



ISSN 1005-8036  
CN 11-3499/N

# 中央民族大学学报

JOURNAL OF THE CENTRAL UNIVERSITY FOR NATIONALITIES

- 用于空间半导体区熔生长仿真系统的数学模型
- 计算机辅助合成三维动态彩色全息图设计
- 番茄耐冷性RAPD分子标记的筛选及特异片段的克隆
- 上海原住民的Y染色体遗传分析

ZHONGYANG MINZU DAXUE XUEBAO

2004 VOL. 13 NO. 1

**Title on the Cover:**

Yang Jun, Li Hui, Jin Jian-Zhong, Jin Li, Lu Da-Ru (2004) Genetics Analysis of Shanghai Aborigine with SNPs on Y Chromosome. *Journal of the Central University for Nationalities* 13(1):60-69.

# 上海原住民的 Y 染色体遗传分析

杨俊, 李辉, 金建中, 金力, 卢大儒

(复旦大学生命科学学院现代人类学研究中心, 上海 200433)

**摘要:** Y染色体是继线粒体之后的研究人类进化的又一热点,通过对其上的多态性位点研究也证实了非洲起源说。Y染色体上的各单核苷酸多态(SNP)位点组合成的单倍型与民族系统的时间发生尺度比较一致,所以是鉴别民族迄今最恰当的遗传材料。M119C这一 SNP突变是百越民族的特色,M119C-M110C和M95T-M88G更是百越的特有种。良渚文化的族属是考古学界的一个争议焦点,有百越和三苗两种说法。良渚文化与马桥文化之间差异极大,其居民遗传上是否继承也是有问题的。本研究通过对马桥地区良渚时期、马桥时期、战国、明代及现代居民 Y染色体若干 SNP位点的检测,重点考察了M119位点,发现马桥地区各时期的居民都有较高的M119C和M95T类型,具有典型的百越特征,而没有苗瑶的特征。这说明自古以来马桥地区居民都是越人,并且一脉传承至今。现代上海的原住民“宕俸”人与台湾原住民的遗传结构最接近,其次接近西南部的侗傣民族。主成分分析及主成分与 Y染色体单倍型的相关分析发现,单倍型H9和H11区分了东越与西越,H7区分了苗瑶,H8区分了氏羌,H11与H5区分了濮越系统与非濮越系统。各单倍型能对各民族系统较清晰地地区分和代表,是族属分析的良好材料,值得深入研究。

**关键词:** Y染色体; SNPs; 主成分分析; 百越; 古DNA; 上海原住民; 良渚

**中图分类号:** Q987; Q986 **文献标识码:** A **文章编号:** 1005-8036(2004)01-0060-10

## 1 引言

随着现代分子生物学研究技术的发展及其在遗传学中应用的日益广泛,对于人类种族之间亲缘关系的研究方法也在不断增加和改进。30年代血清学的出现,使人们开始得以从微观上来研究生物遗传标记。首先是多态基因组分型,如HLA,ABO血型等。到了60年代,出现了第一代遗传标记——限制性片段长度多态性(RFLP, Restriction Fragment Length Polymorphism),它通过酶切——电泳达到检测目的。但是,RFLP在基因组中普遍性不高,多数情况下,所研究的基因附近或者内部并不存在适合的限制性酶切位点。到了90年代,被第二代遗传标记——微卫星(短串联重复序列,short tandem repeat, STR)所取代。微卫星是由串联重复序列组成,每单元长度2-6个碱基之间,广泛存在于真核基因组中,因为重复单位数目不同而呈现高度多态,由Genescan检测STR的重复单位数。STR在染色体上不表达,一般只适用于不涉及基因功能,如人类起源和迁徙等方面的研究,但是,STR的突变率较高,对久远的历史事件的纪录不够精确,因此被第三类遗传标记——单核苷酸多态性(Single Nucleotide Polymorphisms, SNPs)所代替<sup>[1-2]</sup>。

SNPs是指基因组内特定核苷酸位置上存在的两种不同的碱基,其频率最低的一种不低于1%。它是基因组中最普遍、最精细的多态性遗传标记。比较任意两条同源染色体的碱基序列,平均约1000bp就有一个SNP,人类基因组有30亿个碱基,因此有300万以上的SNPs。因此尽管SNP的多态性不如STR,但是其分布频密,故整体而言,其多态性最高。由于SNP位点通常只有2个等位基因(biallelic),因此往往只要进行+/-的分析,可以用PCR-RFLP、等位基因特异性PCR(allele-specific polymerase chain reaction)、变性高效液相层析(DHPLC)和DASH(Dynamic Allele Specific Hybridisation)等方法进行自动化检测。为研究

群体遗传学,特别是人类起源与迁徙,提供了广阔的前景<sup>[1-2]</sup>。

对于 DNA 多态性研究有三种方法,首先是对于常染色体的研究,其含有 3,000,000kb 碱基对,信息量较大,但是其多态性来源除了突变以外,还有可能来自减数分裂中的重组,因此不能完全忠实地反映进化过程中的突变。其次是 1987 年加州大学伯克利分校的 Allan Wilson 和埃摩里大学的 Douglas Wallace 发展的用人类母系遗传的线粒体 DNA 研究人类起源问题,并由此提出了“非洲夏娃”假说。该假说认为,现代人约于 15 万年以前起源于非洲,在 10 万年中散布到欧亚大陆各地,又被称为“走出非洲”假说<sup>[3-5]</sup>。1997 年,Underhill 根据 Y 染色体研究结果提出了“亚当学说”——最早的男性出现在非洲,与“夏娃学说”相吻合<sup>[6]</sup>。虽然线粒体 DNA 突变率较高,多态位点多,且无重组,但所含信息量少,又集中于 D 环区,回复突变率高,对于研究 DNA 多态仍然存在一定的局限性。

