Human Genome Diversity Research in China

Chu Jiayou,
Institute of Medical Biology,
Chinese Academy of Medical Sciences,
Kunming 650118
There are 56 nationalities in China, each of them having independent inhabitation areas. And some of them are genetic isolated population. In the source of nationalities and genetic phenotypes, each nationality has its unique characteristics. There are significant differences in categories, enzyme system, HLA antigen and incidences of some genetic diseases.
## Chinese populations

<table>
<thead>
<tr>
<th>Nationalities</th>
<th>Population</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Achang</strong></td>
<td>27,708</td>
</tr>
<tr>
<td><strong>Bai</strong></td>
<td>1,594,827</td>
</tr>
<tr>
<td>Baoan</td>
<td>12,212</td>
</tr>
<tr>
<td><strong>Blang</strong></td>
<td>82,280</td>
</tr>
<tr>
<td><em>Bouye</em></td>
<td>2,545,059</td>
</tr>
<tr>
<td><strong>Dai</strong></td>
<td>1,025,128</td>
</tr>
<tr>
<td>Daur</td>
<td>121,357</td>
</tr>
<tr>
<td><strong>De’ang</strong></td>
<td>15,462</td>
</tr>
<tr>
<td>Dong</td>
<td>2,514,014</td>
</tr>
<tr>
<td>Dongxiang</td>
<td>373,872</td>
</tr>
<tr>
<td><strong>Dulong</strong></td>
<td>5,816</td>
</tr>
<tr>
<td>Ewenki</td>
<td>26,315</td>
</tr>
<tr>
<td>Gaoshan</td>
<td>400,000</td>
</tr>
<tr>
<td>Gelao</td>
<td>437,997</td>
</tr>
<tr>
<td><em>Han</em></td>
<td>1.2 billion</td>
</tr>
<tr>
<td><strong>Hani</strong></td>
<td>1,253,952</td>
</tr>
<tr>
<td>Hezhe</td>
<td>4,245</td>
</tr>
<tr>
<td><em>Hui</em></td>
<td>8,602,978</td>
</tr>
<tr>
<td>Jing</td>
<td>18,915</td>
</tr>
<tr>
<td><strong>Jingpo</strong></td>
<td>119,209</td>
</tr>
<tr>
<td><strong>Jino</strong></td>
<td>18,000</td>
</tr>
<tr>
<td>Kazak</td>
<td>1,111,718</td>
</tr>
<tr>
<td>Kirgiz</td>
<td>141,549</td>
</tr>
<tr>
<td>Korean</td>
<td>1,920,597</td>
</tr>
<tr>
<td><strong>Lahu</strong></td>
<td>411,476</td>
</tr>
<tr>
<td>Lhoba</td>
<td>2,312</td>
</tr>
<tr>
<td>Li</td>
<td>1.110,000</td>
</tr>
<tr>
<td><strong>Lisu</strong></td>
<td>574,856</td>
</tr>
<tr>
<td><em>Manchu</em></td>
<td>9,821,180</td>
</tr>
<tr>
<td>Maonan</td>
<td>71,968</td>
</tr>
<tr>
<td><em>Miao</em></td>
<td>7,398,035</td>
</tr>
<tr>
<td>Moinba</td>
<td>7,475</td>
</tr>
<tr>
<td><em>Mongolian</em></td>
<td>4,806,849</td>
</tr>
<tr>
<td>Mulao</td>
<td>159,328</td>
</tr>
<tr>
<td><strong>Naxi</strong></td>
<td>278,009</td>
</tr>
<tr>
<td><strong>Nu</strong></td>
<td>21,723</td>
</tr>
<tr>
<td>Oroqen</td>
<td>6,965</td>
</tr>
<tr>
<td><strong>Pumi</strong></td>
<td>29,657</td>
</tr>
<tr>
<td>Qiang</td>
<td>198,252</td>
</tr>
<tr>
<td>Russian</td>
<td>13,504</td>
</tr>
<tr>
<td>Sala</td>
<td>87,697</td>
</tr>
<tr>
<td>She</td>
<td>630,378</td>
</tr>
<tr>
<td><em>Shui</em></td>
<td>345,993</td>
</tr>
<tr>
<td>Tajik</td>
<td>33,583</td>
</tr>
<tr>
<td>Tatar</td>
<td>4,873</td>
</tr>
<tr>
<td><em>Tibetan</em></td>
<td>4,593,330</td>
</tr>
<tr>
<td>Tu</td>
<td>191,624</td>
</tr>
<tr>
<td>Tuja</td>
<td>5,704,223</td>
</tr>
<tr>
<td>Uygur</td>
<td>7,214,431</td>
</tr>
<tr>
<td>Uzbek</td>
<td>14,502</td>
</tr>
<tr>
<td><strong>Wa</strong></td>
<td>351,974</td>
</tr>
<tr>
<td>Xibe</td>
<td>172,847</td>
</tr>
<tr>
<td><em>Yao</em></td>
<td>2,134,013</td>
</tr>
<tr>
<td><em>Yi</em></td>
<td>6,572,173</td>
</tr>
<tr>
<td>Yugur</td>
<td>10,569</td>
</tr>
<tr>
<td><em>Zhuang</em></td>
<td>15,489,630</td>
</tr>
</tbody>
</table>

* and ** distributed in Yunnan province.
** only resided in Yunnan province
Biodiversity is so important, human being should not forget it’s own diversity.

