作为国际人类基因组计划的一部分,已向法国人类基因组计划国际中心(CEPH)递交了我国 15 个民族的 149 株细胞系以作为国际人类基因组多样性细胞库的组成部分。已有超过 150 篇 SCI 论文采用了这些细胞样品,我国也因此获得了世界范围内超过 1 000 份不同民族个体遗传资源的共享权,促进了我国与国际人类遗传资源和遗传多样性研究的交流和成果共享。由于受项目初期科研条件的限制,有的民族仅保存了少量个体的永生细胞株,还需要及时补充更多的个体细胞株以建立更完整标准参考样品库。

3.2 疾病易感基因和环境适应相关基因的鉴定

我国从 20 世纪 80 年代初就开始了疾病相关基因的多态性研究,早期的研究多采用血清或生化方法对血红蛋白病、HLA 类型和一些代谢酶进行了调查。90 年代后,由于我国人类基因组计划的启动以及 DNA 测序等分析方法的迅速发展,以发表基因

多态性最为集中的《中华医学遗传杂志》为例,共发表了疾病相关基因的多态性论文 333 篇,涉及 65 种疾病及 110 多个基因,疾病和医学相关位点多态性研究涉及的群体数达 199 个^[6]。其中涉及最多的疾病依次是冠心病、肿瘤、糖尿病及高血压等复杂性多基因疾病,研究的基因则以 HLA 基因多态最多。较为显著的变化是,在疾病或表型性状相关基因的研究领域,我国学者在近年来作出了突出的、具有中国特色和国际领先水平的研究成果,尤其是在应用 GWAS 分析疾病相关易感基因方面取得了较多显著的进展和成果。2009 年张学军的研究组采用 GWAS 研究方法,通过 15 000 例中国汉族和维吾尔族银屑病病例和健康对照的大样本研究,发现了与银屑病发病机制密切相关的 LCE 基因变异^[52],这是我国国内首个完成的重大疾病 GWAS 研究,随后又进一步发现了 6 个银屑病易感基因,其中 ERAP1 和 ZNF816A 与中国人中早发型病例密切相关,提出了银屑病发病的遗传异质性^[53]。该研究合作组继而进行了目前世界上红斑狼疮全基因组关联分析研究中样本量最大的研究项目,发现了 5 个红斑狼疮易感基因,并确定了 4 个新的易感位点,首次通过遗传学研究证明了红斑狼疮发病机制中的遗传危险因素在不同人种间具有相同和不同的易感基因^[54]。此外,应用 GWAS 研究策略,2011 年宋怀东、黄薇、陈赛娟等鉴定出 RNASE2 和 CHR9A9 两个新的甲状腺机能亢进(Graves 病)易感位点,并发现与治疗相关的两种不同发病机制^[55]。贺林研究组鉴别出了位于 8p12 和 1q24.2 的 2 个与精神分裂症相关的风险位点^[56]。2012 年 Yu 等^[57]鉴定出中国汉族人群 IgA 肾病的两个易感基因, TNFSF13 和 DEFA,并证实 MHC 区域的 rs660895 多态位点与 IgA 肾病临床分型的相关性;林昕东研究组确认了 9 个中国人群食管鳞状细胞癌的易感基因位点,并分析饮酒状况下 ADH 和 ALDH2 基因多态性与患病风险的相互关系^[58];我国学者还用 GWAS 策略鉴定出了白癜风、强直性脊柱炎、肺癌、原发性闭角型青光眼、前列腺癌、非贲门胃癌等疾病的易感基因。除了 GWAS,我国科学家还利用第二代测序技术扫描疾病易感基因突变。陈竺和陈赛娟的研究小组利用外显子测序(Exome sequencing),发现了急性单核细胞白血病(AML-M5)中 DNA 甲基转移酶基因 DNMT3A 突变与高发病率及

预后不良密切相关,同时还发现 13.6%的AML-M4白血病存在DNMT3A Arg882 突变^[59]。我国学者还证实了PRRT2 基因突变有可能导致阵发性运动源性运动障碍的发生^[60]。近年来有等多项GWAS和外显子测序的研究成果在*Nature Genetics*等国际知名期刊上发表,这些研究成果表明中国科学家自主的高通量全基因组研究已跻身国际先进水平。