人类群体遗传学的研究过去很少采用 Y 染色体的遗传标记,因为在 Y 染色体的非重组区很少有二态性的 RFLP 存在。这在广泛运用 RFLP 作为遗传标记的时期很难用作群体分析。Y 染色体特异的 STR 使人们开始将 Y 染色体的多态标记用于群体遗传学的研究。近来,随着基因组测序工作的进展,特别是随着新的检测技术 DHPLC 的应用,SNP 成为新一代遗传标记后,研究者们认识到应用 Y 染色体遗传标记作为研究现代人迁移模式理想工具的重要性,并在实践中进行应用<sup>[7-8]</sup>。

1999 年金力等人开始发展了一套 Y 染色体的单核苷酸多态性标记(SNP)来研究人类起源与迁徙问题<sup>[9-11]</sup>。每个 SNP 位点上有一个碱基存在两种类型,一种是原始型,另一种是突变型。Y 染色体上各位点在历史中陆续发生的突变形成了许多种组合。这就是 Y 染色体的 SNP 单倍型。它们在各民族中有独特的分布格局。Y 染色体拟常染色体非重组区段(pseudoautosomal region)的 SNP 标记是目前公认的研究早期人类进化和迁移最理想的工具。这是由于 Y 染色体有 60000kb,具有足够的信息量,且父系遗传。Y 染色体因为只有男子体内存在,而且正常有生育能力的男子只有一条,所以其非重组区可以有效地排除交换重组的干扰,同时 SNP 突变率低更能稳定地遗传,可以忠实地记录进化事件,再由于以单倍体存在的 Y 染色体,有效群体大小远远小于常染色体位点,同时易产生人群特异性的单倍型<sup>[8]</sup>。目前,这一技术已找到超过 200 个 Y 染色体 SNP 位点,其中在东亚人群中具多态性的位点应用于对中国人群的大规模研究中,得到了这些 SNP 在东亚包括中国各人群中的特征分布<sup>[12]</sup>。许多微妙而鲜为人知的民族形成的过程都在这些位点中昭示于众。在 2000 年的 Nature Review 和 2001 年的 Science 都发表了相关的研究成果<sup>[12-13]</sup>。

东亚人群中分布的 19 种 Y 染色体单倍型,编号从 H1 到 H19。其中 H2, H3 是从中亚进入的类型。从东亚一些民族群体的 Y 染色体单倍型分布看来, H9 分布于傣、壮、侗、黎等侗台民族和高山族中,并向东北延伸; H10 和 H12 几乎只分布于侗台民族和高山族,可以认为是百越民族特有的。H7 分布于瑶、畲等苗瑶民族中,可以认为是其特有类型。国内在苗瑶民族中没有百越的 H9, H10, H12, 而百越民族中也没有 H7。所以这几种单倍型是这两个系统民族的鉴别标准。高山族中 H9、H10、H12 几乎占了 75%, 阿美人占 100%<sup>[9]</sup>, 可见其余百越民族亲缘关系之深。在西部的藏缅民族和汉族中 H8 的比例极高, 在景颇等民族中达到 100%, 可推测 H8 发生于氏羌系统民族, 并随着汉族的扩散而带到各民族群体中。

自从证实了在古人类遗骸中某些坚硬组织及在缺水条件下的软组织存在着少量部分降解的 DNA, 对古代人群的分子遗传学研究便蓬勃开展起来了。

成功地提取古 DNA, 最早是 1980 年湖南医科大学从汉代马王堆女尸中提取了古 DNA 和 RNA<sup>[14]</sup>。古 DNA 研究的突破性进展是利用了分子克隆这一技术。但是, 分子克隆需要的古 DNA 量较大, 而多数古人遗骸数量较少, 并与其他研究价值, 这就限制了古 DNA 的研究。80 年代中期, 聚合酶链式反应(PCR)技术的出现<sup>[15]</sup>使古 DNA 和古 RNA 的研究真正迅速开展起来, 由于 PCR 技术仅需痕量的 DNA 作模板, 它成为几乎所有研究者的有力工具。对古 DNA 的研究使我们有机会了解古生物和古人类的遗传特征及进化、迁移的过程等等。目前的研究进展主要有三个方面: 一, 个体水平遗传信息的获取, 从个体水平我们可以了解到史前人类的遗传信息, 如位点分布等; 二, 独立群体水平遗传信息的获取, 对一个较小范围内的两个或多个个体进行遗传信息的比较, 我们可以得出他们之间的家族亲缘关系, 及古代人群的遗传信

息;三,多群体水平遗传信息的获取,比较不同群体可以得出其进化的过程或迁移的过程,特别是结合考古的人文学方面的研究成果,我们可以从时间和空间等角度较全面了解民族群体的发展史,并且构建进化树。

