

古老人群 mtDNA 研究中存在的问题和对策

姚永刚^{①③} 张亚平^{①②*}

(①中国科学院昆明动物研究所分子进化与基因组多样性实验室, 昆明 650223; ②云南大学云南省科技部共建生物资源保护与利用重点实验室, 昆明 650091; ③中国科学院研究生院, 北京 100039. *联系人, E-mail: zhangyp@public.km.yn.cn)

摘要 由于多种原因, 古老人类标本中的 DNA 一般都有不同程度的降解, 在分析时只能得到短片段的序列. 如何对这些序列片段的真实性进行甄别, 最大限度地挖掘出其中包含的信息, 是目前对古代人群遗传结构分析及其他古 DNA 研究中普遍存在的一个难题. 本文对近期国内古老人群 mtDNA 研究中存在的问题进行了评述, 并从 mtDNA 世系的系统发育关系角度, 对新疆(包括邻近的中亚地区)古老人群数据进行了重新分析. 结果显示, 这些古老 DNA 数据的可靠性存在或多或少的问题; 结合现代不同地理人群中的 mtDNA 变异情况, 能够对获得的距今数千年前的 mtDNA 的真实性进行判别和自检, 并能有效地对序列信息进行解码. 同时, 对中亚地区古老人群 mtDNA 数据重新分析的结果, 支持该地区欧亚人群基因交流融合由来已久的推测.

关键词 古老 DNA mtDNA 系统发育 问题 对策

在缺乏古代人群直接遗传信息的情况下, 根据现代人群的遗传分析和比较, 可为推测人群起源和迁移提供依据和佐证, 这在目前人类学领域较为常见^[1,2], 并且有人倡议建立一个新学科——考古遗传学(Archeogenetics)^[3]. 由于现代人群遗传结构受群体的迁移历史、基因流和遗传漂变等作用的影响, 是否与历史上人群的遗传结构相似不得而知. 同时, 现代人群中不同遗传标记揭示的遗传图景也可能不同. 因此, 关于一些特定人群遗传结构和群体历史的争论颇多. 对古代人群遗传结构进行直接研究, 有助于解决这些争议, 验证一些基于现代人群遗传数据得出的假说. 因而, 古老 DNA 研究一直受到高度重视.

1 古老 DNA 研究中的难题

由于在保存过程中受到各种因素的影响, 古老标本中的 DNA 均存在较为严重的降解现象, 导致古老 DNA 研究中存在两大难题. 难题之一是技术方面, 即如何从标本中提取出 DNA, 扩增出相关的目的片段, 并有效地防止外源 DNA 的污染. 针对这一难题, 目前有较多的研究经验和建议, 如采用比较特殊的 DNA 提取方法和技术; 多次独立地对同一样品、同一标本的不同部位取样进行分析^[4~6]; 在不同实验室独立进行实验操作, 互相印证^[6,7]; 与来自同一遗址中其他标本 DNA 的成功提取作为对照^[8], 或进行其他形式分析^[6,9], 看该遗址是否适合古老 DNA 的保存.

古老 DNA 研究的难题之二, 是如何对所获得的短的古 DNA 序列进行有效地破译, 这些 DNA 序列通常是位于 mtDNA 控制区第 1 高变区(hypervariable segment I of mtDNA control region, HVS I)的 100 ~ 300 bp 片段. 这一难题的准确解决与否, 直接关系到结论的正确性. 目前普遍采用的做法是将古老 DNA 序列和现代人序列放在一起, 构建系统发育树或网络关系图, 比较其中古老样本和现代样本的聚类情况^[10~13]. 这种做法虽然大多基于距离法或简约性原则, 但由于分析的片段非常短, 信息有限, 古老样本和现代人样本的聚类关系在由不同参数设置构建的系统树中可能不同^[14,15]. 而且, 将现代人序列剪切成和古老 DNA 片段一样的长度, 根据距离法构建个体或群体间的系统发育树而不考虑世系(lineage)间真实的系统发育关系, 往往不能充分地揭示片段中区分 mtDNA 单倍型类群(haplogroup)归属的有限信息, 会得出一些非常奇怪的结果. 一个比较典型的例子就是 Wang 等人^[11]的工作. 他们分析了我国山东淄博地区距今 2500 和 2000 年前^[10]及现代人群 mtDNA 185 bp 序列数据, 认为 2500 年前的淄博人群和现代欧洲人群比较相似, 2000 年前的淄博人群则表现为现代欧洲和东亚人群的中间类型; 过去 2500 年内我国北方人群母系基因库发生了由欧洲型向东亚型的替代. 这种“替代”推测一方面和我国人群的迁移历史^[16]有较大出入, 同时也与该小组近期工作^[17]相矛盾. 一方面, 他们推测认为过去 2500 年内, 我国北方