在人群环境适应(Adaptation)与遗传多样性相互作用分子机制的研究领域,近年来我国学者也取得了丰硕的、具有自主性的研究成果。如通过分析碘缺乏的陕西秦巴山区和上海地区汉族的ApoE等位基因频率,发现epsilon2 和epsilon4 的高等位基因频率可能是源于对营养不良和碘缺乏环境因素的适应性,从而导致碘缺乏区有较高的智力缺陷发生率^[61]。通过高密度全基因组SNP连锁分析,发现中国大陆人群的身体质量指数(BMI)存在有在进化上具有保守性的遗传因素,然而同时也有多个位点的多态性差异反映了群体特异性环境适应所导致的“基因-环境”相互作用遗传效应^[62]。在对藏族高海拔适应性的遗传学机制研究中,我国华大基因研究院通过外显子测序研究策略,对 50 例藏族个体的样品进行分析,确定了EPAS1 等基因的SNP多态性与高原世居藏族人群高原适应性密切相关,并发现了一些具有民族特异性的SNP位点。该研究工作发表在*Science*杂志^[63]。随后,金力和宿兵采用GWAS研究策略也证实了EPAS1 和EGLN1 是藏族高海拔适应的关键基因^[64,65]。这些成果对预测、预防与治疗高原缺氧性疾病具有重大意义。宿兵等^[66]在东亚人群中,P53 基因第72 位的氨基酸变异与纬度密切相关,即纬度越高,其所在纬度人群中精氨酸等位基因的频率越高。该研究首次揭示,像P53 和MDM2 这样感受环境压力的关键基因,在东亚人群中受到了很强的自然选择的影响。宿兵等^[67]还在人群对酒精的适应性进化研究中发现,乙醇脱氢酶第 47 位氨基酸多态性(ADH1B 47His)在人群中的分布和起源时间与中国南方新石器早期水稻的驯化及其扩散的模式非常吻合。该研究是目前报道的仅有的几个人类对食物变化分子适应的例子,为了解新石器时代以来人群遗传结构的变迁以及自然选择的分子机制提供了新的证据。

3.3 中国人类基因组计划中的遗传多样性研究

早在 1991 年,以吴旻院士为首的一些科学家就提出,人类基因组不能没有中国各民族基因组的数据,1994 年国家自然科学基金重大项目“中华民族若干位点结构的研究”立项启动了中国人类基因组研究和国际合作,中国不同人群遗传多样性研究是其中的重点内容。1999 年,我国参与了国际人类基因组计划,获得了人类 3 号染色体短臂上一个约 30 Mb区域的测序任务,片段约占人类基因组的 1%,该计划在 2001 年顺利完成。2008 年我国深圳华大基因研究院采用第二代测序技术,自主完成了第一例中国人标准基因组序列图(炎黄一号基因组),这也是首个亚洲个人基因组图谱,发现了 3 百万个 SNPs,其中 13.6%尚未收录在NCBI的dbSNP数据库中。并通过与现有HapMap数据的对比分析,推测中国人的遗传有效种群大小(Effective population size)约为 3300~5700,还发现了具有种群特异性的基因组结构变异和功能基因^[68]。在随后进行的泛基因组(pan-genome)研究中,我国学者通过对炎黄一号进一步的深度测序和拼接,并将其与非洲人基因组比较,发现了约 5 Mb的新序列。此外还发现,人类基因组中除了原先公认的单核苷酸多态性、插入删除多态性和结构性变异以外,还存在着种群特异甚至个体独有的DNA序列和功能基因^[69]。这些研究均从全基因组层面展现了中国不同民族群体间存在的遗传多样性,表明自主构建中国民族群体基因图谱和推动个体化医学研究的必要性,同时也展现了我国科学家在人类遗传多样性研究中的独特贡献和所具有的世界领先地位。

4 结语与展望

遗传多样性的研究无论是对生物多样性的保护还是对生物资源的可持续利用均有重要的意义。不难看出,人类遗传多样性的研究对推动我国的民族资源保存、卫生健康保障水平和经济发展所具有的重要性。中国的人类遗传多样性研究经过多年努力,从初期对人群的描述性观察,到近来全基因组多位点乃至全基因组序列的扫描分析,在近 20 多年间已成为发展最快、取得成果最为显著的研究领域之一。其中在中华民族永生细胞库建立、中华民族的

南方和北方群体遗传结构、东亚人起源、疾病易感基因的全基因组关联分析及高通量序列测定和分析、个体基因组图谱绘制等方面积累较多,做出了中国独特的贡献,成为缩小发达国家与发展国家研究差距的典型范例,而所有这些成果都与充分利用我国独有的民族群体多样性资源这一优势密切相关。尽管与国外相比,中国的人类遗传学研究仍然有一定的差距,但随着我国经济和科技的发展,人类遗传多样性研究领域多获得的基金支持和技术手段也不断增多,加之我国丰富的遗传资源,我国对国际的贡献将会更大。

中国的人类基因组学及疾病研究中仍然有许多问题需要解决,一些领域仍然可以继续发挥我们的特长和优势。中华不同民族遗传资源库仍然需要补充部分民族群体的样本,整个资源库也需要从民族群体种类及每个群体的数量上继续完善,并且加强疾病家系和隔离群体家系的永生细胞库建立,以避免重复性和破坏性采样。从管理上也需完善细胞鉴定等生物指标和数据库管理体系,从而能更好与国际标准细胞库接轨。已有研究表明,人永生B-淋巴细胞株可提供丰富的转录本表达信息及序列信息^[70],可以应用永生淋巴细胞株筛选预测肺癌肿瘤化疗药物反应性的分子标记^[71]。如何拓宽中华民族永生细胞库在药物基因组学、个体化医疗等领域的应用是下一步的努力方向。GWAS和第二代测序技术提供了全基因组水平的分析能力,但目前的研究还只集中在汉族群体,已有的研究结果已经证实了我国各民族群体间存在的遗传差异及其在医学研究中的价值,所以我国仍然需要绘制各个民族群体的全基因组代表数据。发达国家已经发现了一些肿瘤、心血管疾病、自身免疫系统等多种多基因疾病的相关基因位点,但研究结果也同时表明其中的许多位点在不同种族或民族群体中可呈现出完全不同的关联性,这些位点与疾病的关联性仍有必要在我国不同群体中深入研究,国外发现的易感基因位点也需在我国不同发病人群中进行多中心、大样本的反复验证。基因变异导致的基因表达水平变化具有显著的民族群体差异^[72],我国已获得了许多DNA水平的遗传多样性资料,但仍然缺乏各民族群体中数量性状座位(QTL)的表达信息谱,还缺乏系统性、群体性的从DNA到RNA乃至蛋白质多样性的互相验证。对这些

