第8卷 第3期 人类学学报 Vol. VIII, No. 3
1989年8月 ACTA ANTHROPOLOGICA SINICA Aug., 1989

# 中国人群遗传结构分析

翁自力
（北京师范大学生物系）
袁义达 杜若甫
（中国科学院遗传研究所）

**关键词** 群体遗传结构；遗传距离；主坐标分析

**内 容 提 要**

本文根据红细胞血型基因频率，用 Harpending 和 Jenkins(1973) 方法计算了中国22个人群间的遗传距离，同时在国内首次运用主坐标分析及其排序方法展示了中华民族的遗传结构，反映出中国东西人群与南北人群间的基因流。

## 一、前 言

遗传结构指的是群体中各种基因的频率以及由于不同的交配体制所产生的各种基因型在数量上的分布。国内有关中国人群遗传结构的研究起步较晚，但目前已经积累了较多的红细胞血型分布资料，有可能从这些资料中揭示出中华民族的源和流。

八十年代以来，主坐标分析及其多维排序方法日渐成为人类遗传结构研究中重要的手段之一。在遗传距离的各种测度中，Harpending 和 Jenkins 提出的 $d^2$ 比较适于作为人类学研究的遗传距离测度。这些分析手段国内尚未见报道。本文使用该分析方法对中国人群的遗传结构作一初步分析。

## 二、材 料

资料取自国内外 21 篇文献及中国科学院遗传研究所人类群体遗传研究室待发表资料。收集了北方汉族（袁义达等，1982；袁义达等，1984c；郝露萍等，1986；贾静涛等，1987），宁夏回族（血型调查组，1980a、1980b、1981；袁义达等，1985a），内蒙古蒙古族（血型调查组，1980a、1980b、1981；袁义达等，1985 b；Yuan *et al.*, 1984）；新疆维吾尔族（血型调查组，1980a、1980b、1981；表义达等，1984a）；辽宁满族（艾琼华等，1985c），吉林延边朝鲜族（Yuan *et al.*, 1984），西藏藏族（艾琼华等，1985 b；兰炯采等，1987b；Mourant *et al.*, 1976），西南地区彝族（血型调查组，1980a、1980b、1981；艾琼华等，1985a；兰炯采等，1987a），广西侗族（袁义达等，1984b），云南白族（血型调查组，1980a、1980b、1981；金锋等，1984b），广西壮族（血型调查组，1980a、1980b、

1981; Yuan *et al.*, 1984),湖南湘西土家族 (金锋等, 1984 a),湖南湘西苗族(金锋等, 1985) 以及新疆乌兹别克族,柯尔克孜族,塔塔尔族,哈萨克族, 锡伯族,云南苗族, 景颇族,傣族,佤族 (血型调查组, 1980a、1980b、1981) 共二十二个人群红细胞血型基因频率(表 1、2)。 其中汉族人群除 MNSs 系统单倍型频率采用了辽宁汉族人群的数据外,其它各红细胞血型系统或遗传位点上的基因频率或单倍型频率(包括 MN 血型位点及 S 血型位点) 均采用华北汉族人群的数据。当一个人群的一个血型系统有两篇以上报道时一般均取其基因频率的算术平均值。