Prof. Min Wu, A founder of Chinese Human Genome Project
Why human genome diversity research in China is focused in Yunnan province?

- Among 56 nationalities of China, there are 26 nationalities with populations of more than 5000 living together in Yunnan province. Among 26 nationalities, 15 are only living in Yunnan province: Bai, Hani, Dai, Lisu, Lahu, Wa, Naxi, Jingpo, Bulang, Pumi, Nu, Achang, Deang, Jino and Dulong.
Geographical Distribution of Sino-Tibetan Populations
Why are so many nationalities in Yunnan province?
1. Specific historic reasons make many nationalities migrating to Yunnan.

During the 5th century B.C. to 3rd century B.C., Xiongnu people in Mongolian Plain became stronger gradually and began to expand its territory to Qinhai and Gansu province. As a result, two ancient nationalities of Di and Qiang were expelled to the southwest China. In addition, ancient wars made some nationalities migrate to Yunnan to seek asylum. From then on, these nationalities settled down in Yunnan province.
2. Complicated geography and climate of Yunnan.

The complicated geographic environment, climate and rich natural resources in Yunnan province just fit for those different nationalities’ living customs. Therefore, Yunnan becomes an ideal place for many nationalities to settle down.
3. **Yunnan is isolated from outside for a period of time.**

Yunnan province is isolated from outside because of long-term feudal rule, regional domination and inaccessibility in transportation. Furthermore, separating from outside, backward mode of production and lower needs for living made many nationalities keeping in a independent and isolated state.
At present, more and more youths are married each other among different nationalities, genome of some nationalities face the danger of extinction. Therefore, the reservation of different nationalities genome become an urgent project. If measures have not been taken from now on, we will lose the opportunity forever.
Three focuses of our study

1. Reservation of different Chinese populations’ genome (cell bank and genome DNA bank).
2. Research on genetic diversity.
Co-operation groups

- Institute of Medical Biology, Chinese Academy of Medical Sciences, Kunming
- Harbin Medical University, Harbin
- Institute of Genetics, Chinese Academy of Sciences, Beijing
- Chinese Western Medical College, Chengdu
- Fudan University, Shanghai
Examples of our Current Work 1

- Sample collection and cell line Establishment
For this study, consulting with nationality scholars, blood samples of healthy individuals with no relative tie belong to different populations were collected at ethnic groups’ living areas, and 3 generations are investigated and recalled. This can guarantee the representative of samples.
More attentions will be paid to collect some samples from special genetic disease cases and cases with special phenotypes. To sample large families with specific characteristics of the ethnic group is also in the plan.
- Based on the principle of “informed consent”