问题的深入了解和探索,将更有助于推动我国人类遗传多样性资源的保护和利用,推动对疾病病因、诊断技术和治疗策略的研究,同时也推动我国的生物技术利用和发展。

参考文献(References):

- [1] 张振标. 现代中国人体质特征及其类型的分析. 人类学学报, 1988, 7(4): 314–322. DOI
- [2] 胡兴宇, 汪澜, 黎彦才. 中国 33 个少数民族体质特征类型的研究. 解剖学杂志, 1993, 16(1): 71–73. DOI
- [3] 李实喆, 毛钟荣, 徐玖瑾, 崔梅影, 王永发, 陈良忠, 袁义达, 李绍武, 杜若甫. 中国十一个少数民族的皮纹研究 I. 指纹. 人类学学报, 1984, 3(1): 37–42. DOI
- [4] 张海国, 丁明, 焦云萍, 汪宪平, 颜中, 金刚, 孟秀莲, 白崇显, 陆振虞, 陈仁彪. 中国人肤纹研究——III. 中国 52 个民族的肤纹聚类. 遗传学报, 1998, 25(5): 381–391. DOI
- [5] 党洁, 霍正浩, 彭亮, 陈银涛, 焦海燕, 陆宏, 钟慧军, 赵巍. 精神分裂症患者皮纹a-b嵴线数波动性不对称的研究. 人类学学报, 2007, 26(1): 64–69. DOI
- [6] 金力, 褚嘉祐. 中华民族遗传多样性研究. 上海: 上海科学技术出版社, 2006. DOI
- [7] 陈稚勇, 赵桐茂, 张工梁. 中国人ABO血型分布. 遗传, 1982, 4(2): 4–7. DOI
- [8] 杜若甫, 肖春杰, Cavalli-Sforza LL. 用 38 个基因座的基因频率计算中国人群间遗传距离. 中国科学C辑, 1998, 28(1): 83–89. DOI
- [9] 赵桐茂, 张工梁, 朱永明, 郑素琴, 顾文娟, 陈琦, 章霞, 刘鼎元. 中国人免疫球蛋白同种异型的研究: 中华民族起源的一个假说. 遗传学报, 1991, 18(2): 97–108. DOI
- [10] 潘犁, 刘祖洞. 中国人十四个群体中Gm和Km因子的分布. 复旦学报 (自然科学版), 1988, 27(4): 382–388. DOI
- [11] 袁飒英, 蒋伟宏, 陆嫣, 费虹明, 陈仁彪. 我国汉族和彝族rRNA基因多态性的研究. 遗传学报, 1994, 21(3): 173–178. DOI
- [12] 王晓玲, 范中杰, 黄建凤, 宿少勇, 赵建功, 顾东风. PON基因簇潜在功能多态位点与冠心病的关联研究. 遗传学报, 2005, 32(7): 675–681. DOI
- [13] 况少青, 王建民, 黄薇, 张宇舟, 陆林, 陈竺, 金力. 应用多重PCR进行微卫星荧光标记-半自动基因组扫描. 中华医学遗传学杂志, 1998, 15(2): 104–107. DOI
- [14] 袁文涛, 徐红岩, 赵进英, 丁伟, 江宏铨, 顾鸣敏, 薛京伦, 陈家伦, 方福德, 陈竺, 金力, 黄薇. 微卫星位点在基因组扫描中的信息表现. 中华医学遗传学杂志, 2000, 17(2): 65–71. DOI
- [15] 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. Genetic relationship of populations in China. *Proc Natl Acad Sci USA*, 1998, 95(20): 11763–11768. DOI
- [16] 班贵宏, 褚嘉祐, 许绍斌, 杨昭庆, 钱亚屏, 俞建昆, 纳剑波, 刘晓娟, 张思仲. MICA基因微卫星多态性在中国 13 个群体中的分布. *遗传学报*, 2001, 28(12): 1085–1092. DOI
- [17] 俞建昆, 褚嘉祐, 钱亚屏, 孙浩, 史磊, 史荔, 初正韬, 杨昭庆, 林克勤, 陶玉芬, 黄薇, 金力. 应用 30 个常染色体STR位点研究中国 6 个民族群体的遗传关系. *遗传学报*, 2001, 28(8): 699–706. DOI
- [18] 童大跃, 孙宏钰, 伍新尧, 陆惠玲, 李建金, 陈丽娴. 中国南方汉族人群 9 个STR基因座多态性分析. *中山大学学报 (医学科学版)*, 2009, 30(4): 400–403. DOI
- [19] 李莉, 赵书民, 张素华, 李成涛, 柳燕, 林源, 刘俊宏. X染色体上 16 个STR基因座的分型检测和多态性分析. *法医学杂志*, 2012, 28(1): 36–40, 43. DOI
- [20] 刘烜, 单可人, 齐晓岚, 何燕, 赵艳, 吴昌学, 李毅, 褚迅, 任锡麟. 贵州布依族、仡佬族、仫佬族、毛南族、壮族Y-SNP的初步研究. *遗传*, 2006, 28(11): 1350–1354. DOI
- [21] 王瑞恒, 刘利民, 赵金玲. 我国 3 个民族 13 个SNPs位点多态性及遗传学关系的比较. *遗传*, 2009, 31(3): 273–279. DOI
- [22] Abdulla MA, Ahmed I, Assawamakin A, Bhak J, Brahmachari SK, Calacal GC, Chaurasia A, Chen CH, Chen J, Chen YT, Chu J, Cutiongco-de la Paz EM, De Ungria MCA, Delfin FC, Edo J, Fuchareon S, Ghang H, Gojobori T, Han J, Ho SF, Hoh BP, Huang W, Inoko H, Jha P, Jinam TA, Jin L, Jung J, Kangwanpong D, Kampuansai J, Kennedy GC, Khurana P, Kim HL, Kim K, Kim S, Kim WY, Kimm K, Kimura R, Koike T, Kulawonganuchai S, Kumar V, Lai PS, Lee JY, Lee S, Liu ET, Majumder PP, Mandapati KK, Marzuki S, Mitchell W, Mukerji M, Naritomi K, Ngamphiw C, Niikawa N, Nishida N, Oh B, Oh S, Ohashi J, Oka A, Ong