人群的母系基因库发生了由欧洲型向东亚型的替代^[11]; 另一方面, 他们又认为东亚不同地理人群母系遗传结构存在较大的同源性, 并且将这种同源性归于新石器时期农业人口的扩张^[17]. 显然, 新石器时期农业的扩张在时间上要早于 2500 多年前的史载时期. 如果新石器时期人类的活动就界定了后期人群的遗传结构, 那么又怎会有后期母系基因库替代发生? 最近, 我们对这批数据进行了重新分析. 结果提示, 山东古代群体 mtDNA 类型, 基本上都可以归属到报道的东亚人群 mtDNA 类型中^[18,19], 遗传结构总体上和我国南方人群相似性较大^[20]. 因此, 在对得到的短的序列片段进行分析和解释时, 一定要相当谨慎^[20].

既然短片段含有的信息量有限, 一个自然的想法就是尽量获取长的序列片段来进行分析. 由于古老样本中现实情况是 DNA 大多被降解, 每次扩增只能获得短的片段, 通过多个短片段的重叠扩增和测序, 理论上可以拼接出一个长一点的片段. 而且, 我们还可以借鉴目前现代人群 mtDNA 的研究结果, 针对不同 mtDNA 类群的一些编码区特征变异位点设计引物, 得到含有该位点的短片段, 进而分析获得更多的信息. 但在实际操作过程中, 这种多次扩增及对多个不同区域片段进行分析, 往往也会增加污染或其他出错的机会, 不能保证每次得到的片段都是古老样本的真实 DNA, 这无疑也为古老 DNA 研究增添了判别真伪的难度. 目前, 还没有方法能彻底解决这一难题. 从 mtDNA 世系的系统发育关系角度, 对得到的古 DNA 序列信息进行综合考虑和分析, 能够为判别这些片段的真伪提供线索, 从而对一些值得怀疑的序列在发表前重新分析证实, 避免将不准确的数据公布于众.

2 从 mtDNA 的系统发育关系角度解码古老 DNA

mtDNA 所特有的遗传特性, 如母系遗传、缺乏重组等, 决定了 mtDNA 序列上发生的每一次突变事件(回复突变和突变热点除外)都能在序列中记录下来. 通过对序列上累积的这些突变进行分析, 理论上可以重建 mtDNA 真实的系统树. 对于每一个 mtDNA, 根据其是否含有系统发育树中某个聚类枝(亦即单倍型类群)的特征界定突变, 也就是相对较古老一些的突变, 同时和数据库中经过编码区信息准确证实的

单倍型类群归属的其他 mtDNA 世系进行序列匹配或相近匹配(matching or near-matching)分析, 可以将它们划归到系统发育树的各个聚类枝中. 然后, 通过比较不同群体中各单倍型类群的分布情况和序列变异, 可以很好地显示出群体的地理分化、遗传结构的时空变化, 以及估算出群体扩张或发生迁移的大致时间. 这种可称之为系统发育地理分析(phylogeographic analysis)的方法, 在近期的工作中有较多的使用和报道^[18,19,21,22], 并且对于诠释古老 DNA 也非常有效^[18,20]. 下面就以最近崔银秋等人^[13]发表的数据作为例子, 具体阐明如何从 mtDNA 的系统发育关系来判别和解释古老 DNA 数据.