**表 1 十一个人口均在一百万以上的民族红细胞血型系统基因频率**(或单倍型频率, $\times 10^{-4}$)

| 血型系统 | 等位基因(单倍型) | 维吾尔 | 满 | 蒙古 | 回 | 汉 | 朝鲜 | 藏 | 彝 | 白 | 侗 | 壮 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ABO | A | 2216 | 1647 | 1979 | 2106 | 2076 | 2374 | 1296 | 2152 | 2230 | 2042 | 1478 |
| | B | 2376 | 2774 | 2334 | 2290 | 2135 | 2193 | 2464 | 2114 | 2248 | 1672 | 1922 |
| MNS | MS | 1580 | 0087 | 0547 | 0281 | 0434 | 0181 | 0368 | 0557 | 0448 | 0000 | 0118 |
| | Ms | 3970 | 5842 | 4957 | 4422 | 4442 | 4902 | 5931 | 6419 | 6034 | 6294 | 7275 |
| | NS | 0390 | 0246 | 0112 | 0313 | 0310 | 0040 | 0237 | 0157 | 0054 | 0124 | 0000 |
| Rhesus | cde | 1610 | 0000 | 0452 | 0654 | 0000 | 0000 | 0224 | 0000 | 0243 | 0000 | 0252 |
| | Cde | 0128 | 0980 | 0152 | 0070 | 0000 | 0000 | 0604 | 0166 | 0164 | 0000 | 0124 |
| | cdE | 0192 | 0000 | 0000 | 0000 | 0545 | 0680 | 0104 | 0000 | 0000 | 0000 | 0000 |
| | cDe | 0542 | 0601 | 0558 | 0554 | 0638 | 1214 | 0978 | 0550 | 0594 | 0843 | 0312 |
| | CDe | 5239 | 6547 | 5904 | 6036 | 6475 | 6424 | 4666 | 7056 | 7118 | 7532 | 7871 |
| | cDE | 2138 | 1823 | 2874 | 2390 | 2210 | 0953 | 2716 | 2048 | 1193 | 1396 | 1351 |
| | CDE | 0172 | 0215 | 0060 | 0296 | 0132 | 0729 | 0710 | 0178 | 0188 | 0229 | 0090 |
| P | $P_1$ | 3830 | 1577 | 2374 | 2246 | 2132 | 1420 | 2138 | 2849 | 1640 | 1333 | 1134 |
| Duffy | $Fy^a$ | 8982 | 9405 | 9225 | 9270 | 9330 | 9398 | 9236 | 9434 | 9275 | 9651 | 9810 |
| Kidd | $Jk^a$ | 3140 | 4262 | 4094 | 4428 | 3487 | 4306 | 4447 | 4722 | 4700 | 4628 | 3649 |
| Diego | $Di^a$ | 0392 | 0168 | 0342 | 0349 | 0402 | 0402 | 0545 | 0917 | 0101 | 0226 | 0436 |
| ABH | Se | 5192 | 4457 | 5553 | 5095 | 4761 | 4908 | 5860 | 4824 | 6479 | 4865 | 4734 |

**表 2 十一个少数民族人群红细胞血型系统基因频率**(或单倍型频率, $\times 10^{-4}$)

| 血型系统 | 等位基因(单倍型) | 乌兹别克 | 柯尔克孜 | 塔塔尔 | 哈萨克 | 锡伯 | 土家 | 苗(湖南) | 苗(云南) | 景颇 | 傣 | 佤 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ABO | A | 2044 | 1383 | 2561 | 1746 | 1978 | 2843 | 2061 | 1225 | 2370 | 1625 | 3079 |
| | B | 2855 | 2753 | 2197 | 2271 | 2988 | 1897 | 1722 | 2580 | 1478 | 2084 | 1967 |
| MN | M | 6318 | 6492 | 6622 | 6441 | 5800 | 5950 | 6213 | 7555 | 8084 | 7594 | 8308 |
| P | P1 | 3837 | 2223 | 4072 | 2893 | 2224 | 3036 | 1734 | 2659 | 1022 | 0811 | 2069 |
| Rhesus | cde | 2421 | 1292 | 2638 | 1298 | 0383 | 0000 | 0000 | 0000 | 0000 | 0000 | 0000 |
| | Cde | 4330 | 0181 | 1163 | 0000 | 0000 | 0707 | 0000 | 0000 | 0707 | 0889 | 0000 |
| | cdE | 0000 | 0000 | 0000 | 0000 | 0000 | 0000 | 0000 | 0510 | 0000 | 0000 | 0000 |
| | cDe | 1684 | 4841 | 1670 | 0537 | 2652 | 0701 | 1415 | 0896 | 1694 | 0981 | 0734 |
| | CDe | 3380 | 3044 | 1962 | 5925 | 4607 | 6654 | 6356 | 6850 | 6034 | 6626 | 8102 |
| | cDE | 1722 | 0642 | 2271 | 2190 | 2079 | 0974 | 2150 | 1501 | 0942 | 1100 | 0977 |
| | CDE | 0390 | 0000 | 0296 | 0000 | 0279 | 0014 | 0079 | 0243 | 0623 | 0404 | 0187 |
| ABH | Se | 5019 | 6521 | 4547 | 5428 | 5755 | 4865 | 5958 | 6370 | 6544 | 4995 | 5402 |