R, Padilla CD, Palittapongarnpim P, Perdigon HB, Phipps ME, Png E, Sakaki Y, Salvador JM, Sandraling Y, Scaria V, Seielstad M, Sidek MR, Sinha A, Srikummool M, Sudoyo H, Sugano S, Suryadi H, Suzuki Y, Tabbada KA, Tan A, Tokunaga K, Tongsima S, Villamor LP, Wang E, Wang Y, Wang HF, Wu JY, Xiao HS, Xu SH, Yang JO, Shugart YY, Yoo HS, Yuan WT, Zhao GP, Zilfalil BA. Mapping human genetic diversity in Asia. *Science*, 2009, 326(5959): 1541–1545. DOI
- [23] Xu SH, Yin XY, Li SL, Jin WF, Lou HY, Yang L, Gong XH, Wang HY, Shen YP, Pan XD, He YG, Yang YJ, Wang Y, Fu WQ, An Y, Wang JC, Tan JZ, Qian J, Chen XL, Zhang X, Sun YF, Wu BL, Jin L. Genomic dissection of population substructure of Han Chinese and its implication in association studies. *Am J Hum Genet*, 2009, 85(6): 762–774. DOI
- [24] Chen JM, Zheng HF, Bei JX, Sun LD, Jia WH, Li T, Zhang FR, Seielstad M, Zeng YX, Zhang XJ, Liu JJ. Genetic structure of the Han Chinese population revealed by genome-wide SNP variation. *Am J Hum Genet*, 2009, 85(6): 775–785. DOI
- [25] Jakobsson M, Scholz SW, Scheet P, Gibbs JR, VanLiere JM, Fung HC, Szpiech ZA, Degnan JH, Wang K, Guerreiro R, Bras JM, Schymick JC, Hernandez DG, Traynor BJ, Simon-Sanchez J, Matarin M, Britton A, van de Leemput J, Rafferty I, Bucan M, Cann HM, Hardy JA, Rosenberg NA, Singleton AB. Genotype, haplotype and copy-number variation in worldwide human populations. *Nature*, 2008, 451(7181): 998–1003. DOI
- [26] 杜仁骞, 金力, 张锋. 基因组拷贝数变异及其突变机理与人类疾病. *遗传*, 2011, 33(8): 857–869. DOI
- [27] Lou HY, Li SL, Yang YJ, Kang LL, Zhang X, Jin WF, Wu BL, Jin L, Xu SH. A map of copy number variations in Chinese populations. *PLoS One*, 2011, 6(11): e27341. DOI
- [28] Sheng HH, Zeng AP, Zhu WX, Zhu RF, Li HM, Zhu ZD, Qin Y, Jin W, Liu Y, Du YL, Sun J, Xiao HS. Allelic distributions of CYP2D6 gene copy number variation in the Eastern Han Chinese population. *Acta Pharmacol Sin.*, 2007, 28(2): 279–286. DOI
- [29] Yang TL, Guo Y, Li SM, Li SK, Tian Q, Liu YJ, Deng HW. Ethnic differentiation of copy number variation on chromosome 16p12.3 for association with obesity phenotypes in European and Chinese populations. *Int J Obes (Lond)*, 2012, doi: 10.1038/ijo.2012.31. DOI
- [30] 陈仁彪, 赵桐茂, 叶根耀, 裴冀, 史秉璋, 苏炳华. 中国大陆HLA多态性联合报告. *现代免疫学*, 1987, 7(6): 321–324. DOI
- [31] 陈仁彪, 叶根跃, 庚镇城, 王增慧, 孔繁华, 田丁, 包丕云, 刘若英, 刘杰, 宋芳吉, 范丽安, 张工梁, 郭实士, 徐林敏, 徐星培, 程定珍, 赵修竹. 我国大陆主要少数民族HLA多态性聚类分析和频率分布对中华民族起源的启示. *遗传学报*, 1993, 20(5): 389–393. DOI
- [32] 吴国光, 邓志辉, 高素青, 程良红, 金士正, 周丹, 李桢, 邹红岩, 张旋, 魏天莉, 程曦, 王大明. 6965 名汉族骨髓供者HLA多态性分析. *中华血液学杂志*, 2004, 25(8): 473–477. DOI
- [33] Yao Y, Shi L, Tao Y, Kulski JK, Lin K, Huang X, Xiang H, Chu J. Distinct HLA allele and haplotype distributions in four ethnic groups of China. *Tissue Antigens*, 2012, 80(5): 452–461. DOI