- Giving medical consulting service for local people
人类基因组多样性计划
研究对象个人情况登记表

<table>
<thead>
<tr>
<th>序号</th>
<th>内容</th>
<th>备注</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>地名 中甸</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>取样日期</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>研究对象序号 KM95001 - 2A</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>研究者</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>研究对象姓名 专玛拉姆</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>曾用名、别名和头衔</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>年龄（周岁） 40A</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>性别 男</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>出生地点 迪庆州中甸县兰村</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>部族、支系、等级或民族 藏族</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>母亲（生母）姓名 专玛拉姆</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>母亲序号 ZA001</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>母亲出生地点 迪庆州中甸县</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>母亲民族 藏族</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>父亲（生父）姓名 旺堆</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>父亲序号 ZA002</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>父亲出生地点 迪庆州中甸县</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>父亲民族 藏族</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>配偶姓名 仓决</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>配偶序号 ZA001</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>配偶出生地点 迪庆州中甸县</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>配偶民族 藏族</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>研究对象家族或家系名称</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>宗族</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>在亲族及团体（胞族、血统、武士社团或妇女俱乐部等）内的其它成员身份</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>研究对象第一语言 藏语</td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>其它会说或能听懂的语言 汉语</td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>注释</td>
<td></td>
</tr>
</tbody>
</table>
DNA were extracted from blood samples and qualified. Meanwhile, use EB virus transformation technique to establish the genetic specific immortalized lymphocyte lines of different nationalities.
<table>
<thead>
<tr>
<th>Ethnic groups</th>
<th>Sampling place</th>
<th>cell line</th>
<th>DNA Samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sala</td>
<td>Xunhua County, Qinghai Province</td>
<td>71</td>
<td>100</td>
</tr>
<tr>
<td>She</td>
<td>Fuan City, Fujian Province</td>
<td>49</td>
<td>100</td>
</tr>
<tr>
<td>Tu</td>
<td>Huzhu County, Qinghai Province</td>
<td>39</td>
<td>100</td>
</tr>
<tr>
<td>Naxi</td>
<td>Lijiang City, Yunnan Province</td>
<td>36</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>Jinghong City, Xishuangbanna Autonomous Region &amp; Luxi county,</td>
<td>62</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>Dehong, Dai and Jingpo Autonomous Region, Yunnan Province</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dai</td>
<td>Luxi County, Dehong, Dai and Jingpo Autonomous Region, Yunnan Province</td>
<td>49</td>
<td>100</td>
</tr>
<tr>
<td>Jingpo</td>
<td>Luxi County, Dehong, Dai and Jingpo Autonomous Region, Yunnan Province</td>
<td>55</td>
<td>100</td>
</tr>
<tr>
<td>De’ang</td>
<td>Luxi County, Dehong, Dai and Jingpo Autonomous Region, Yunnan Province</td>
<td>51</td>
<td>100</td>
</tr>
<tr>
<td>Tibetan</td>
<td>Lhasa City, Tibet Autonomous Region</td>
<td>55</td>
<td>100</td>
</tr>
<tr>
<td>Zhuang</td>
<td>Baise County, Guangxi Zhuang Autonomous Region</td>
<td>55</td>
<td>100</td>
</tr>
<tr>
<td>Bulang</td>
<td>Luxi County, Dehong, Dai and Jingpo Autonomous Region, Yunnan Province</td>
<td>45</td>
<td>100</td>
</tr>
<tr>
<td>Li</td>
<td>Baisha County, Hainan Province</td>
<td>57</td>
<td>100</td>
</tr>
<tr>
<td>Moso</td>
<td>Yongning Xiang, Ninglang County, Yunnan Province</td>
<td>43</td>
<td>100</td>
</tr>
<tr>
<td>Bai</td>
<td>Wangqiao Xiang, Dali, Yunnan Province</td>
<td>61</td>
<td>100</td>
</tr>
<tr>
<td>Yi</td>
<td>Paomaping Xiang, Ninglang County, Yunnan Province</td>
<td>47</td>
<td>100</td>
</tr>
<tr>
<td>Han-GD</td>
<td>Shaoguan City, Guangdong Province</td>
<td>61</td>
<td>100</td>
</tr>
<tr>
<td>Hui</td>
<td>Najjiaying Town, Tonghai County, Yunnan Province</td>
<td>50</td>
<td>100</td>
</tr>
<tr>
<td>Ethnic groups</td>
<td>Sampling place</td>
<td>cell line</td>
<td>DNA Samples</td>
</tr>
<tr>
<td>---------------</td>
<td>------------------------------</td>
<td>-----------</td>
<td>-------------</td>
</tr>
<tr>
<td>Han_SD</td>
<td>Zhouping County, Shandong</td>
<td>61</td>
<td>100</td>
</tr>
<tr>
<td>Lisu</td>
<td>Fugong County, Nujiang, Yunnan Province</td>
<td>56</td>
<td>100</td>
</tr>
<tr>
<td>Nu</td>
<td>Fugong County, Nujiang, Yunnan Province</td>
<td>57</td>
<td>100</td>
</tr>
<tr>
<td>Dai-Huayao</td>
<td>Xinping County, Yunnan Province</td>
<td>5</td>
<td>100</td>
</tr>
<tr>
<td>Oroqen</td>
<td>Ali river Oroqen Autonomous County, Nei Mongol Province</td>
<td>39</td>
<td>120</td>
</tr>
<tr>
<td>Ewenki</td>
<td>Na river County, Heilongjiang Province</td>
<td>43</td>
<td>120</td>
</tr>
<tr>
<td>Daur</td>
<td>Hongyan Daur Autonomous county, Nei Mongol Province</td>
<td>58</td>
<td>120</td>
</tr>
<tr>
<td>Mongol</td>
<td>Haila’er City, Nei Mongol Province</td>
<td>62</td>
<td>120</td>
</tr>
<tr>
<td>Koreans</td>
<td>Yanbian Koreans Autonomous Region, Jilin Province</td>
<td>58</td>
<td>120</td>
</tr>
<tr>
<td>Hezhen</td>
<td>Tongjiang County, Heilongjiang Province</td>
<td>18</td>
<td>120</td>
</tr>
<tr>
<td>Man</td>
<td>Youyian Man Autonomous county, Liaoning Province</td>
<td>65</td>
<td>120</td>
</tr>
<tr>
<td>Xibe</td>
<td>Chabu’er Xibe Autonomous County, Xinjiang Province</td>
<td>43</td>
<td>120</td>
</tr>
<tr>
<td>Han-Sichuan</td>
<td>Leyu County, Sichuan Province</td>
<td>51</td>
<td>120</td>
</tr>
<tr>
<td>Han_Fujian</td>
<td>Minqing County, Fujian Province</td>
<td>58</td>
<td>120</td>
</tr>
<tr>
<td>Ethnic groups</td>
<td>Sampling place</td>
<td>cell line</td>
<td>DNA Samples</td>
</tr>
<tr>
<td>---------------</td>
<td>----------------------------------------------------</td>
<td>-----------</td>
<td>-------------</td>
</tr>
<tr>
<td>Hui</td>
<td>Tongxin County, Ningxia Hui Autonomous Region, Xinjiang Province</td>
<td>63</td>
<td>120</td>
</tr>
<tr>
<td>Kazak</td>
<td>Yinin Kazak Autonomous Region</td>
<td>64</td>
<td>120</td>
</tr>
<tr>
<td>Bouyei</td>
<td>Qiannan Bouyei Autonomous Region, Guizhou Province</td>
<td>67</td>
<td>120</td>
</tr>
<tr>
<td>Dong</td>
<td>Sanjiang County, Guangxi Province</td>
<td>50</td>
<td>100</td>
</tr>
<tr>
<td>Qiang</td>
<td>Mao County, Sichuang Province</td>
<td>48</td>
<td>100</td>
</tr>
<tr>
<td>Kirgiz</td>
<td>Wuqia County, Xinjiang Province</td>
<td>57</td>
<td>100</td>
</tr>
<tr>
<td>Han-Hunan</td>
<td>Shuangfeng County, Hunan Province</td>
<td>59</td>
<td>100</td>
</tr>
<tr>
<td>Han-Shanxi</td>
<td>Linyou County, Shanxi Province</td>
<td>58</td>
<td>100</td>
</tr>
<tr>
<td>Han-Guangdong</td>
<td>Guangning County, Guangdong Province</td>
<td>55</td>
<td>100</td>
</tr>
<tr>
<td>Yi</td>
<td>Butuo County, Sichuan Province</td>
<td>49</td>
<td>100</td>
</tr>
<tr>
<td>Miao</td>
<td>Songtao County, Guizhou Province</td>
<td>43</td>
<td>100</td>
</tr>
<tr>
<td>Tujia</td>
<td>Laifeng County, Hubei Province</td>
<td>56</td>
<td>100</td>
</tr>
<tr>
<td>Han-Zhejiang</td>
<td>Jinyun County, Zhejiang Province</td>
<td>59</td>
<td>100</td>
</tr>
<tr>
<td>Han-Henan</td>
<td>Fangcheng County, Henan Province</td>
<td>60</td>
<td>100</td>
</tr>
<tr>
<td>Uygur</td>
<td>Yining City, Xinjiang Province</td>
<td>60</td>
<td>100</td>
</tr>
<tr>
<td>Yao</td>
<td>Du’an County, Guangxi Province</td>
<td>58</td>
<td>100</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td><strong>2411</strong></td>
<td><strong>4860</strong></td>
</tr>
</tbody>
</table>
A Human Genome Diversity Cell Line Panel