# 三、方　法

## (一) 遗传距离 $d^2$

Harpending 和 Jenkins(1973) 提出的遗传距离 $d^2$ 的计算方法如下：

对于 S 个人类亚群，首先构建一个 S × S 的关系矩阵 (relation matric) $R$。其第 $i$ 行第 $j$ 列元素 $r_{ij}$ 可由下列公式求得：

$$r_{ij}=\frac{1}{m}\sum_{k=1}^{m}(P_{ik}-\bar{P}_k)(P_{jk}-\bar{P}_k)/\bar{P}_k(1-\bar{P}_k)$$

式中 $m$ 是各位点等位基因总数，$\bar{P}_k$ 是第 $k$ 个等位基因频率的平均值，$P_{ik}$ 和 $P_{jk}$ 分别是第 $i$ 个亚群和第 $j$ 个亚群的第 $k$ 个等位基因的频率。

由关系矩阵 $R$，可以构建遗传距离矩阵 $D$。第 $i$ 个亚群与第 $j$ 个亚群的遗传距离 $d_{ij}^2=r_{ii}+r_{jj}-2r_{ij}$。

由遗传距离阵 $D$ 也可反推出关系矩阵 $R$，其方法如下。

首先求出 $D$ 矩阵各行之和：

$$D_{i\cdot}=\sum_{j=1}^{S}d_{ij}^2(j=1,2,\cdots,S),$$

各列之和

$$D_{\cdot j}=\sum_{j=1}^{S}d_{ij}^2(i=1,2,\cdots,S),$$

以及 $D$ 中全部元素之和

$$D_{\cdot\cdot}=\sum_{i=1}^{S}\sum_{j=1}^{S}d_{ij}^2$$

则

$$r_{ij}=\frac{1}{2}d_{ij}^2+\frac{1}{2S}D_{i\cdot}+\frac{1}{2S}D_{\cdot j}-\frac{1}{2S^2}D_{\cdot\cdot}$$
$$(i,j=1,2,\cdots,S)$$

## (二) 主坐标分析

构建关系矩阵 $R$ 后，进一步用主坐标分析的方法将群体间的遗传距离关系在二维平面或三维平面上近似地表示出来。其步骤是先求出 $R$ 矩阵的 $S$ 个特征根，将它们从大到小排列，依次为 $\lambda_1、\lambda_2、\cdots、\lambda_S$，并分别求出它们的特征根向量，依次作为行排成正交矩阵 $E$。矩阵 $E$ 的每一行元素乘以它们相应的特征根的平方根得到矩阵 $C$，$C$ 矩阵中第 $i$ 行第 $j$ 列元素 $c_{ij}=e_{ij}\lambda_i^{\frac{1}{2}}$，式中 $e_{ij}$ 是矩阵 $E$ 第 $i$ 行第 $j$ 列元素。第 $i$ 个主坐标上保存的变异信息量为 $\lambda_i\Big/\sum_{i=1}^{S}\lambda_i$。选取前两个主坐标即可绘出这些群体的主坐标分析二维排序图。当信息量较集中时，各群体在排序图上的距离的平方近似地等于它们之间的遗传距

离 $d^2$(Harpending 和 Jenkins, 1973; 阳含熙、卢泽愚, 1981)。

主坐标分析方法不同于目前国内人类学研究中经常使用的主成分分析 (Principal components analysis) 不同。例如在我们研究 $S$ 个人群的关系时，主成分分析方法当求调查这 $S$ 个人群的 $m$ 项属性，然后从这样得到的 $S\times m$ 的原始数据矩阵出发对这些人群进行分析。而主坐标分析则只需从这 $S$ 个人群的 $S\times S$ 的相异矩阵(对于人类群体遗传学来讲也就是遗传距离矩阵) 出发就可对它们进行排序。无论我们在衡量人群间遗传差异时使用哪一种遗传距离测度，都可由遗传距离矩阵推算出关系矩阵，进而依据前述各步骤作出人群的主坐标分析二维或三维排序图。