- [34] Shi L, Ogata S, Yu JK, Ohashi J, Yu L, Shi L, Sun H, Lin KQ, Huang XQ, Matsushita M, Horai S, Muramatsu M, Chu JY, Tokunaga K. Distribution of HLA alleles and haplotypes in Jinuo and Wa populations in Southwest China. *Hum Immunol*, 2008, 69(1): 58–65. [DOI](#)
- [35] Shen CM, Zhu BF, Deng YJ, Ye SH, Yan JW, Yang G, Wang HD, Qin HX, Huang QZ, Zhang JJ. Allele polymorphism and haplotype diversity of HLA-A, -B and -DRB1 loci in sequence-based typing for Chinese Uyghur ethnic group. *PLoS One*, 2010, 5(11): e13458. [DOI](#)
- [36] 高路, 董永利, 郝肇菁, 王欧, 杨智丽, 苏艳华, 郑冰蓉, 咎瑞光, 肖春杰. 云南 16 个少数民族群体的线粒体 DNA 多态性研究. *遗传学报*, 2005, 32(2): 118–123. [DOI](#)
- [37] 李彬彬, 钟复光, 易红生, 王先然, 李良芳, 王丽兰, 吴立甫. 贵州四个民族人群线粒体 DNA 编码区的多态性. *中国比较医学杂志*, 2010, 20(3): 8–11. [DOI](#)
- [38] 木耶塞尔·伊斯马依力, 古丽娜·艾山, 马合木提·哈力克. 新疆 8 个地域维吾尔族群体线粒体 DNA 9 bp 缺失频率与 Y-染色体 DYS287 位点多态性研究. *中山大学学报 (自然科学版)*, 2011, 50(4): 100–105. [DOI](#)
- [39] 贺林, 严明, 王世浚. 中国汉族人线粒体 DNA RFLP 的初步研究. *科学通报*, 1987, (23): 1826–1828. [DOI](#)
- [40] Yao YG, Nie L, Harpending H, Fu YX, Yuan ZG, Zhang YP. Genetic relationship of Chinese ethnic populations revealed by mtDNA sequence diversity. *Am J Phys Anthropol*, 2002, 118(1): 63–76. [DOI](#)
- [41] Li H, Cai XY, Winograd-Cort ER, Wen B, Cheng X, Qin ZD, Liu WH, Liu YF, Pan SL, Qian J, Tan CC, Jin L. Mitochondrial DNA diversity and population differentiation in southern East Asia. *Am J Phys Anthropol*, 2007, 134(4): 481–488. [DOI](#)
- [42] Peng MS, He JD, Liu HX, Zhang YP. Tracing the legacy of the early Hainan Islanders—a perspective from mitochondrial DNA. *BMC Evol Biol*, 2011, 11(1): 46. [DOI](#)
- [43] 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](#)
- [44] 许丽萍, 徐玖瑾, 朱苏玲, 包维东, 杜若甫. 中国 10 个人群中 Y 染色体 Alu 序列的多态分布. *科学通报*, 1998, 43(8): 843–846. [DOI](#)
- [45] 柯越海, 宿兵, 肖君华, 陈华, 黄薇, 陈竺, 褚嘉佑, 谈家桢, 金力, 卢大儒. Y 染色体单倍型在中国汉族人群中的多态性分布与中国人群的起源及迁移. *中国科学 (C 辑: 生命科学)*, 2000, 30(6): 614–620. [DOI](#)
- [46] Zheng LH, Sun HM, Wang JW, Li SL, Bai J, Jin Y, Yu Y, Chen F, Jin L, Fu SB. Y chromosomal STR polymorphism in northern Chinese populations. *Biol Res*, 2009, 42(4): 497–504. [DOI](#)