A RESOURCE OF 1064 CULTURED
haploid/derived cell lines (LCL) from
individuals from diverse world populations
and corresponding nucleotide sequences of
DNA is deposited at the Foundation for the
Harding Human Genome Project (HHGP)
and CEPII in Périgueux, France. LCLs were
selected from various laboratories by the
Humangenomproject (HGP) and CEPII to
provide nucleotide supplies of DNA for
studies of sequence diversity and history of
diverse human populations. Information for each
LCL is included in a project-specific database and
includes a CEPH-generated, numeric identifier
that labels the LCL, its corresponding DNA, and
advice information. Every product derived from
an LCL is labeled with a unique identifier;
therefore, LCLs were tested for genetic variation by a
gene-specific, high-throughput method. All data
shows that genetic variation was discerned.
Examples of our Current Work 2

Research on genetic diversity

STR
Y Chromosome
mtDNA
SNP
Y Chromosome markers

- **SNPs** (Single Nucleotide Polymorphisms)
  - Mutation rate is low
  - Trace the ancient events

- **Microsatellites** (Tandem repeat of 2-5 nt)
  - Mutation rate is high
  - Trace the recent events
  - Estimate the age of MRCAs
Allelic Specific PCR in Multiplex of Five Y Chromosome SNPs

M   1      2      3      4      5      6      7      8

M89 (365bp)  
M110 (250bp) 
M50 (219bp)  
M7 (174bp)   
M119 (144bp)
## Y Chromosome Haplotype Frequency Distributions in East Asian Populations

<table>
<thead>
<tr>
<th>Continent</th>
<th>Population</th>
<th>Size</th>
<th>H1</th>
<th>H2</th>
<th>H3</th>
<th>H4</th>
<th>H5</th>
<th>H6</th>
<th>H7</th>
<th>H8</th>
<th>H9</th>
<th>H10</th>
<th>H11</th>
<th>H12</th>
<th>H13</th>
<th>H14</th>
<th>H15</th>
<th>H16</th>
<th>H17</th>
</tr>
</thead>
<tbody>
<tr>
<td>East Asia</td>
<td>Buryat (1)</td>
<td>4</td>
<td>75</td>
<td>25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ewenki (2)</td>
<td>8</td>
<td>50</td>
<td>13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Manchurian (3)</td>
<td>18</td>
<td>17</td>
<td>22</td>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mongolian (4)</td>
<td>24</td>
<td>28</td>
<td>13</td>
<td>4.2</td>
<td>8.3</td>
<td></td>
<td></td>
<td>13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Korean (5)</td>
<td>7</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Japanese (6)</td>
<td>29</td>
<td>21</td>
<td>28</td>
<td>11</td>
<td>21</td>
<td></td>
<td></td>
<td>28</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hui (7)</td>
<td>20</td>
<td>10</td>
<td>30</td>
<td>13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Tibetan (8)</td>
<td>8</td>
<td>13</td>
<td>13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hui (9)</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Tuja (10)</td>
<td>10</td>
<td>10</td>
<td>20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Yao-Nandan (11)</td>
<td>10</td>
<td>50</td>
<td>20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Yao-Jinxiu (12)</td>
<td>10</td>
<td>10</td>
<td>30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Zhuang (13)</td>
<td>29</td>
<td>3.6</td>
<td>3.6</td>
<td>7.1</td>
<td>3.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Dong (14)</td>
<td>10</td>
<td>20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bulang (15)</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Lahu (16)</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Yi (17)</td>
<td>14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>She (18)</td>
<td>11</td>
<td>9.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Atayal (19)</td>
<td>24</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Yami (20)</td>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Riwian (21)</td>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ami (22)</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Li (23)</td>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cambodian (24)</td>
<td>26</td>
<td>3.8</td>
<td>3.8</td>
<td>12</td>
<td>23</td>
<td>3.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>NE. Thai (25)</td>
<td>23</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Malaysian (26)</td>
<td>13</td>
<td>7.7</td>
<td>7.7</td>
<td>31</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Baliak (27)</td>
<td>18</td>
<td>5.6</td>
<td>5.6</td>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Javanese (28)</td>
<td>11</td>
<td>9.1</td>
<td>9.1</td>
<td>27</td>
<td>9.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
YAP+ Diversity