崔等人^[13]分析了距今 2000 ~ 2500 年前新疆吐鲁番交河故城一号台地 4 个个体 mtDNA HVSI 序列 363 bp 的变异情况(其实是标准序列^[23]中 16056 ~ 16378 区域间的 323 bp 片段, 应当去除两端引物序列), 同时还检测了单倍型类群 A(+ 663Hae III), B(CO II 和 tRNA^{Lys} 基因间 9 bp 片段的缺失(9 bp del)), C(-13259Hinc II), D(-5176Alu I)和 M(+10394Dde I , +10397Alu I)的特征变异位点. 这样对同一个样本同时进行控制区和编码区序列分析, 理论上比仅仅依据一段短的 HVS I 序列信息进行 mtDNA 单倍型类群的划分更准确、可靠, 值得提倡. 这其中, 类群 C 和 D 是 M 的子类群或亚类群(sub-haplogroup). 也就是说, 在 mtDNA 系统树中, C 和 D 聚类枝是包含在一个较大的称为 M 的大聚类枝中, M 类群的 4 个特征界定位点, 489 ~ 10400 ~ 14783 ~ 15043(相对于作为外群(outgroup)的非洲 L3 类群^[19]), 在类群 C 和 D 的个体中都会出现. 类群 C 和 D 的区别则体现在各自的单倍型类群的特异性突变上, 如 D 类群含有 5178A(变异位点加上突变后的碱基以突出该点是颠换; -5176Alu I)和 4883 等突变, 而类群 C 含有 3552A, 9545, 11914, 13263(-13259Hinc II)和 14318 等突变^[18,19]. 类群 A 在系统发育关系上属于一个大的类群 N, 类群 B 在系统发育关系上属于类群 R(类群 R 本身又是类群 N 的一个子类群). 而且, 大多数单倍型类群在控制区序列中也有特征变异模式(motif), 如类群 A 含有 16223 ~ 16290 ~ 16319 的变异模式, D 含有 16223 ~ 16362, C 含有 16223 ~ 16298 ~ 16327, B 含有 16189. 这些类群的巢式(nested)归属系统关系, 反映在崔等人^[13]检测的酶切位点和 HVS I 变异组合中, 只会出现以下结果: 类群 A, + 663Hae III ,

+ 5176*Alu* I, 9 bp non-del, -10394*Dde* I, -10397*Alu* I, +13259*Hinc* II, 16223 ~ 16290 ~ 16319; 类群 B, -663*Hae* III, + 5176*Alu* I, 9 bp del, -/+10394*Dde* I, -10397 *Alu* I, + 13259*Hinc* II, 16189; 类群 C, -663*Hae* III, +5176 *Alu* I, 9 bp non-del, +10394*Dde* I, +10397*Alu* I, -13259*Hinc* II, 16223 ~ 16298 ~ 16327; 类群 D, -663*Hae* III, -5176*Alu* I, 9 bp non-del, -10394*Dde* I, -10397*Alu* I, +13259*Hinc* II, 16189 ~ 16223 ~ 16362 (10397 位置上的突变会导致 10394*Dde* I 和 10397*Alu* I 酶切位点同时消失, 该变异界定了 D 类群的一个子类群 D5^[18,19])或-663*Hae* III, -5176*Alu* I, 9 bp non-del, +10394*Dde* I, +10397*Alu* I, +13259*Hinc* II, 16223 ~ 16362. 如果出现其他的一些酶切组合类型(不包括 CO II 和 tRNA^{Lys} 基因间 9 bp 片段的缺失, 该缺失在其他单倍型类群中可能会多次独立发生^[24]), 最大可能就是由于污染, 或者是实验过程中人为因素导致不同样本混淆造成不同 mtDNA 世系间的人工重组.