- [47] Shi H, Dong YL, Wen B, Xiao CJ, Underhill PA, Shen PD, Chakraborty R, Jin L, Su B. Y-chromosome evidence of southern origin of the East Asian-specific haplogroup O3-M122. *Am J Hum Genet*, 2005, 77(3): 408–419. [DOI](#)
- [48] Shi H, Zhong H, Peng Y, Dong YL, Qi XB, Zhang F, Liu LF, Tan SJ, Ma RZ, Xiao CJ, Wells RS, Jin L, Su B. Y chromosome evidence of earliest modern human settlement in East Asia and multiple origins of Tibetan and Japanese populations. *BMC Biol*, 2008, 6(1): 45. [DOI](#)
- [49] Zhong H, Shi H, Qi XB, Duan ZY, Tan PP, Jin L, Su B, Ma RZ. Extended Y chromosome investigation suggests postglacial migrations of modern humans into East Asia via the northern route. *Mol Biol Evol*, 2011, 28(1): 717–727. [DOI](#)
- [50] 黄小琴, 褚嘉佑, 林克勤, 陶玉芬, 俞建昆, 史磊, 史荔, 于亮, 石铁流. 建立不同民族永生细胞株质量控制方法的探讨. *中国优生与遗传杂志*, 2005, 13(9): 10–13. [DOI](#)
- [51] 褚嘉佑, 徐玖瑾, 傅松滨, 林克勤, 朱苏玲, 黄小琴, 陶玉芬, 薛雅丽, 孙艳阳, 杨昭庆, 钱亚屏, 李璞. 中华民族永生细胞库的建立. *国际遗传学杂志*, 2008, 31(4): 241–247, 301. [DOI](#)
- [52] Zhang XJ, Huang W, Yang S, Sun LD, Zhang FY, Zhu QX, Zhang FR, Zhang C, Du WH, Pu XM, Li H, Xiao FL, Wang ZX, Cui Y, Hao F, Zheng J, Yang XQ, Cheng H, He CD, Liu XM, Xu LM, Zheng HF, Zhang SM, Zhang JZ, Wang HY, Cheng YL, Ji BH, Fang QY, Li YZ, Zhou FS, Han JW, Quan C, Chen B, Liu JL, Lin D, Fan L, Zhang AP, Liu SX, Yang CJ, Wang PG, Zhou WM, Lin GS, Wu WD, Fan X, Gao M, Yang BQ, Lu WS, Zhang Z, Zhu KJ, Shen SK, Li M, Zhang XY, Cao TT, Ren W, Zhang X, He J, Tang XF, Lu S, Yang JQ, Zhang L, Wang DN, Yuan F, Yin XY, Huang HJ, Wang HF, Lin XY, Liu JJ. Psoriasis genome-wide association study identifies susceptibility variants within LCE gene cluster at 1q21. *Nat Genet*, 2009, 41(2): 205–210. [DOI](#)
- [53] Sun LD, Cheng H, Wang ZX, Zhang AP, Wang PG, Xu JH, Zhu QX, Zhou HS, Ellinghaus E, Zhang FR, Pu XM, Yang XQ, Zhang JZ, Xu AE, Wu RN, Xu LM, Peng L, Helms CA, Ren YQ, Zhang C, Zhang SM, Nair RP, Wang HY, Lin GS, Stuart PE, Fan X, Chen G, Tejasvi T, Li P, Zhu J, Li ZM, Ge HM, Weichenthal M, Ye WZ, Shen SK, Yang BQ, Sun YY, Li SS, Lin Y, Jiang JH, Li CT, Chen RX, Cheng J, Jiang X, Zhang P, Song WM, Tang J, Zhang HQ,