上海是拥有上千万人口国际性大都市,而上海郊区各地的居民却保留着各式各样迥异的口语和民俗,特别是闵行、奉贤、金山、松江四区接合部的原住民,其自称“宕傣”(Dangdaic)与“扬越”的古音接近<sup>[16]</sup>。我们对上海南郊原住民进行了随机采样,进行 Y 染色体 SNP 检测。希望借此能较深入地调查弄清“上海人”的历史源流。

## 2 材料与方法

### 2.1 研究对象

由上海博物馆提供的马桥出土的良渚时期(约 4000B. C.)骨骼 5 块、马桥时期(约 2000B. C.) 2 块、战国时期(约 500B. C.) 2 块、松江等地明代(约 16 世纪)墓葬骨骼 3 块(主要是白齿)、马桥现代居民血样 26 份、马桥东南以黄浦江相隔的金汇镇居民血样 48 份,全部为男性。

### 2.2 材料

(1)引物选择 引物都根据测序得到的 SNP 进行设计,详见方法。

(2)实验所用试剂及其配方 抗凝剂 ACD(1.92g 柠檬酸,5.28g 柠檬酸钠,5.88g 葡萄糖,加水至 400ml);DNA 抽提缓冲液[Tris·Cl(pH8.0)10mmol/L,EDTA(pH8.0)1mol/L,胰 RNA 酶 20μg/ml,SDS0.5%];新洁而灭;EDTA 5ml/ml pH 8.0;TE(10mM Tri-HCL pH 7.5,1mM EDTA pH8.0);Proteinase K;饱和酚;氯仿;醋酸钠;无水乙醇;75%乙醇;PCR 体系(dNTP 2mM;引物 10mM;酶 5 U/μl;10 PCR buffer;MgCl<sub>2</sub>25mM);10×TBE(107.8g Tris,55.0g 硼酸,8.2g EDTA 二钠盐);琼脂糖;EB;电泳上样液(1% 溴酚蓝,30% 甘油)。

(3)主要仪器 PTC-100 PCR 仪(MJ Research. Inc);SCR-300 电泳仪(上海万达生物工程有限公司);Electro-4 Gel Tank 电泳槽(Hybaid 公司);FR-200 紫外与可见分析装置(上海复日生物实验技术研制所)。

### 2.3 方法

(1)血样 DNA 的抽提 将血样与 DNA 提取缓冲液按 2:3 混合,加 10 微升蛋白酶 K,55℃ 水浴 8 小时——加等体积饱和酚,室温剧烈震荡 15 分钟,12 000rpm 离心 10 分钟——取上清,加等体积氯仿:异戊醇(24:1),剧烈震荡 15 分钟,12 000rpm 离心 10 分钟——取上清,加入 1/10 体积的 3M 醋酸钠及 2 体积的无水乙醇,-20℃ 沉淀 30 分钟——12 000rpm 离心 15 分钟,去清液,加 500 微升 75%乙醇洗涤沉淀,12 000rpm 离心 5 分钟,去清液,烘干——加 50 微升 TE 溶解沉淀,-20℃ 保存待用。

(2)古代 DNA 提取 清洗,用手术刀片将化石表面刮去,酒精消毒,紫外灭菌——用灭过菌的刀片将化石表面刮去,去深层骨粉,装入 1.5ml 离心管,加入 1 毫升新洁而灭稀释液,震荡过夜——去清液,用 1 毫升无菌水清洗,打散骨粉,13 000rpm 离心 5 分钟,去上清,重复 2 次,假如 1 毫升无菌水,震荡过夜——用无菌水清洗,打散骨粉,4 000rpm 离心 10 分钟,去清液,加入 1 毫升裂解液(SDS),震荡过夜——将样品加入液氮冷冻 5 分钟,然后在 55℃ 烘箱中解冻(20~30 分钟),反复 4 次,放入液氮冷冻过夜——将样品在 40℃ 烘箱中解冻 40 分钟,3 000rpm 离心 5 分钟,去清液,搅碎骨粉,加 1 毫升裂解液和 50 微升蛋白酶 K,震荡过夜 2 天——酚氯仿抽提 DNA——乙醇沉淀 DNA,-20℃ 保存待用

表 1 Long-PCR 体系

Tab. 1 Prescription of Long-PCR

药品	dNTP	10 buf	Mg <sup>++</sup>	Prim	Temp	Antibody	AbBuf	Taq	H <sub>2</sub> O	Vol
体积(μl)	1	1	1	0.3×N	1	0.1	0.4	0.1	补足体积	10

(3) DNA 多聚酶链式反应 本论文在实验过程中,先对目标片段进行一次长扩增(Long-PCR),再进行一次巢式扩增(Nest-PCR).扩增产物用相应的限制性内切酶切割.