在崔等人^[13]文章的表 2 中, 来自不同墓地的样本 2 和 4 的 HVS I 序列仅含有 16234 变异(表 2 提到 16234 发生 C→T 转换是正确的, 但在第 4 页却将其误写为 T→C 转换和 A→C 颠换), 酶切结果相同(-663*Hae* III, +5176*Alu* I, 9 bp non-del, -10394*Dde* I, -10397*Alu* I, +13259*Hinc* II), 显示该类型属于 N 类群或 R 类群(包括 N 和 R 的子类群). 由于该个体的序列在 Richards 等人^[22]统计的欧洲人群中出现在两个 H 类型个体中, 并且还可能存在 3 个相差一次突变的相近类型(H: 16234 ~ 16293, 16234 ~ 16349 del; J: 16126 ~ 16234), 所以该 mtDNA 类型很可能属于欧洲型单倍型类群 H. 样本 3 的 HVS I 序列变异为 16129 ~ 16223 ~ 16298 ~ 16327, 具有类群 C 的变异模式, 并且与报道的新疆汉族 1 个 C 类型个体(XJ8435)^[18]和维吾尔族 1 个 C 类型个体^[25]的序列完全匹配. 此外, 在我国西安汉族人群^[17]、中亚人^[26]、蒙古人(只有编号 17.22 的个体, 而非表 3 中提到的 2 个个体)^[27]、西伯利亚 Koryak 人群^[28]和 Mansi 人群^[29]以及泰国 Akah 人群^[30]中也有与该样本序列完全匹配的 C 类型个体. 然而, 对该样本编码区的分析结果(-663*Hae* III, + 5176*Alu* I, 9 bp non-del, -10394*Dde* I, -10397*Alu* I, + 13259*Hinc* II)表明, 该个体属于 N 类群(包括其子

类群), 不属于类群 M, 更不属于类群 C. 因此, 该个体是 N 类群个体编码区和 C 类群个体控制区序列间的人为重组. 样本 1 的 HVS I 序列为 16298 ~ 16319, 在我们收集的我国人群 mtDNA 数据库(超过 2200 人)中没有发现完全匹配类型, 但和新疆地区维吾尔族^[25]、回族¹⁾以及武汉汉族样本^[18]中变异为 16223 ~ 16298 ~ 16319, 属于单倍型类群 M8a 个体的序列相近. 如果样本 1 也属于 M8a 类群, 那么就一定存在 5176*Alu* I 酶切位点, 而不是文献[13]表 2 中观察到的仅 D 类群中才出现的-5176*Alu* I. 所以, 该序列可能是类群 M8a 和 D 的一个重组, 并且 16223 位置上的突变被漏读. 这种不同 mtDNA 单倍型类群间序列人为重组的现象, 在近期报道的数据中并不少见, 譬如 Lee 等人^[31]报道的 98 个韩国人 mtDNA 控制区和细胞色素 B 基因全序列中, 就存在较多的不同单倍型类群间的序列重组和碱基漏读现象; Silva 等人^[32]测定的 8.8 kb mtDNA 序列中也存在人为的序列重组现象, 并且还漏读了相当多的变异位点. 采用本文介绍的从 mtDNA 系统发育关系角度对序列进行自检, 也许可以发现数据中存在的问题, 从而在发表之前重新核实.

3 中亚地区古老 DNA 的重新分析

根据上文提到的序列匹配和相近匹配分析, 还可以对报道的中亚地区其他古老人群 mtDNA^[33, 34]试着进行单倍型类群的划分, 并根据这些古老 mtDNA 在现今人群中是否有分布和分布的情况来推测更多的历史. Clisson 等人^[33]研究了距今 2000 多年前同一个墓穴发掘出的两个样本. 其中, 样本 1 的 HVS I 序列(对应于标准序列 16021 ~ 16417 区域)和标准序列完全相同, 很可能是欧洲人群中高频率分布的 H 类型, 而且该 H 类型在中亚人群^[26]和我国新疆地区人群^[25]也有分布. 值得注意的是, 我国广东汉族^[19]和云南傣族^[21,35]样本中也有 HVS I 序列和标准序列完全相同的个体, 但编码区序列分析表明, 这些个体属于东亚人群特有的类群 F2 或 M7. 样本 2 序列变异为 16223 ~ 16278 ~ 16362, 具有典型的 G 类群 HVS I 变异模式, 和我们测定的新疆地区蒙古族样本中 G 类型(+4831*Hha* I)个体序列¹⁾完全一致. 此外, 和样本 2 序列相差一次碱基突变的 G 类型, 如 16223 ~ 16227 ~ 16278 ~ 16362, 16223 ~ 16234 ~ 16278 ~ 16362 和