- Sun L, Cui J, Zhang LJ, Tang B, Huang F, Qin Q, Pei XP, Zhou AM, Shao LM, Liu JL, Zhang FY, Du WD, Franke A, Bowcock AM, Elder JT, Liu JJ, Yang S, Zhang XJ. Association analyses identify six new psoriasis susceptibility loci in the Chinese population. *Nat Genet*, 2010, 42(11): 1005–1009. [DOI](#)
- [54] Han JW, Zheng HF, Cui Y, Sun LD, Ye DQ, Hu Z, Xu JH, Cai ZM, Huang W, Zhao GP, Xie HF, Fang H, Lu QJ, Li XP, Pan YF, Deng DQ, Zeng FQ, Ye ZZ, Zhang XY, Wang QW, Hao F, Ma L, Zuo XB, Zhou FS, Du WH, Cheng YL, Yang JQ, Shen SK, Li J, Sheng YJ, Zuo XX, Zhu WF, Gao F, Zhang PL, Guo Q, Li B, Gao M, Xiao FL, Quan C, Zhang C, Zhang Z, Zhu KJ, Li Y, Hu DY, Lu WS, Huang JL, Liu SX, Li H, Ren YQ, Wang ZX, Yang CJ, Wang PG, Zhou WM, Lv YM, Zhang AP, Zhang SQ, Lin D, Low HQ, Shen M, Zhai ZF, Wang Y, Zhang FY, Yang S, Liu JJ, Zhang XJ. Genome-wide association study in a Chinese Han population identifies nine new susceptibility loci for systemic lupus erythematosus. *Nat Genet*, 2009, 41(11): 1234–1237. [DOI](#)
- [55] Chu X, Pan CM, Zhao SX, Liang J, Gao GQ, Zhang XM, Yuan GY, Li CG, Xue LQ, Shen M, Liu W, Xie F, Yang SY, Wang HF, Shi JY, Sun WW, Du WH, Zuo CL, Shi JX, Liu BL, Guo CC, Zhan M, Gu ZH, Zhang XN, Sun F, Wang ZQ, Song ZY, Zou CY, Sun WH, Guo T, Cao HM, Ma JH, Han B, Li P, Jiang H, Huang QH, Liang L, Liu LB, Chen G, Su Q, Peng YD, Zhao JJ, Ning G, Chen Z, Chen JL, Chen SJ, Huang W, Song HD. A genome-wide association study identifies two new risk loci for Graves' disease. *Nat Genet*, 2011, 43(9): 897–901. [DOI](#)
- [56] Shi YY, Li ZQ, Xu Q, Wang T, Li T, Shen JW, Zhang FY, Chen JH, Zhou GQ, Ji WD, Li BJ, Xu YF, Liu DT, Wang P, Yang P, Liu BX, Sun WS, Wan CL, Qin SY, He G, Steinberg S, Cichon S, Werge T, Sigurdsson E, Tosato S, Palotie A, Nöthen MM, Rietschel M, Ophoff RA, Collier DA, Rujescu D, Clair DS, Stefansson H, Stefansson K, Ji J, Wang QZ, Li WJ, Zheng LQ, Zhang HR, Feng GY, He L. Common variants on 8p12 and 1q24.2 confer risk of schizophrenia. *Nat Genet*, 2011, 43(12): 1224–1227. [DOI](#)
- [57] Yu XQ, Li M, Zhang HQ, Low HQ, Wei X, Wang JQ, Sun LD, Sim KS, Li Y, Foo JN, Wang W, Li ZJ, Yin XY, Tang XQ, Fan L, Chen J, Li RS, Wan JX, Liu ZS, Lou TQ, Zhu L, Huang XJ, Zhang XJ, Liu ZH, Liu JJ. A genome-wide association study in Han Chinese identifies multiple susceptibility loci for IgA nephropathy. *Nat Genet*, 2012, 44(2): 178–182. [DOI](#)
- [58] Wu C, Kraft P, Zhai K, Chang J, Wang ZM, Li Y, Hu Z, He ZB, Jia WH, Abnet CC, Liang LM, Hu N, Miao XP, Zhou YF, Liu ZH, Zhan QM, Liu Y, Qiao Y, Jin GF, Guo CH, Lu CH, Yang HJ, Fu JH, Yu DK, Freedman ND, Ding T, Tan W, Goldstein AM, Wu TC, Shen HB, Ke Y, Zeng YX, Chanock SJ, Taylor PR, Lin DX. Genome-wide association analyses of esophageal squamous cell carcinoma in Chinese identify multiple susceptibility loci and gene-environment interactions. *Nat Genet*, 2012, 44(10): 1090–1097. [DOI](#)
- [59] Yan XJ, Xu J, Gu ZH, Pan CM, Lu G, Shen Y, Shi JY, Zhu YM, Tang L, Zhang XW, Liang WX, Mi JQ, Song HD, Li KQ, Chen Z, Chen SJ. Exome sequencing identifies somatic mutations of DNA methyltransferase gene DNMT3A in acute monocytic leukemia. *Nat Genet*, 2011, 43(4): 309–315. [DOI](#)
- [60] Chen WJ, Lin Y, Xiong ZQ, Wei W, Ni W, Tan GH, Guo SL, He J, Chen YF, Zhang QJ, Li HF, Murong SX, Xu J, Wang N, Wu ZY. Exome sequencing identifies truncating mutations in PRRT2 that cause paroxysmal kinesigenic dyskinesia. *Nat Genet*, 2011, 43(12): 1252–1255. [DOI](#)
- [61] Gao JJ, Zhang FC, Guo TW, Gao XC, Duan SW, Wang HY, Zheng ZJ, Huang TZ, Feng GY, St Clair D, He L. Distribution of apolipoprotein E allele frequencies of the Han Chinese in an iodine-deficient mountainous area. *Ann Hum Biol*, 2004, 31(5): 578–585. [DOI](#)