- **YAP+ (%)**
  - **Han**
    - Han(Guangdong) 0
    - Han(Shanxi) 0
    - Han(North) 0
    - Han(Yunnan) 0
  - **Northern populations**
    - Ewenki 0
    - Oraqen 0
    - Daur 0
    - Man 0
    - Uygur 4
    - Hui 8

- **YAP+ (%)**
  - **Southern populations**
    - Zhuang 0
    - Lahu 0
    - Wa 0
    - Dai 0
    - Tujia 2
    - Yao 2
    - Yi 15
    - Tibeten(Yunnan) 39
    - Tibeten(Tibet) 49
Multiple origins of Tibetan

- The YAP+ is an ancient polymorphism that originated from Central Asia (Altheide and Hammer, 1997)
- Our previous study on extant East Asian populations showed that M122C (defining H6, H7 and H8) is predominant in East Asian populations, especially in Han Chinese (54.1% on average) (Su et al., 1999) but it is nearly absent in other world populations, including Central Asians (RS Wells, personal communication).
Based on the fact of presentation with high frequency of both YAP + (41.3% in Zang and 44.4% in Khamba) and M122C (>35%), reflects a genetic source of Tibetan Y-chromosomes may have been derived from two different gene pools, one from Central Asia and the other from East Asia.
SNP (Single Nucleotide Polymorphisms)

- SNP application
  - Linkage analysis
  - Association study
  - Population genetics
  - Evolution
  - Complex disease study
  - Pharmacogenomics
  - Individual medical
The 126 genes of 21 chromosome were sequenced

<table>
<thead>
<tr>
<th></th>
<th>Gene</th>
<th></th>
<th>Gene</th>
<th></th>
<th>Gene</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ABCG1</td>
<td>45</td>
<td>DSCR5</td>
<td>89</td>
<td>PKNOX1</td>
</tr>
<tr>
<td>2</td>
<td>ADAMTS1</td>
<td>24</td>
<td>CBS</td>
<td>68</td>
<td>KCNJ15</td>
</tr>
<tr>
<td>3</td>
<td>ADAMTS5</td>
<td>25</td>
<td>CCT8</td>
<td>69</td>
<td>KCNJ6</td>
</tr>
<tr>
<td>4</td>
<td>ADARB1</td>
<td>26</td>
<td>CHAF1B</td>
<td>70</td>
<td>KIAA0179</td>
</tr>
<tr>
<td>5</td>
<td>AIRE</td>
<td>27</td>
<td>CLDN14</td>
<td>71</td>
<td>KIAA0184</td>
</tr>
<tr>
<td>6</td>
<td>APP</td>
<td>28</td>
<td>CLDN17</td>
<td>72</td>
<td>KIAA0539</td>
</tr>
<tr>
<td>7</td>
<td>ATP50</td>
<td>29</td>
<td>CLDN8</td>
<td>73</td>
<td>KIAA0653</td>
</tr>
<tr>
<td>8</td>
<td>ATP5A</td>
<td>30</td>
<td>COL18A1</td>
<td>74</td>
<td>KIAA0958</td>
</tr>
<tr>
<td>9</td>
<td>B3GALT5</td>
<td>31</td>
<td>COL6A1</td>
<td>75</td>
<td>KIAA0136</td>
</tr>
<tr>
<td>10</td>
<td>BACE2</td>
<td>32</td>
<td>COL6A2</td>
<td>76</td>
<td>LSS</td>
</tr>
<tr>
<td>11</td>
<td>BACH1</td>
<td>33</td>
<td>CRYAA</td>
<td>77</td>
<td>MCM3</td>
</tr>
<tr>
<td>12</td>
<td>BTG3</td>
<td>34</td>
<td>CRYZL1</td>
<td>78</td>
<td>MX1</td>
</tr>
<tr>
<td>13</td>
<td>C21orf1</td>
<td>35</td>
<td>CSTB</td>
<td>79</td>
<td>MX2</td>
</tr>
<tr>
<td>14</td>
<td>C21orf18</td>
<td>36</td>
<td>CTBP2</td>
<td>80</td>
<td>N143</td>
</tr>
<tr>
<td>15</td>
<td>C21orf2</td>
<td>37</td>
<td>CXADR</td>
<td>81</td>
<td>NCAM2</td>
</tr>
<tr>
<td>16</td>
<td>C21orf258</td>
<td>38</td>
<td>D21S2056E</td>
<td>82</td>
<td>NDUFV3</td>
</tr>
<tr>
<td>17</td>
<td>C21orf30</td>
<td>39</td>
<td>DNMT3L</td>
<td>83</td>
<td>NRIP1</td>
</tr>
<tr>
<td>18</td>
<td>C21orf33</td>
<td>40</td>
<td>DSCAM</td>
<td>84</td>
<td>PCNT</td>
</tr>
<tr>
<td>19</td>
<td>C21orf4</td>
<td>41</td>
<td>DSCR1</td>
<td>85</td>
<td>PCP4</td>
</tr>
<tr>
<td>20</td>
<td>c21orf5</td>
<td>42</td>
<td>DSCR2</td>
<td>86</td>
<td>PDE9A</td>
</tr>
<tr>
<td>21</td>
<td>C21orf6</td>
<td>43</td>
<td>DSCR3</td>
<td>87</td>
<td>PDXK</td>
</tr>
<tr>
<td>22</td>
<td>CBR1</td>
<td>44</td>
<td>DSCR4</td>
<td>88</td>
<td>PKL</td>
</tr>
<tr>
<td>23</td>
<td>CBR3</td>
<td>67</td>
<td>KCNE2</td>
<td>111</td>
<td>TFF2</td>
</tr>
</tbody>
</table>
The individual samples