1) 姚永刚. 中华民族源流探讨-线粒体 DNA 的研究. 中国科学院研究生院博士学位论文. 2002

16093 ~ 16223 ~ 16278 ~ 16362 等, 在中亚人^[26]、我国新疆的回族、蒙古族和维吾尔族^[25]中也有分布。

Ovchinnikov 等人^[34]分析了距今约 2000 年前 Jeti-Asar 人群的 3 个标本, 这些标本的头骨都表现出欧洲人的特征。他们测定的 HVS I 序列对应于标准序列 16208 ~ 16401 区域, 酶切分析相对于崔等人^[13]多检测了 16517*Hae* III 位点, 样本 1 的变异为-663*Hae* III, + 5176*Alu* I, 9 bp non-del, -10394*Dde* I, -10397*Alu* I, +13259*Hinc* II, +16517*Hae* III, 16362。该个体和新疆维吾尔族中 1 个 H 类型个体的 HVS I 序列(16362 ~ 16482)相同^[25], 在新疆汉族的 3 个 D 类型 mtDNA 中(16362; 16174~16362; 16093~16362)^[18]中也出现。进一步和 Richards 等人^[22]汇总的欧洲人群数据库比较后发现, 该样本的序列在欧洲人群中的 H 类型(16095G ~ 16188 ~ 16362, 16111 ~ 16362, 16129 ~ 16362, 16129 ~ 16153 ~ 16362, 16150 ~ 16362, 16173 ~ 16362, 16189 ~ 16362, 16192 ~ 16362, 16362, 16362 ~ 16482), HV 类型(16168 ~ 16362, 16192A ~ 16362, 16362), J 类型(16126 ~ 16362, 16126 ~ 16193 ~ 16362)和 U 类型(16362, 16189 ~ 16362, 16153 ~ 16362, 16129 ~ 16362, 16129 ~ 16189 ~ 16362, 16129C ~ 16154 ~ 16362, 16129C ~ 16189 ~ 16362, 16129T ~ 16189 ~ 16362)等也出现。显然, 这里序列匹配分析不能准确提示样本 1 到底属于哪个类群。样本 2 变异为-663*Hae* III, + 5176*Alu* I, 9 bp non-del, -10394*Dde* I, -10397*Alu* I, +13259*Hinc* II, -16517*Hae* III, 16223 ~ 16257A ~ 16261, 属于典型的 N9a 类群。该序列在我国人群中分布频率为 0% ~ 8%, 且在南方和西北地区人群中频率较高一些^[18,21,25,35]。另外, 在中亚人^[26]、伊朗人^[22]、蒙古人^[27]、韩国人^[31]和日本人^[36]等都有分布, 而在欧洲人群中没有发现。所以, 该样本 mtDNA 属于亚洲人群特有类型。样本 3 编码区信息(-663*Hae* III, -5176*Alu* I, 9 bp non-del, -10394*Dde* I, -10397*Alu* I, +13259*Hinc* II, +16517 *Hae* III)显示该个体属于 D5 类群, 也就是说, 该个体 HVS I 序列理论上会有 16189 ~ 16223 ~ 16362 的特征变异模式, 但 Ovchinnikov 等人^[34]观察到的 HVS I 变异为 16223 ~ 16311 ~ 16357, 与云南汉族(YN164)和辽宁汉族(LN7593)中的 M10 类型序列^[18]完全匹配, 而在我们数据库的 D5 类型中找不到任何匹配或相近的匹配类型。所以, 一个可能的解释是该个体存在 D5 和 M10 类型序列间的人为重组。

4 小结

对来自中亚地区(包括我国新疆地区)距今 2000 多年前 mtDNA 的重新分析显示, 这些古老 DNA 数据的可靠性存在或多或少的问題。将报道的这些古老样本凑在一起(共 9 个), 可以从“群体”这一角度来窥视 2000 多年前中亚地区人群的母系基因库组成。显然, 2000 多年前中亚地区人群的 mtDNA 类型都可以归属到现代欧亚人群特异性的单倍型类群中^[18,19,22], 和现今该地区人群在母系遗传组分上类似^[25~26], 而且, 大多数古老 mtDNA 类型在当地现代人群中都能够找到序列完全一致的匹配个体, 这支持中亚地区欧亚人群基因交流和融合由来已久的推测。