## (三) 基因的排序

在完成了人群的排序之后，还可进一步将各基因在各主坐标轴上的排序关系表示出来。这样对照人群的主坐标分析排序图和与之相应的基因排序图，我们便可综合地了解各个基因频率在这些人群中大致的分布趋势。

第 $k$ 个等位基因在第 $i$ 个主坐标轴上的负荷量 (Loading) $L_{ki}$ 由下式给出:

$$L_{ki}=\sum_{j=1}^{S} Z_{jk}\cdot e_{ji}\text{(Jorde } et\ al.\text{, 1982)}$$

式中 $e_{ji}$ 是矩阵 $E$ 的元素，

$$Z_{jk}=(P_{jk}-\bar{P}_k)/\sqrt{\bar{P}_k(1-\bar{P}_k)}\ \text{(Harpending 和 Jenkins, 1973)}$$

当我们需要将基因的排序情况和人群的排序情况在同一坐标系中表示出来时，则第 $k$ 个基因在第 $i$ 个主坐标轴上的排序坐标 $M_{ki}$ 一般可由下列公式计算:

$$M_{ki}=L_{ki}\cdot\left(\lambda_i\Big/\sum_{i=1}^{m}L_{ki}^2\right)^{\frac{1}{2}}\ \text{(参考 Crawford and Enciso, 1982)}$$

式中 $\lambda_i$ 为第 $i$ 个主坐标所对应的特征根，$m$ 为各位点上等位基因的总数。

# 四、结 果

我们用 ABO、MNSs、Rhesus、P、Duffy、Kidd、Diego、ABH 等五个遗传系统上 25 个等位基因(或单倍型)频率，计算了我国十一个人口在一百万以上的民族间的遗传距离矩阵和关系矩阵(表 3)。然后用主坐标分析方法绘出这十一个民族人群的主坐标分析二维排序图(图 1)。第一主坐标(用 $e_1\lambda_1^{\frac{1}{2}}$表示)和第二主坐标(用 $e_2\lambda_2^{\frac{1}{2}}$表示)包含的信息量分别是 37.7% 和 19.2%，合计为 56.9%。我们还计算了各基因在第一和第二主坐标上的排序坐标，并将它们一一在图 1 上标出。

为了根据血型位点上的基因频率探讨更多人群间的遗传关系，我们又根据 ABO、MN、Rhesus、P、ABH 等 5 个血型系统上共 17 个等位基因(或单倍型)频率，计算了 22

表 3　十一个人群的遗传距离 $d^2$ 矩阵和关系矩阵*　($\times 10^{-3}$)

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 维吾尔 | 64 | −13 | 10 | 10 | 3 | −13 | −2 | −3 | −11 | −24 | −22 |
| 满 | 120 | 29 | −1 | −2 | −5 | −2 | 14 | −6 | −8 | −2 | −4 |
| 蒙古 | 51 | 40 | 8 | 5 | 1 | −6 | 0 | −2 | −3 | −5 | −7 |
| 回 | 52 | 41 | 53 | 8 | 1 | −2 | −2 | −4 | −4 | −3 | −8 |
| 汉 | 73 | 52 | 19 | 19 | 13 | 9 | −6 | −3 | −8 | −2 | −4 |
| 朝鲜 | 114 | 57 | 43 | 36 | 19 | 24 | −4 | −6 | −2 | 5 | −2 |
| 藏 | 84 | 18 | 23 | 28 | 41 | 47 | 16 | 0 | −4 | −6 | −7 |
| 彝 | 87 | 58 | 29 | 33 | 37 | 53 | 34 | 18 | −3 | 2 | 5 |
| 白 | 119 | 78 | 46 | 50 | 62 | 61 | 58 | 57 | 33 | 5 | 4 |
| 侗 | 129 | 50 | 35 | 31 | 34 | 31 | 45 | 30 | 39 | 16 | 16 |
| 壮 | 136 | 66 | 50 | 52 | 50 | 55 | 57 | 35 | 53 | 13 | 28 |
| | 维吾尔 | 满 | 蒙古 | 回 | 汉 | 朝鲜 | 藏 | 彝 | 白 | 侗 | 壮 |