<table>
<thead>
<tr>
<th>No1.</th>
<th>property</th>
<th>No2.</th>
<th>Sample</th>
<th>National</th>
<th>Sample location</th>
<th>age</th>
<th>Sex</th>
</tr>
</thead>
<tbody>
<tr>
<td>A01</td>
<td>individual</td>
<td>I1</td>
<td>Sun L</td>
<td>Han</td>
<td>Shandong</td>
<td>11</td>
<td>M</td>
</tr>
<tr>
<td>A02</td>
<td>individual</td>
<td>I2</td>
<td>Han F</td>
<td>Han</td>
<td>Shandong</td>
<td>9</td>
<td>M</td>
</tr>
<tr>
<td>A03</td>
<td>individual</td>
<td>I3</td>
<td>Liu H</td>
<td>Han</td>
<td>Jilin</td>
<td>37</td>
<td>M</td>
</tr>
<tr>
<td>A04</td>
<td>individual</td>
<td>I4</td>
<td>Zhu L</td>
<td>Han</td>
<td>Haerbin</td>
<td>24</td>
<td>M</td>
</tr>
<tr>
<td>A05</td>
<td>individual</td>
<td>I5</td>
<td>BF28</td>
<td>Han</td>
<td>Shanxi</td>
<td></td>
<td>F</td>
</tr>
<tr>
<td>A06</td>
<td>individual</td>
<td>I6</td>
<td>BF29</td>
<td>Han</td>
<td>Shanxi</td>
<td></td>
<td>F</td>
</tr>
<tr>
<td>A07</td>
<td>individual</td>
<td>I7</td>
<td>Meng S</td>
<td>Zhuang</td>
<td>Guanxi</td>
<td>22</td>
<td>M</td>
</tr>
<tr>
<td>A08</td>
<td>individual</td>
<td>I8</td>
<td>Huang g</td>
<td>Zhuang</td>
<td>Guanxi</td>
<td>8</td>
<td>M</td>
</tr>
<tr>
<td>A09</td>
<td>individual</td>
<td>I9</td>
<td>Ni M</td>
<td>Zang</td>
<td>Tibet</td>
<td>59</td>
<td>M</td>
</tr>
<tr>
<td>A10</td>
<td>control(-)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A11</td>
<td>individual</td>
<td>I11</td>
<td>Ni D</td>
<td>Bulang</td>
<td>Yunnan</td>
<td>18</td>
<td>M</td>
</tr>
<tr>
<td>A12</td>
<td>individual</td>
<td>I12</td>
<td>Yan Z</td>
<td>Bulang</td>
<td>Yunnan</td>
<td>30</td>
<td>M</td>
</tr>
<tr>
<td>B01</td>
<td>individual</td>
<td>I13</td>
<td>Piao Y</td>
<td>Caoxian</td>
<td>Jilin</td>
<td>24</td>
<td>M</td>
</tr>
<tr>
<td>B02</td>
<td>individual</td>
<td>I14</td>
<td>Yan D</td>
<td>Dai</td>
<td>Yunnan</td>
<td>43</td>
<td>M</td>
</tr>
<tr>
<td>B03</td>
<td>individual</td>
<td>I15</td>
<td>Liu Z</td>
<td>Han</td>
<td>Fujian</td>
<td>17</td>
<td>M</td>
</tr>
<tr>
<td>B04</td>
<td>individual</td>
<td>I16</td>
<td>Huang F</td>
<td>Han</td>
<td>Fujian</td>
<td>18</td>
<td>M</td>
</tr>
<tr>
<td>B05</td>
<td>individual</td>
<td>I17</td>
<td>Su Y</td>
<td>Menggu</td>
<td>Heilongjiang</td>
<td>14</td>
<td>M</td>
</tr>
<tr>
<td>B06</td>
<td>individual</td>
<td>I18</td>
<td>Hu R</td>
<td>Menggu</td>
<td>Heilongjiang</td>
<td>14</td>
<td>M</td>
</tr>
<tr>
<td>B07</td>
<td>individual</td>
<td>I19</td>
<td>GSM1</td>
<td>Miao</td>
<td>Guizhou</td>
<td></td>
<td>M</td>
</tr>
<tr>
<td>B08</td>
<td>individual</td>
<td>I20</td>
<td>GSM2</td>
<td>Miao</td>
<td>Guizhou</td>
<td></td>
<td>M</td>
</tr>
<tr>
<td>B09</td>
<td>individual</td>
<td>I21</td>
<td>SBY11</td>
<td>Yi</td>
<td>Sichuan</td>
<td></td>
<td>M</td>
</tr>
<tr>
<td>B10</td>
<td>individual</td>
<td>I22</td>
<td>SBY12</td>
<td>Yi</td>
<td>Sichuan</td>
<td></td>
<td>M</td>
</tr>
<tr>
<td>B11</td>
<td>individual</td>
<td>I23</td>
<td>XYUR1</td>
<td>Weiwaer</td>
<td>Xinjiang</td>
<td></td>
<td>M</td>
</tr>
<tr>
<td>B12</td>
<td>individual</td>
<td>I24</td>
<td>XYUR2</td>
<td>Weiwaer</td>
<td>Xinjiang</td>
<td></td>
<td>M</td>
</tr>
<tr>
<td>C01</td>
<td>individual</td>
<td>I25</td>
<td>Yan J</td>
<td>Dai</td>
<td>Yunnan</td>
<td></td>
<td>M</td>
</tr>
<tr>
<td>C02</td>
<td>individual</td>
<td>I26</td>
<td>HGG21</td>
<td>Han</td>
<td>Guangdong</td>
<td></td>
<td>F</td>
</tr>
<tr>
<td>C03</td>
<td>individual</td>
<td>I27</td>
<td>HGG22</td>
<td>Han</td>
<td>Guangdong</td>
<td></td>
<td>F</td>
</tr>
<tr>
<td>C04</td>
<td>individual</td>
<td>I28</td>
<td>He C</td>
<td>Han</td>
<td>Shanghai</td>
<td>39</td>
<td>F</td>
</tr>
<tr>
<td>C05</td>
<td>individual</td>
<td>I29</td>
<td>SLE4F</td>
<td>Han</td>
<td>Shanghai</td>
<td></td>
<td>M</td>
</tr>
<tr>
<td>C06</td>
<td>individual</td>
<td>I30</td>
<td>SLE4M</td>
<td>Han</td>
<td>Shanghai</td>
<td></td>
<td>F</td>
</tr>
<tr>
<td>C07</td>
<td>individual</td>
<td>I31</td>
<td>SLE4S1</td>
<td>Han</td>
<td>Shanghai</td>
<td></td>
<td>M</td>
</tr>
<tr>
<td>C08</td>
<td>individual</td>
<td>I10</td>
<td>Zhan D</td>
<td>Zang</td>
<td>Tibet</td>
<td>28</td>
<td>M</td>
</tr>
</tbody>
</table>
# The pooling samples