短的 HVS I 序列提供的信息毕竟有限。一些个体的 HVS I 序列相同, 但也可能属于不同单倍型类群, 如上文提到 HVS I 序列和标准序列相同的个体可以是欧洲型的 H 类型, 也可以是东亚型的 M7 或 F2 等类型。分析的片段越短, 序列出现相同的机会就越多。这在客观上也为我们依据序列匹配和相近匹配分析带来难度和不确定性。所以, 对于古老样本, 在获取控制区序列后, 先依据得到的信息进行单倍型类群归属的预测, 进而对相关编码区类群特征性位点进行检测和验证是完全必要的^[20]。如果编码区提供的信息和控制区变异模式有矛盾, 这就提示该样本的 DNA 很可能有污染, 需要重新提取 DNA 进行分析。对于因材料受限制而不能够再次分析的样本, 对数据的矛盾之处在结果中作一个说明不失为明智的做法。本文倡导的这种对数据进行自检的思路, 对于动物古老 DNA 分析同样也有借鉴意义。我们建议, 在不能够完全排除实验方面存在缺陷可能带来的污染时, 一定要对得到的 mtDNA 数据从世系的系统发育关系角度进行自检, 解释也要谨慎, 这样才能够最大限度地避免错误的结论。

致谢 本工作为国家自然科学基金(批准号: 30024004)、云南省自然科学基金(批准号: 2000C0077M)和中国科学院重要方向(批准号: KSCX2-SW-2010)资助项目。

参 考 文 献

1 姚永刚, 张亚平. 线粒体 DNA 和人类进化. 动物学研究, 2000, 21: 392~406

2 Jin L, Seielstad M, Xiao C, et al. Genetic, linguistic and archaeological perspectives on human diversity in Southeast Asia.