* 左下方三角型内为遗传距离阵，右上方三角形内为关系矩阵。

个人群间的遗传距离 $d^2$ (限于篇幅未列出)进而进行主坐标分析并绘出二维排序图 (图 2)。前两个主坐标上共保留信息量 56.1%。

## 五、讨　论

可以看出，图 1 所展示的各个人群的遗传关系同它们的地理位置关系有一定的联系。

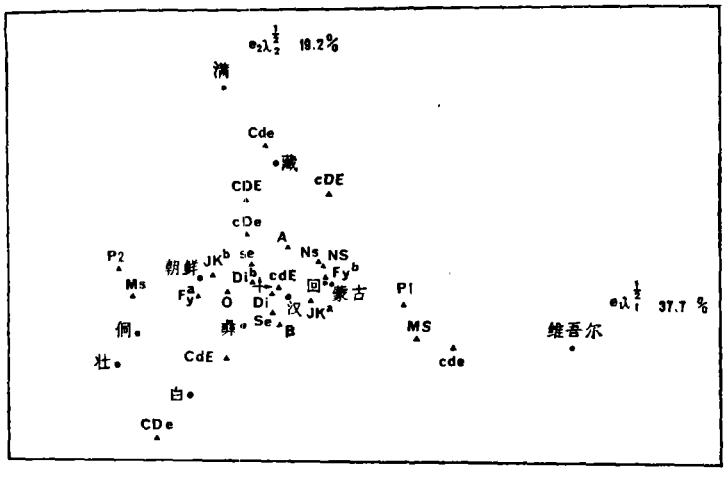


图 1　十一个人群主坐标分析二维排序图

11 Chines ethnic groups plotted on first two scaled eigenvector

北方人群在二维排序图上的位置偏于右方及上方，其中蒙古族、回族、汉族间的遗传

关系较近。维吾尔族与其它人群关系较远，其排序坐标位于图的最右端。南方人群在坐标系上的分布偏于左下方，其中壮族、侗族和白族间相互关系比较近，而彝族比较接近于北方人群，这同彝族原来是从中国西北迁至中国南部的历史背景有关。

对照基因排序情况可以看出，总得说来，北方人群中在地理位置上最靠西北的维吾尔族的 cde、MS 等单倍型及 $P_1$ 基因频率较高，满族及藏族则 Cde、CDE 和 cDE 单倍型及 B 基因频率相对较高。在南方人群中则上述单倍型频率较低，而 Ms、CDe 等单倍型频率较高。在图 1 中间的一些等位基因或单倍型，如 cdE、$Di^a$、Ns、NS、Se、O 等，在各民族中频率差异不很明显或变化不特别规则。对照表 1 中的数字可以看出，图 1 的确十分形象而准确地表示出不同民族这些单倍型及基因频率分布的综合特点。