<table>
<thead>
<tr>
<th>No1.</th>
<th>property</th>
<th>No2.</th>
<th>Sample</th>
<th>Sample location</th>
<th>sum</th>
<th>Sex</th>
</tr>
</thead>
<tbody>
<tr>
<td>C09</td>
<td>pool</td>
<td>P01</td>
<td>Han North</td>
<td>Shanxi</td>
<td>39</td>
<td></td>
</tr>
<tr>
<td>C10</td>
<td>pool</td>
<td>P02</td>
<td>Han, South</td>
<td>Guangdong</td>
<td>49</td>
<td></td>
</tr>
<tr>
<td>C11</td>
<td>pool</td>
<td>P03</td>
<td>High choleserol</td>
<td>Nanjing</td>
<td>43</td>
<td></td>
</tr>
<tr>
<td>C12</td>
<td>pool</td>
<td>P04</td>
<td>Coronary heart</td>
<td>Shanghai</td>
<td>239</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Disease</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D01</td>
<td>pool</td>
<td>P05</td>
<td>SLE</td>
<td>Shanghai</td>
<td>84</td>
<td></td>
</tr>
<tr>
<td>D02</td>
<td>pool</td>
<td>P06</td>
<td>Diabetes I</td>
<td>Shanghai</td>
<td>75</td>
<td></td>
</tr>
<tr>
<td>D03</td>
<td>pool</td>
<td>P07</td>
<td>NIDDM</td>
<td>Beijing</td>
<td>96</td>
<td></td>
</tr>
<tr>
<td>D04</td>
<td>pool</td>
<td>P08</td>
<td>Hypertension</td>
<td>Shanghai</td>
<td>96</td>
<td></td>
</tr>
<tr>
<td>D05</td>
<td>pool</td>
<td>P09</td>
<td>Control</td>
<td>Shanghai</td>
<td>96</td>
<td></td>
</tr>
<tr>
<td>D06</td>
<td>pool</td>
<td>P10</td>
<td>Hypertension</td>
<td>Henan</td>
<td>36</td>
<td></td>
</tr>
<tr>
<td>D07</td>
<td>pool</td>
<td>P11</td>
<td>Hypotension</td>
<td>Henan</td>
<td>36</td>
<td></td>
</tr>
<tr>
<td>D08</td>
<td>pool</td>
<td>P12</td>
<td>Graves' disease</td>
<td>Liaoning</td>
<td>96</td>
<td></td>
</tr>
<tr>
<td>D09</td>
<td>pool</td>
<td>P13</td>
<td>Schizophrenia</td>
<td>Shanghai</td>
<td>200</td>
<td></td>
</tr>
<tr>
<td>D10</td>
<td>pool</td>
<td>P14</td>
<td>Schizophrenia</td>
<td>Beijing</td>
<td>200</td>
<td></td>
</tr>
<tr>
<td>D11</td>
<td>pool</td>
<td>P15</td>
<td>Affrican</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D12</td>
<td>pool</td>
<td>P16</td>
<td>European</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D12</td>
<td>pool</td>
<td>P16</td>
<td>European</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Y-Chromosome Evidence for a Northward Migration of Modern Humans into Eastern Asia during the Last Ice Age.