- [62] Zhang DF, Pang ZC, Li SX, Thomassen M, Wang SJ, Jiang WJ, Hjelmberg J v B, Kruse TA, Kyvik KO, Christensen K, Tan QH. High-resolution genome-wide linkage mapping identifies susceptibility loci for BMI in the Chinese population. *Obesity (Silver Spring)*, 2012, 20(4): 830–833. [DOI](#)
- [63] Yi X, Liang Y, Huerta-Sanchez E, Jin X, Cuo ZX, Pool JE, Xu X, Jiang H, Vinckenbosch N, Korneliussen TS, Zheng H, Liu T, He W, Li K, Luo R, Nie X, Wu H, Zhao M, Cao H, Zou J, Shan Y, Li S, Yang Q, Asan, Ni P, Tian G, Xu J, Liu X, Jiang T, Wu R, Zhou G, Tang M, Qin J, Wang T, Feng S, Li G, Huasang, Luosang J, Wang W, Chen F, Wang Y, Zheng X, Li Z, Bianba Z, Yang G, Wang X, Tang S, Gao G, Chen Y, Luo Z, Gusang L, Cao Z, Zhang Q, Ouyang W, Ren X, Liang H, Huang Y, Li J, Bolund L, Kristiansen K, Li Y, Zhang Y, Zhang X, Li R, Yang H, Nielsen R, Wang J. Sequencing of 50 human exomes reveals adaptation to high altitude. *Science*, 2010, 329(5987): 75–78. [DOI](#)
- [64] Peng Y, Yang ZH, Zhang H, Cui CY, Qi XB, Luo X, 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](#)
- [65] Xu SH, Li SL, Yang YJ, Tan JZ, Lou HY, Jin WF, Yang L, Pan X, Wang JC, Shen YP, Wu BL, Wang HY, Jin L. A genome-wide search for signals of high-altitude adaptation in Tibetans. *Mol Biol Evol*, 2011, 28(2): 1003–1011. [DOI](#)
- [66] Shi H, Tan SJ, Zhong H, Hu WW, Levine A, Xiao CJ, Peng Y, Qi XB, Shou WH, Ma RL, Li Y, Su B, Lu X. Winter temperature and UV are tightly linked to genetic changes in the p53 tumor suppressor pathway in Eastern Asia. *Am J Hum Genet*, 2009, 84(4): 534–541. [DOI](#)
- [67] Peng Y, Shi H, Qi XB, Xiao CJ, Zhong H, Ma RL, Su B. The ADH1B Arg47His polymorphism in east Asian populations and expansion of rice domestication in history. *BMC Evol Biol*, 2010, 10: 15, doi: 10.1186/1471-2148-10-15. [DOI](#)
- [68] Wang J, Wang W, Li RQ, Li YR, Tian G, Goodman L, Fan W, Zhang JQ, Li J, Zhang JB, Guo YR, Feng BX, Li H, Lu Y, Fang XD, Liang HQ, Du ZL, Li D, Zhao YQ, Hu YJ, Yang ZZ, Zheng HC, Hellmann I, Inouye M, Pool J, Yi X, Zhao J, Duan JJ, Zhou Y, Qin JJ, Ma LJ, Li GQ, Yang ZT, Zhang GJ, Yang B, Yu C, Liang F, Li WJ, Li SC, Ni PX, Ruan J, Li QB, Zhu HM, Liu DY, Lu ZK, Li N, Guo GW, Zhang JG, Ye J, Fang L, Hao Q, Chen Q, Liang Y, Su YY, San A, Ping C, Yang S, Chen F, Li L, Zhou K, Zheng HK, Ren YY, Yang L, Gao Y, Yang G, Li Z, Feng XL, Kristiansen K, Wong GKS, Nielsen R, Durbin R, Bolund L, Zhang XQ, Li SG, Yang HM, Wang J. The diploid genome sequence of an Asian individual. *Nature*, 2008, 456(7218): 60–65. [DOI](#)
- [69] Li RQ, Li YR, Zheng HC, Luo RB, Zhu HM, Li QB, Qian WB, Ren YY, Tian G, Li JX, Zhou GY, Zhu X, Wu HD, Qin JJ, Jin X, Li DF, Cao HZ, Hu XD, Blanche H, Cann H, Zhang XQ, Li SG, Bolund L, Kristiansen K, Yang HM, Wang J, Wang J. Building the sequence map of the human pan-genome. *Nat Biotechnol*, 2010, 28(1): 57–63. [DOI](#)
- [70] Toung JM, Morley M, Li M, Cheung VG. RNA-sequence analysis of human B-cells. *Genome Res*, 2011, 21(6): 991–998. [DOI](#)
- [71] Tan XL, Moyer AM, Fridley BL, Schaid DJ, Niu N, Batzler AJ, Jenkins GD, Abo RP, Li L, Cunningham JM, Sun Z, Yang P, Wang L. Genetic variation predicting cis-platin cytotoxicity associated with overall survival in lung cancer patients receiving platinum-based chemotherapy. *Clin Cancer Res*, 2011, 17(17): 5801–5811. [DOI](#)
- [72] Spielman RS, Bastone LA, Burdick JT, Morley M, Ewens WJ, Cheung VG. Common genetic variants account for differences in gene expression among ethnic groups. *Nat Genet*, 2007, 39(2): 226–231. [DOI](#)

•综合信息•

生物衰老：研究方法实验方案

生命科学实验指南系列

书号：978-7-03-034363-5/Q · 2919 作者：王钊 等译 定价：98

内容简介：本书详细介绍生物学衰老的前沿研究技术，旨在为读者提供诸多实用、高效且适宜的研究方案，用于研究调控生物学衰老过程的机制及对其的干预方法。书中主要内容分为三部分：研究衰老过程的基本方法，涵盖基础领域的细胞培养与分选、衰老标志物检测、端粒长度分析和 DNA 甲基转移酶研究等；干预衰老过程的技术，包括热量限制及其模拟、营养基因组学、外源端粒酶使用和药物作用评估等；分析生物衰老的实验方案，如基因的高通量筛选、DNA 微阵列技术、消减杂交技术和二维凝胶蛋白质组学等。利用本书综述的各项强大工具，读者可根据研究需要将不同技术紧密结合，从而更好地推进生物学衰老这一新兴综合学科的不断发展。

科学出版社 科学销售中心

联系人：周文宇 电话：010-64022646, 010-64017321 E-mail:zhouwenyu@mail.sciencep.com