表 2 Long-PCR 引物序列

Tab. 2 Sequences of Long-PCR primers

位点	上游引物(5'~3')	下游引物(5'~3')
M89	TTCCAAACAGCAAAGACCAT	CAAGTGAACCTCCCTAACA
M95	CAACTTTCAGGTCAAGGAGAT	TTGGTTTTGFTCATTTCTTT
M122	TGGGTTTTTCTTCTTCTTA	CCTTCAGAATGTGTGCCTAA
M9	GCAACATATAAAACTTTTCAAGG	CACAAAGGACATAAACAAAGAAC
M45	GGAGAAAAAAGGATGTACA	AGGTGGATCAGGGTGAAGAAC
M134	ATGGAGAGGGAGGATGATA	TTCGTATTGTGTGGGAGTTAC
M130	TAGATTTTGTGGTGGTGGTGG	TTTTTGTACTTCTGTTTGGGA
M119	GGACCTAGCTAAACCTGAAATA	TAACCATGAGAAAAAGAAAGACA
M110	AAACCAAAGGCTACCCATTC	TGTAGGGGGATTTTGTCTTA
M88	TGAACTCTCCCTGCCAAAAA	GGGGTCTTGTGTGTGTGCTA

循环条件: 94 ×3min (94 ×40sec 63 [- 0.5 /cycles] ×40sec 72 ×1min) ×14 cycles  
(94 ×40sec 56 ×40sec 72 ×1min) ×20 cycles 72 ×5min

将 Long-PCR 产物稀释 5~10 倍,作为 Nest-PCR 的模板

表 3 Nest-PCR 体系

Tab. 3 Prescription of Nest-PCR

药品	dNTP	10 buf	Mg <sup>++</sup>	Primer	Temp	Taq	H <sub>2</sub> O	Vol
体积(μl)	1	1	0.8	0.3	1	0.1	5.8	10

表 4 各位点的扩增引物和扩增产物

Tab. 4 Primers and productions of Nest-PCR

位点	退火温度	巢式扩增引物(5'~3')	产物长度	内切酶	原始型	突变型	单倍型
M130	55	TATCTCTCTTCTATTGCAAG CCACAAAGGGGAAAAAACAC	205	BslI	C 切	T	H1
M1(YAP)	53	CAGGGGAAAGATAAA GAAATA ACTGCTAAAAAGGGGATGGAT			无 Alu	Alu 插入	H2,3
M89	52	GAAA GTGGGCCCCACAGAA GGA GCAAAATCAGGCAAA GTGAGACAT	100	NlaIII	C 切	T	H4
M9	55	GAAACGGCCTAAAGATGGTGGAT AAACTGAATCTTTTTTCTTCATTTTTG	210	BamHI	C 切	G	H5
M119	55	AGGTAAATGACTCACCCTAAGGAAG GGGTATTCCAAATTCAGCATACACCC	161	BstI	A	C 切	H9
M110	55	AACATCTCTGTGAGACTACTGG ATTGACTTCTTTTCCCC	200	NlaIII	T	C 切	H10
M95	55	ATAAGGAAA GACTACCATATTA CCG TTTGAAAGCCCCAGTTGTGAG	202	HhaI	C 切	T	H11
M88	55	TCTTATCTCTGCTTCTTCCCG CATGTGATGGTTCAGTGGTGTGA	146	BstI	A	G 切	H12
M122	55	TAGAAAA GCAATTGAGATACTAATICA CCGATGCTGATATGCTAGTICAG	122	NlaIII	C 切	T	H6
M134	55	AAGGACCAAGAAAATGATGATCG TTTGATGATCTTCTTTGGGCTTC	207	NlaIII	G 切	缺失	H8
M45	55	ATTGGCAGTGA AAAATTATAGCTA TGCCTTTGCTACAACCTCTCCTA	162	BfaI	G 切	A	H13

用琼脂糖凝胶电泳检测扩增和酶切产物的长度,分析各位点是否突变.

对照检验:实验者 DNA 作阳性对照,空白作阴性对照,同样品一起作检测。  
实验结果用对应的酶切检测,如图 1 所示。

135 bp →  
161 bp →

样品号 M210 M209 M205 M204 M102 M201 M203 M88 天钥 松江 明墓

图 1 墓葬人骨 Y 染色体 DNA M119 酶切电泳图像

Fig. 1 Electrophoretic photo of enzymatic reaction on ancient DNA M119 segments

每个泳道为一个个体的扩增产物, M119 突变为 A C, 没有发生突变的个体, 其扩增产物酶切不动, 突变的个体, 能被酶切开。

#### 2.4 数据的处理与分析

本实验数据除了用常用的 EXCEL 电子表格软件记录储存, 并作简单的数字图形分析以外, 还由 SPSS10.0 统计软件做主成分分析, 把各群体分类。各主成分与各单倍型作相关分析, 解析各单倍型的分类意义。

## 3 结 果

### 3.1 马桥地区古代遗骨的 Y 染色体 SNP 位点类型

马桥古代样品中除良渚时期 M207 的一块腰椎骨以外, 其余 11 颗臼齿都成功地扩增和检测。阴性对照无产物, 阳性对照与样品结果不一致, 所以可知样品没有污染, 结果是可信的。各样品各位点的类型见表 5。

表 5 马桥古代个时期样品的 Y 染色体若干 SNP 位点类型

Tab. 5 Types of Several Y-SNPs of Maqiao ancient samples

样品 位点	良渚文化时期				马桥文化时期		战国时代		明代		
	M204	M205	M209	M210	M201	M102	M203	M88	松江	天钥	松江
M130	C	C	C	C	C	C	C	C	C	C	T
YAP	-	-	-	-	-	-	-	-	-	-	-
M89	T	T	T	T	T	T	T	T	T	T	C
M9	G	G	G	G	G	G	G	G	G	C	C
M119	C	A	C	C	C	A	A	A	C	A	A
M110	T	T	T	T	T	T	T	T	T	T	T
M95	C	T	C	C	C	T	T	C	C	C	C
M88	A	A	A	A	A	A	A	A	A	A	A
M122	C	C	C	C	C	C	C	T	C	C	C
M134	G	G	G	G	G	G	G	A	G	G	G
M45	G	G	G	G	G	G	G	G	G	G	G
单倍型	H9	H11	H9	H9	H9	H11	H11	H8	H9	H4	H1

各样品 SNP 位点类型的组合都符合图 1 的系统,所以结果是可信的. 良渚、马桥两个时期,马桥地区的居民都有一半是 M119 突变的,战国没有突变,明代墓葬中松江的有突变,其他地区没有发现 M119C 突变,而出现了 M130T 突变和 M89T-M9C 类型. M95 T 出现于良渚、马桥、战国三个时期.