图 2 所展示的各个人群间的遗传距离关系也同他们在地理分布上的位置关系相当吻合。

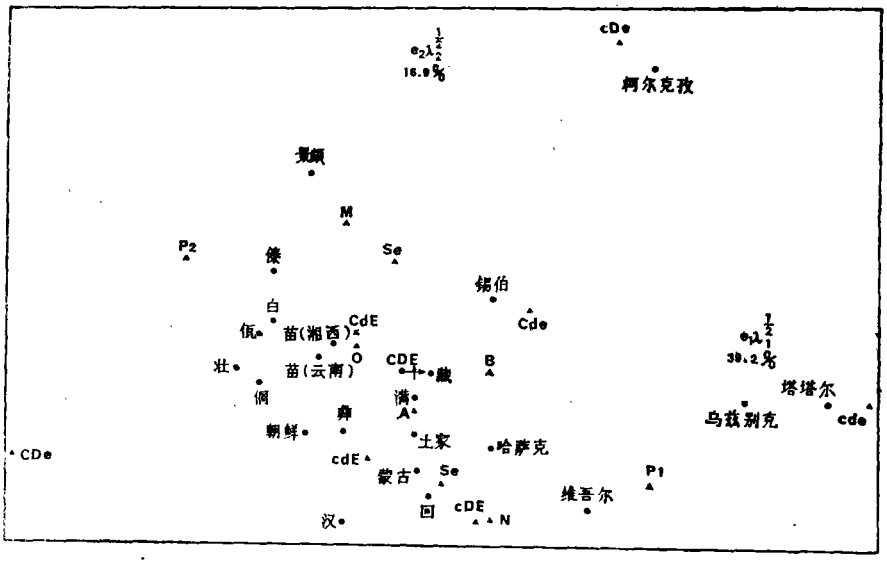


图 2 二十二个人群主坐标分析二维排序图

22 Chinese populations plotted on first two scaled eigenvector

可以看出，第一主坐标(即图中的横坐标)反映的主要是高加索人种和蒙古人种间的基因流，以高加索人种血缘较多的塔塔尔、乌兹别克和柯尔克孜族等为一极，在图的右边；而典型的蒙古人种人群，如朝鲜、汉、白、傣、壮、侗、佤等族则集中于另一极，在图的左边；维吾尔族、哈萨克族的位置比较居中。高加索人种血缘较多的民族在基因或单倍型频率方面的特点是 cde、cDe、$P_1$ 等的数值比较高，而典型的蒙古人种人群的特点则是 CDe、$P_2$、cdE 等的频率比较高。

而第二主坐标反映的则主要是我国北方人群和南方人群间的基因流。典型的南方人群，如傣族、佤族、壮族多集中于坐标系的左上方，它们与北方人群的关系较远。这些民族的一般特点是 $P_2$、M、Se、O、CdE 等基因或单倍型的频率较高。北方诸民族则多集中于坐标系的右下方，如汉族、回族、蒙古族、哈萨克族等，其特点主要是 cDE、N、Se 等

基因或单倍型的频率比较高。彝族和土家族在二维排序图上的位置显示出它们在遗传组成上比较接近于北方人群。据文献记载，彝族是古代北方的羌人南迁融合一部分西南土著人血缘而形成的，这与我们主坐标分析二维排序图上的情况吻合。

乌兹别克族、塔塔尔族和柯尔克孜族等三个民族在图 2 上的位置很特殊，表明它们与其它民族的血缘关系较远。结合基因排序坐标可以看出，柯尔克孜族 cDe 频率特别高，而乌兹别克族和塔塔尔族则是 cde 频率特别高。维吾尔族的位置比较特殊，介乎于乌兹别克、塔塔尔族人群与北方人群之间，其 $P_1$ 基因频率特别高。

本研究所涉及的遗传标记还仅限于少数几个红细胞血型位点，由主坐标二维排序图反映出的各人群的关系还比较粗糙。为了进一步研究国内各人群复杂的遗传结构，还需增加更多的遗传资料方能作出比较正确的分析，得出更为精确的结论。

## 参　考　文　献


艾琼华、袁义达、杜若甫，1985a。彝族的红细胞血型分布。中国科学院遗传研究所研究工作年报，80 页。

艾琼华、赵红、战文惠、杜若甫，1985b。藏族中红细胞血型系统的分布。中国科学院遗传研究所研究工作年报，82 页。

艾琼华、李实喆、杜若甫，1985c。满族的红细胞血型调查。中国科学院遗传研究所研究工作年报，83 页。

兰炯采、杨士英、郑世荣、李维根、范伯澄、文素华、蔡志英，1987a。四川地区彝族 ABO、Duffy、Lewis、Kidd 和 Diego 血型系统的分布。中华血液学杂志，**8**：93。

兰炯采、杨世英、郑世荣、达娃扎西，1987b。藏族 ABO、Duffy、Lutheran、Lewis、Diego 和 Xg 血型分布。中华血液学杂志，**8**：487—489。