- Singapore: World Scientific Publishing Co Pte Ltd, 2001. 41~171
- 3 Renfrew C. From molecular genetics to archaeogenetics. *Proc Natl Acad Sci USA*, 2001, 98: 4830~4832
- 4 Krings M, Stone A C, Schmitz R W, et al. Neandertal DNA sequences and the origin of modern humans. *Cell*, 1997, 90: 19~30
- 5 Stone A C, Stoneking M. mtDNA analysis of prehistoric Oneota population: implications for the peopling of the new world. *Am J Hum Genet*, 1998, 62: 1153~1170
- 6 Hofreiter M, Serre D, Poinar H N, et al. Ancient DNA. *Nat Rev Genet*, 2001, 2: 353~359
- 7 Handt O, Krings M, Ward R H, et al. The retrieval of ancient human DNA sequences. *Am J Hum Genet*, 1996, 59: 368~376
- 8 Cooper A. Reply to Stoneking: ancient DNA——how do you really know when you have it? *Am J Hum Genet*, 1997, 60: 1001~1002
- 9 Smith C I, Chamberlain A T, Riley M S, et al. Neanderthal DNA. Not just old but old and cold? *Nature*, 2001, 410: 771~772
- 10 Oota H, Saitou N, Matsushita T, et al. Molecular genetic analysis of remains of a 2000-year-old human population in China-and its relevance for the origin of the modern Japanese population. *Am J Hum Genet*, 1999, 64: 250~258
- 11 Wang L, Oota H, Saitou N, et al. Genetic structure of a 2,500-year-old human population in China and its spatiotemporal changes. *Mol Biol Evol*, 2000, 17: 1396~1400
- 12 Adcock G J, Dennis E S, Easteal S, et al. Mitochondrial DNA sequences in ancient Australians: implications for modern human origins. *Proc Natl Acad Sci USA*, 2001, 98: 537~542
- 13 崔银秋, 段然慧, 季朝能, 等. 交河故城古车师人的线粒体 DNA 分析. *高等学校化学学报*, 2002, 23: 1510~1514
- 14 Cooper A, Rambaut A, Macaulay V, et al. Human origins and ancient human DNA. *Science*, 2001, 292: 1655~1656
- 15 Gutiérrez G, Sánchez D, Marín A. A reanalysis of the ancient mitochondrial DNA sequences recovered from Neandertal bones. *Mol Biol Evol*, 2002, 19: 1359~1366
- 16 葛剑雄, 曹树基, 吴松弟. *中国移民史*. 福州: 福建人民出版社, 1997
- 17 Oota H, Kitano T, Jin F, et al. Extreme mtDNA homogeneity of continental Asian populations. *Am J Phys Anthropol*, 2002, 118: 146~153
- 18 Yao Y G, Kong Q P, Bandelt H J, et al. Phylogeographic differentiation of mitochondrial DNA in Han Chinese. *Am J Hum Genet*, 2002, 70: 635~651
- 19 Kivisild T, Tolk H, Parik J, et al. The emerging limbs and twigs of the east Asian mtDNA tree. *Mol Biol Evol*, 2002, 19: 1737~1751
- 20 Yao Y G, Kong Q P, Man X Y, et al. Reconstructing the evolutionary history of China: a caveat about inferences drawn from ancient DNA. *Mol Biol Evol*, 2003, 20(2): 214~219
- 21 Yao Y G, Zhang Y P. Phylogeographic analysis of mtDNA variation in four ethnic populations from Yunnan Province: new data and a reappraisal. *J Hum Genet*, 2002, 47: 311~318
- 22 Richards M, Macaulay V, Hickey E, et al. Tracing european founder lineages in the near Eastern mtDNA pool. *Am J Hum Genet*, 2000, 67: 1251~1276
- 23 Anderson S, Bankier A T, Barrell B G, et al. Sequence and organization of the human mitochondrial genome. *Nature*, 1981, 290: 457~465
- 24 Yao Y G, Watkins W S, Zhang Y P. Evolutionary history of the mtDNA 9-bp deletion in Chinese populations and its relevance to the Peopling of East and Southeast Asia. *Hum Genet*, 2000, 107: 504~512
- 25 Yao Y G, Lü X M, Luo H R, et al. Gene admixture in the silk road of China——evidence from mtDNA and melanocortin 1 receptor polymorphism. *Genes Genet Syst*, 2000, 75: 173~178
- 26 Comas D, Calafell F, Mateu E, et al. Trading genes along the Silk Road: mtDNA sequences and the origin of Central Asian populations. *Am J Hum Genet*, 1998, 63: 1824~1838
- 27 Kolman C, Sambuughin N, Bermingham E. Mitochondrial DNA analysis of Mongolian populations and implications for the origin of new world founders. *Genetics*, 1996, 142: 1321~1334
- 28 Schurr T G, Sukernik R I, Starikovskaya Y B, et al. Mitochondrial DNA variation in Koryaks and Itel'men: population replacement in the Okhotsk Sea-Bering Sea region during the Neolithic. *Am J Phys Anthropol*, 1999, 108: 1~39
- 29 Derbeneva O A, Starikovskaya E B, Wallace D C, et al. Traces of early Eurasians in the Mansi of northwest Siberia revealed by mitochondrial DNA analysis. *Am J Hum Genet*, 2002, 70: 1009~1014
- 30 Oota H, Settheetham-Ishida W, Tiwawech D, et al. Human mtDNA and Y-chromosome variation is correlated with matrilineal versus patrilineal residence. *Nat Genet*, 2001, 29: 20~21
- 31 Lee S D, Lee Y S, Lee J B. Polymorphism in the mitochondrial cytochrome B gene in Koreans: an additional marker for individual identification. *Int J Legal Med*, 2002, 116: 74~78
- 32 Silva Jr W A, Bonatto S L, Holanda A J, et al. Mitochondrial genome diversity of Native Americans supports a single early entry of founder populations into America. *Am J Hum Genet*, 2002, 71: 187~192
- 33 Clisson I, Keyser C, Francfort H P, et al. Genetic analysis of human remains from a double inhumation in a frozen kurgan in Kazakhstan (Berel site, Early 3rd Century BC). *Int J Legal Med*, 2002, 116: 304~308
- 34 Ovchinnikov I, Buzhilova A, Mednikova M, et al. Ethnic affinities of the ancient human Jety-Asar population by mitochondrial DNA analysis. *Electrophoresis*, 1999, 20: 1729~1732
- 35 Yao Y G, Nie L, Harpending H, et al. Genetic relationship of Chinese ethnic populations revealed by mtDNA sequence diversity. *Am J Phys Anthropol*, 2002, 118: 63~76
- 36 Imaizumi K, Parsons T J, Yoshino M, et al. A new database of mitochondrial DNA hypervariable regions I and II sequences from 162 Japanese individuals. *Int J Legal Med*, 2002, 116: 68~73

(2002-11-14 收稿, 2003-01-16 收修改稿)