### 3.2 上海郊区群体的分析结果

对马桥和金汇两个镇的样品检测结果,各位点突变分布见下表.

表 6 马桥镇与金汇镇居民 Y 染色体 SNP 突变个体数和 M1Alu 插入个体数(比较其他民族<sup>[10]</sup>),

Tab. 6 Frequencies of Y-SNP mutations of Maqiao and Jinhui populations

群体	样本量	H1	H2	H3	H4	H5	H6	H7	H8	H9	H10	H11	H12	H13	H14
		M130	YAP	M15	M89	M9	M122	M7	M134	M119	M110	M95	M88	M120	M45
金汇	48	2.08	6.25				4.17		27.08	39.58	12.5	8.33			
马桥	26	3.85	3.85				3.85		34.62	34.62	11.54	7.69			
台湾:															
布农	9									11.11	66.67		22.22		
泰雅	24						29.2	4.17	4.17	54.17	8.33				
排湾	11								18.18	54.55	27.27				
阿美	6									100					
侗族:															
侗族	10	20					10		20	20	10	20			
黎族	11								9.09	27.27		54.6	9.09		
壮族	28	3.6		3.6	7.1	3.6	3.6		25	17.9		25	10.7		
汉族:															
浙江	50	12				6	24		26	26		6			
江苏	55	7.27	1.82		1.8	16	23.6	3.64	21.82	16.36		3.64		1.82	
上海	30	6.67	3.33			13	23.3	3.33	16.67	26.67				3.33	
福建	13	7.69				7.7	38.5		38.46				7.69		
广东	15						40		26.67	20			6.67	6.67	
南洋:															
马来	13			7.7	15	23.1				38.5		7.7			7.7
爪哇	15				53	13.3				20		13.3			
关岛	6	16.7	16.7			33	33.3								
萨摩亚	29	48.3				6.9	41.4								3.45

马桥和金汇的各 SNP 突变频率分布比较相似. 金汇的 M119C 频率特别高,占总数的一半以上. 这与高山族,侗族,黎族等群体结构相似,但远高于浙南,福建的群体. 马桥的 M119C 占 1/3 左右,与浙南,福建相近.

表 7 是对南方各群体根据 Y 染色体 SNP 多态位点单倍型主成分分析各主成分具体数值,图 2 中画出了各族的散点图. 表 8 分离了构成主成分的单倍型.

表 7 各民族 Y 染色体数据主成分

Tab. 7 Principle components of Y-SNPs of the populations

	PC1	PC2	PC3	PC4	PC5	PC6	PC7
阿美	0.563	-0.585	-0.004	-0.305	0.414	0.179	-0.165
泰雅	0.585	-0.642	0.001	-0.250	0.399	0.075	-0.083
布朗	0.533	0.757	-0.225	-0.185	0.045	-0.020	0.121
布农	0.076	-0.393	0.016	-0.190	-0.121	-0.682	0.550
高棉	0.663	0.445	-0.206	0.053	-0.456	-0.067	0.025
侗	0.784	-0.035	-0.326	0.282	0.276	-0.047	0.253
黎	0.740	0.235	-0.514	-0.305	-0.018	0.116	0.036

续表 7

	PC1	PC2	PC3	PC4	PC5	PC6	PC7
马来	0.583	-0.077	0.439	0.364	-0.431	-0.241	0.207
美拉尼西亚	-0.243	-0.019	0.151	-0.126	0.009	0.685	0.646
泰	0.349	0.075	-0.442	0.403	-0.563	0.103	-0.018
北方汉	0.613	-0.017	0.533	0.511	-0.122	0.109	-0.119
玻利尼西亚	0.021	0.332	0.734	-0.047	0.066	0.418	0.389
排湾	0.627	-0.722	0.003	-0.105	0.209	-0.103	0.125
畲	0.258	0.362	0.048	0.679	0.349	-0.086	0.000
南方汉	0.756	-0.203	0.204	0.546	-0.044	0.177	-0.084
雅美	0.640	0.380	-0.500	-0.390	0.025	0.088	0.058
瑶-布努	0.174	0.687	-0.085	0.235	0.607	-0.155	0.083
拉伽	0.280	0.649	-0.469	-0.113	0.203	-0.047	0.088
瑶-勉	0.005	0.416	0.286	0.432	0.688	-0.179	0.041
彝	0.434	0.528	0.601	-0.083	-0.228	-0.030	-0.132
壮	0.836	0.042	-0.391	0.175	-0.214	0.152	-0.010
金汇	0.765	-0.593	-0.141	0.120	0.085	0.067	0.040
马桥	0.758	-0.543	-0.141	0.296	0.010	0.080	0.063
巴达克	0.721	0.091	0.334	-0.457	0.101	0.120	-0.249
爪哇	0.587	0.267	0.683	-0.269	-0.030	0.023	-0.165
达雅克	0.596	-0.126	0.548	-0.397	-0.049	-0.356	0.138
通拉加	0.685	0.477	0.141	-0.409	-0.068	-0.102	-0.013