阳含熙、卢泽愚，1981。植物生态学的数量分类方法。科学出版社。262—269 页。

血型调查组，1980a。我国十六个民族的血型调查报告，I ABO 血型及 ABH 分泌型调查结果。中华血液学杂志，**1**：261—263。

血型调查组，1980b。我国十六个民族的血型调查报告，II MN 及 P 血型调查结果。中华血液学杂志，**1**：352—356。

血型调查组，1981。我国十六个民族的血型调查报告，III Rh 血型调查结果。**2**：209—211。

金锋、赵红、杜若甫，1984a。湘西土家族四种红细胞血型系统的表型分布。中国科学院遗传研究所研究工作年报，110 页。

金锋、郝露萍、杜若甫。1984b。白族九种红细胞血型系统的表型分布。中国科学院遗传研究所研究工作年报，111 页。

金锋、赵红、杜若甫，1985。湘西苗族五种红细胞血型系统的分布。中国科学院遗传研究所研究工作年报，75 页。

郝露萍，1986。北京地区人群中六种红细胞抗原分布的调查。中华血液学杂志，**7**：63—69。

贾静涛、王秀玲，1987。辽宁地区汉族 MNSs 血型分布。中华血液学杂志，**8**：99。

袁义达、徐玖瑾、张志、杜若甫，1982。华北汉族 Kell、Kidd、Diego、Duffy、Lutheran 和 Xg 血型系统的分布。遗传学报，**9**：395—401。

袁义达、乌云、艾绍堂、金锋、杜若甫，1984a。新疆维吾尔族的红细胞血型系统的研究。中华血液学杂志，**5**：305—309。

袁义达、金锋、杜若甫、龙邕泉、蔡瑞霖，1984b。侗族九个红细胞血型系统和 ABH 分泌型的分布。人类学学报，**3**：277—284。

袁义达、郝露萍、杜若甫，1984c。华北地区汉族的 Lewis、ABO、MN、Rh、P 等血型系统和 ABH 分泌型的分布。人类学学报，**3**：181—187。

袁义达、杜若甫、李长清，1985a。宁夏回族红细胞血型的研究。人类学学报，**4**：385—393。

袁义达、金锋、赵红、杜若甫，1985b。蒙古族 ABO、Lewis、Diego 和 Xg 血型的分布。中华血液学杂志，**6**：523—525。

Crawford, M. H. and V. B. Enciso, 1982. "Population structure of circumpolar groups", in *Current development in anthropological genetics*, Vol. 2, Ed. M. H. Crawford and J. H. Mielke, pp. 51—91, Plenum, New York.

Harpending, H. and T. Jenkins, 1973. "Genetic distance among southern African populations", in *Methods and theories of anthropological genetics*, Ed. M. H. Crawford and P. L. Workman, pp. 179—199, Univ. of New Mexico Press, Albuquerque.

Jorde, L. B., P. L. Workman and A. W. Eriksson, 1982. "Genetic microevolution in the Aland Island, Finland", in *Current development in anthropological genetics*, Vol. 2, Ed. M. H. Crawford and J. H. Mielke, pp. 333—356, Plenum, New York.

Mourant, A. E., A. C. Copee, And K. Domanieskasabczak, 1976. *The distribution of the human blood groups and other biochemical polymorphisms*, 2nd ed., Oxford Univ. Press, Oxford.

Yuan, Y., R. Du, L. Chen, J. Xu, Y. Wang, S. Li, H. G. Benkmann, P. Bogdanski, G. Kriese, and H. W. Goedde, 1984. Distribution of eight blood group systems and ABH secretion of Mongolian, Korean, Zhuang nationalities in China. *Annals of Hum. Biol.*, **11**: 377—388.

# ANALYSIS ON GENETIC STRUCTURE OF HUMAN POPULATIONS IN CHINA

Weng Zili

*(Biology Department, Beijing Normal University)*

Yuan Yida Du Rofu

*(Institute of Genetics, Academia Sinica)*

**Key words** Genetic structure of population; Genetic distance; Principal coordinate analysis

## Abstract

The genetic structure of 22 ethnic groups in China was analysed by using gene frequency data of red cell blood groups. The Harpending and Jenkins' (1973) topological methods of representing population structure were applied to the study on relationship between these ethnic groups. The plots of these ethnic groups showed clearly the gene flow between Caucasian and Mongoloid and between northern Mongoloid populations and southern ones.