表 8 各 Y 染色体单倍型与主成分的相关性

Tab. 8 Correlation ships between Y-SNPs and principle components

	H1	H2	H3	H4	H5	H6	H7	H8	H9	H10	H11	H12	H13	H14	H15	H16	H17	
PC1	- *						- *	*	*									- **
PC2	*		*		*			- *	- **	- *	**							
PC3					**	*					- **							
PC4	*						*	**										
PC5	**	- *		- *	**	- *						- **						

各主成分都有其相应的主要单倍型。

## 4 讨 论

### 4.1 良渚时期和马桥时期先民的族属分析

虽然因为古代遗骸的样品难得,我们检测的数目相当少,我们还是看到了良渚时期和马桥时期马桥居民的主要遗传结构。这两个时期的 Y 染色体 SNP 单倍型都是以 M119C 和 M95T 两类突变型为主。这一方面说明,古代马桥居民是典型的百越民族群体,与汉藏群体及三苗集团蚩尤部落(主要为 M122T 下各支)毫无关系;另一方面也说明了至少在马桥地区,良渚时期与马桥时期的居民在遗传上是有一致性的,仅是文化在其他因素影响下发生了较大的变化。但是,即便在良渚鼎盛时期,马桥地区与良渚中心之间文化风俗上也是有差异的。所以,良渚时期马桥居民的遗传结构是否能代表整个良渚文化先民的遗传结构是值得商榷的。就良渚时期与马桥时期遗传上的一致性而言,也只能说明这两个时期居民都属于百越。而百越在南方分布是极广泛的,浙南闽北肩头弄文化等被认为对马桥文化有贡献的类群是否也属于



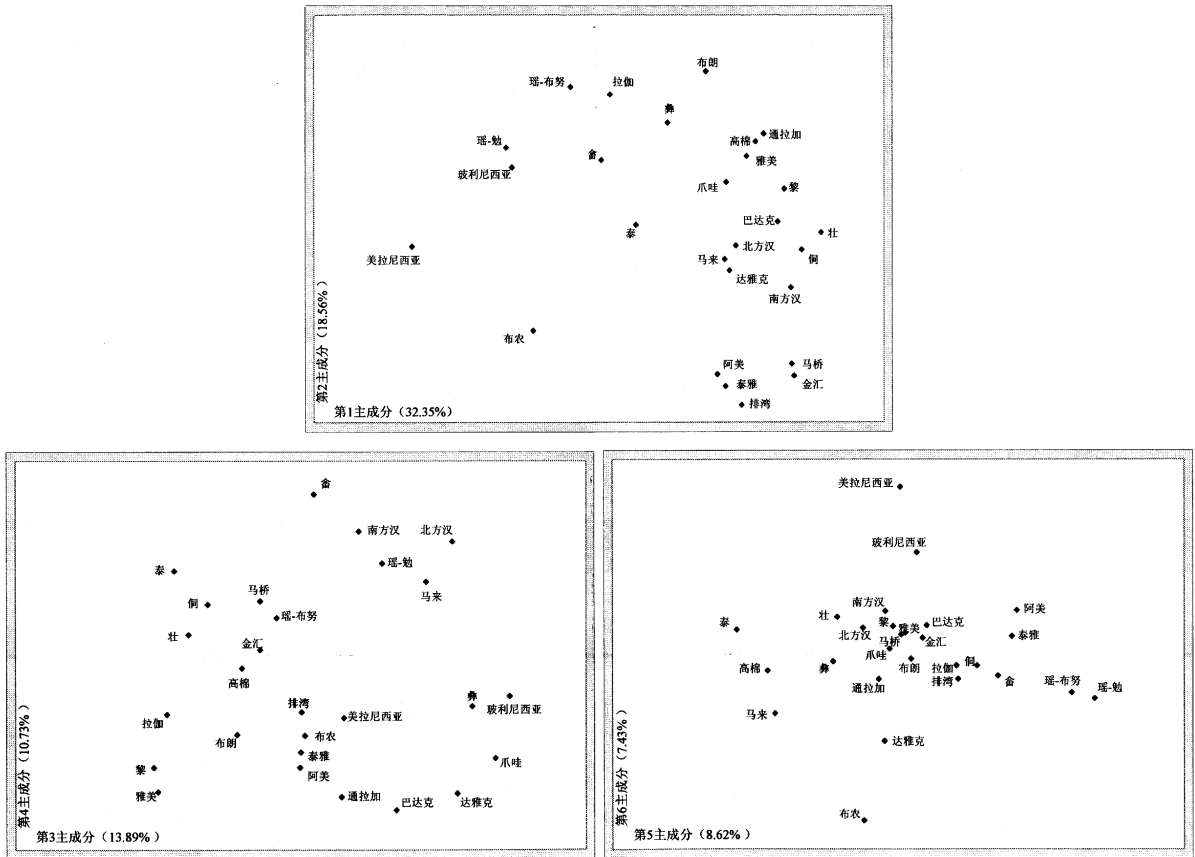


图 2 上海郊区人群与其他群体比较的主成分分析

Fig. 2 Component maps of Shanghai aborigines and other populations

百越也待研究,不能排除他们对马桥文化先民血脉也有贡献.但是马桥先民与北方二里头文化等先民之间遗传上无关系应是较明确的.

### 4.2 马桥地区历史时期居民遗传传承

我们的数据非常倾向于良渚时期马桥时期居民遗传上有继承性.在战国和明代样品只有两三个的情况下仍发现了 M95T 和 M119C,说明这两种突变类型在群体中可能仍是较主要的,史前的血脉是一直传承下来的.在战国时期出现了之前未发现的 M122T 突变. M122T 可能在原先并不存在于马桥地区,是春秋战国时期从西部传入的.当时中国大部分地区战乱频繁,太湖流域经历了吴越争霸和楚国扩张的事件,荆吴人口的流入是很有可能的.而荆吴应带有西部的 M122T.明代是中国与外交流较频繁的历史时期,所以上海地区遗传结构也开始复杂,但不排除本来马桥地区就有 H4 和 H1 类型.

### 4.3 现代上海郊区居民的遗传结构

我们就上海郊区人群与其他人群进行了比较,现代马桥地区居民的 M119C 依然是主要类型,特别是金汇镇,这表明马桥先民的血脉仍是当地人的主要构成,当地人多数在遗传上属于百越系统的.此外 M1A1u 插入的发现说明西北民族对本地也有影响,这种影响可能主要发生于元代. M122T 的频率已经相当可观,在马桥镇已占了近半.这很可能是被汉族同化的结果.在浙南和福建的群体中同样有这样的遗传结构,暗示着东南部的许多汉族群体是来源于百越的.金汇的遗传结构与现代侗台民族和高山族相似,证明这一群体是较纯的百越群体,是与良渚、马桥时期一脉相承的族系.由于黄浦江以西地区流动性更大,所以马桥地区虽然 7000 年来发生了许多大事件和社会文化变革,但是当地居民的血脉应是一贯传承下来的.

主成分分析表明百越系统分成两群,金汇、马桥人群和台湾原住民族聚成东越群体,西南地区的侗

傣民族聚成西越群体. 而图中西越群体与百濮的布朗、高棉关系也很密切. 这与史书上百濮和百越民族关系密切, 濮越难分的说法相符合. 而且我们也可以清楚地看到, 上海郊区人群与氏羌、马来民族的关系较远.

#### 4.4 Y 染色体单倍型在中国的迁徙路线

金力教授通过对东亚人群包括中国人群的大规模研究, 不仅得出了各 Y 染色体单倍型在地域上的分布, 更进一步总结了各单倍型在中国的迁徙路线. 在研究过程中, 我们也发现这个迁徙路线的正确性和实用性. 我们的研究结果都基本符合这个路线. 但对某些单倍型、在某些地区的迁徙路线还不是十分明确. 如图 3 所示. 这都有待于将来的研究来解决. 从而完善用 Y 染色体 SNP 方法研究人类起源和迁徙的方法.

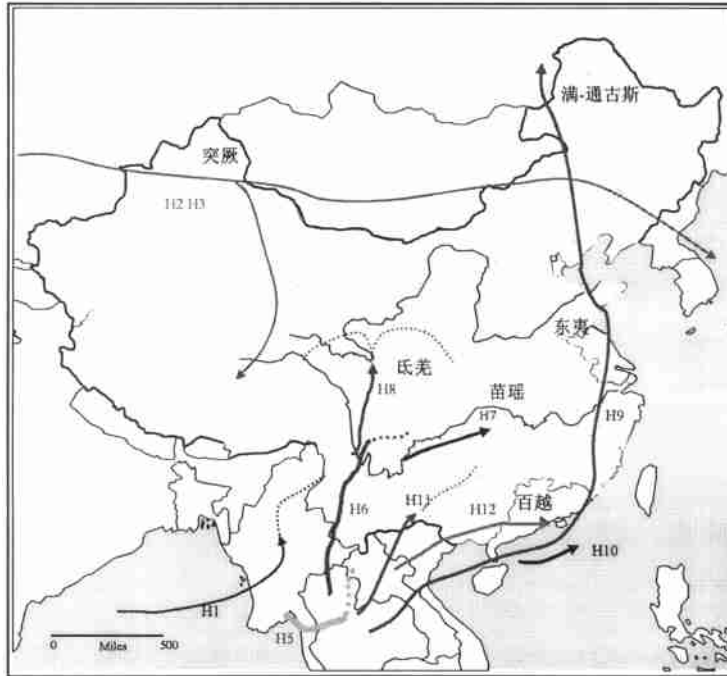


图 3 Y 染色体单倍型在中国的迁徙方向

Fig. 3 Migration routes of Y-CNP in China

#### 4.5 数据处理与主成分分析

由上海群体与南方各群体的主成分分析图及各主成分与单倍型的相关性可见, 第 1 主成分 (32.35%) 中 H17 起主要作用, 将大洋洲与大陆之间差异表现出来; 第 2 主成分 (18.56%) 以 H9 与 H11 为主, 区分东越和西越, 东越群体多 H9, 西越群体多 H11; 第 3 主成分 (13.89%) 中 H5 与 H11 其主要作用, 区别濮越和非濮越, 濮越多 H11, 非濮越多 H5; 第 4 主成分 (10.73%) 区分苗瑶、氏羌和非苗瑶, 苗瑶氏羌多 H8, H7, H1; 最后第 5 主成分 (8.62%) 将中国东部与西部区分开来, 东部略多 H1, H5, 西部略多 H12. 因此, 深入地分析这些 Y 染色体单倍型在各群体中的分布, 将在族源分析中起非常关键的作用, 能够提供给我们非常多的信息, 解决许多细节问题.

#### 参考文献:

- [ 1 ] 张思仲. 人类基因组的 SNP 及其医学应用[J]. 中华医学遗传学杂志, 1999, 16(2): 119 - 121.
- [ 2 ] 吴昕彦, 张庆华, 陈竺. 单核苷酸多态性研究及应用[J]. 中华医学遗传学杂志, 2000, 17(1): 57 - 59.
- [ 3 ] CANN R L, STONEKING M, WILSON A C. Mitochondria DNA and human evolution[J]. Nature, 1987, 325: 31 - 36.

- [ 4 ] WILSON A C, CANN R L. The recent African genesis of humans[J]. *Sci. Am*, 1992, 266:68 - 73.
- [ 5 ] 理查德·利基著, 吴汝康等译, 人类的起源[M]. 上海: 上海科学技术出版社, 1995.
- [ 6 ] GIBBONS A Y. Y chromosome shows that Adam was African[J]. *Science*, 1997, 278(5399): 804 - 805.
- [ 7 ] JOBLING M A, TYLER-SMITH C. Fathers and sons: the Y chromosome and human evolution[J]. *Trends Genet*, 1995, 11: 449 - 455.
- [ 8 ] UNDERHILL P A, JIN L, *et al.* A Pre-Columbian Human Y Chromosome Specific Transition and its Implications for Human Evolution[J]. *PNAS USA*, 1996, 93: 196 - 200.
- [ 9 ] SU B, *et al.* Y-chromosome evidence for a northward migration of modern human into East Asia during the last ice age[J]. *Am. J. Hum. Genet*, 1999, 65: 1718 - 1724.
- [ 10 ] SU B, *et al.* Polynesian origins: insights from the Y chromosome[J]. *Proc. Natl. Acad. Sci. USA*, 2000, 97: 8225 - 8228.
- [ 11 ] SU BING, XIAO CHUNJIE, DEKA RANJAN, *et al.* Y chromosome haplotypes reveal prehistorical migrations to the Himalayas [J]. *Hum Genet*, 2000, 107: 582 - 590.
- [ 12 ] JIN LI, SU BING. Natives or immigrants: modern human origin in East Asia[J]. *Nature Reviews Genetics*, 2000, 1(2): 126 - 133.
- [ 13 ] KE YUEHAI, SU BING, SONG XIU FENG, *et al.* African Origin of Modern Humans in East Asia: A Tale of 12000 Y Chromosome[J]. *Science*, 2001, 292: 1151 - 1153.
- [ 14 ] HUNAN MEDICAL COLLEGE. Study of an ancient cadaver in Mawangtui tomb No. 1 of the Han Dynasty in Changsha[M]. New York: Ancient Memorial Press, 1981, 184 - 187.
- [ 15 ] PÄÄBO S, HIGUCHI R G, WILSON A C. Ancient DNA and the Polymerase Chain Reaction[J]. *Journal of Chemistry*, 1989, 264(17): 9709 - 9712.
- [ 16 ] 徐立群, 李辉, 奚慧峰, 等. 上海郊区姓氏和通婚分析[J]. *遗传学报*, 2002, 29(8): 666 - 673.

## Genetics Analysis of Shanghai Aborigine with SNPs on Y Chromosome

YANG Jun, LI Hui, JIN Jian-zhong, JIN Li, LU Da-ru

(Center for Anthropological Studies, School of Life Sciences, Fudan University, Shanghai, 200433, China)

**Abstract:** Y chromosome is another focus of human evolutionary studies after mitochondrial DNA, which has also approved the "out of Africa" hypothesis. The mutation rate of haplotypes of single nucleic polymorphisms (SNP) on Y chromosome matches that of ethnic evolvments well, so Y-SNPs become the most measurable genetic materials in ethnicity nowadays. Among hundreds of Y-SNPs found, M119C mutation is the characteristic of Daic, and haplotypes of M119C-M110C and M95T-M88G can hardly be found out of Daic and Austronesian. Maqiao Culture lies on but is quite different from Liangzhu Culture, and it is also a question whether the people of these two cultures are the same population. So Y-SNPs are tested in the populations of Liangzhu Period, Maqiao Period, Warring States Period, Ming Dynasty and modern population in Maqiao district, and high frequencies of M119C and M95T are found. The results show that people in Maqiao district are always Daic people genetically from 6000 year ago to now, never Hmong-Mien. These Shanghai aborigines self-proclaimed "Dangdaic" are most similar to Taiwan aborigines on genetic structure, and also fairly close to Kam-Tai peoples in Southwest China. Principle component analysis and correlation analysis between components and haplotypes shows that H9 and H11 haplotypes differentiates West and East Daic, H7 recognizes Hmong-Mien, H8 Sino-Tibetan. Y-SNP haplotypes divide the ethnic groups clearly, that shows these are good materials of ethnicity, and worth of further studies.

**Key words:** Y chromosome; SNPs; principle component; Daic; ancient DNA; Shanghai aborigine; Liangzhu