Bing Su,1 Junhua Xiao,2 Peter Underhill,5 Ranjan Deka,7 Weiling Zhang,2 Joshua Akey,1 Wei Huang,3,4 Di Shen,1 Daru Lu,2 Jingchun Luo,2 Jiayou Chu,8 Jiazhen Tan,2 Peidong Shen,5 Ron Davis,5,6 Luca Cavalli-Sforza,5 Ranajit Chakraborty,1 Momiao Xiong,1 Ruotu Du,2 Peter Oefner,5,6 Zhu Chen,1,4 and Li Jin1,2,3

1Human Genetics Center, University of Texas-Houston, Houston; 2Morgan-Tan International Center for Life Sciences and Institute of Genetics, Fudan University, 3National Human Genome Center at Shanghai, and 4Rui-Jin Hospital, Shanghai Second Medical University, Shanghai; 5Department of Genetics, Stanford University, and 6Stanford DNA Sequencing and Technology Center, Palo Alto; 7Department of Environmental Health, University of Cincinnati, Cincinnati; 8Institute of Medical Biology, The Chinese Academy of Medical Sciences, Kunming, Yunnan, China; and 9Institute of Genetics, The Chinese Academy of Sciences, Beijing

Summary

The timing and nature of the arrival and the subsequent expansion of modern humans into eastern Asia remains controversial. Using Y-chromosome biallelic markers, we investigated the ancient human-migration patterns in eastern Asia. Our data indicate that southern populations in eastern Asia are much more polymorphic than northern populations, which have only a subset of the southern haplotypes. This pattern indicates that the first settlement of modern humans in eastern Asia occurred in mainland Southeast Asia during the last Ice Age, coinciding with the arrival there (Brooks and Wood 1990; Li and Etler 1992; Wu and Poirier 1995; Etler 1996; Wolpoff 1996), which has been claimed to be a challenge to the well-known “out-of-Africa” hypothesis of modern human evolution (Cann et al. 1987; Vigilant et al. 1991). However, not all paleoanthropologists agree that the Asian fossil record shows a clear continuity from Homo erectus to H. sapiens sapiens (Stringer and Andrews 1988; Wilson and Cann 1992), thereby casting doubt on the in situ Asian-origin hypothesis.

A recent study of Asian populations, which used microsatellite markers, questioned the ABO theory.
Short Report

Multiple origins of Tibetan Y chromosomes

Yaping Qian · Bingzhi Qian · Bing Su · Jiankun Yu · Yuehai Ke · Zhengtao Chu · Lei Shi · Daru Lu · Jiayou Chu · Li Jin

Y. Qian · J. Yu · Z. Chu · L. Shi · J. Chu
Institute of Medical Biology, The Chinese Academy of Medical Sciences, Kunming, Yunnan, China

B. Qian · D. Lu · L. Jin
Institute of Genetics and Morgan-Tan International Center for Life Sciences, Fudan University, Shanghai, China

B. Su · Y. Ke · L. Jin
Human Genetics Center, University of Texas Houston, 6901 Bertner, Houston, Texas, 77030, U.S.A.

E-mail: ljin@utshph.sph.uth.tmc.edu
Phone: +1-713-5009846
Fax: +1-713-5000900
Polynesian origins: Insights from the Y chromosome


*Human Genetics Center, University of Texas Health Science Center, Houston, TX 77225; †Institute of Genetics, Fudan University, Shanghai, China 200433; ‡Department of Genetics, Stanford University, Stanford, CA 94305; †§Department of Genetics, University of Nottingham, Nottingham, NG7 2UH, United Kingdom; ‡Department of Molecular Medicine, Australian National University, Canberra, Australia ACT 2601; ‡§Department of Medicine and International Health Institute, Brown University School of Medicine, Providence, RI 02912; †**Department of Anthropology, Pennsylvania State University, University Park, PA 16802. ††Department of Medical Biology, Chinese Academy of Medical Sciences, Kunming, China 650107; and ‡§§Department of Environmental Health, University of Cincinnati, Cincinnati, OH 45267-0056

Communicated by Francisco Salzano, University of Rio Grande do Sul, Porto Alegre, Brazil, March 3, 2000 (received for review January 27, 2000)

The question surrounding the colonization of Polynesia has remained controversial. Two hypotheses, one postulating Taiwan as the putative homeland and the other asserting a Melanesian origin of the Polynesian people, have received considerable attention. In this work, we present haplotype data based on the distribution of 19 biallelic polymorphisms on the Y chromosome in a sample of 551 male individuals from 36 populations living in Southeast Asia, Taiwan, Micronesia, Melanesia, and Polynesia. Surprisingly, nearly none of the Taiwanese Y haplotypes were found in Micronesia and Polynesia. Likewise, a Melanesian-specific haplotype was not found among the Polynesians. However, all of the Polynesian, Micronesian, and Taiwanese haplotypes are present in the extant Southeast Asian populations. Evidently, the Y-chromosome data do not lend support to either of the prevailing hypotheses. Rather, we postulate that Southeast Asia provided a genetic source for two independent migrations, one toward Taiwan and the other toward Polynesia through island Southeast Asia.

The major prehistoric events leading to the settlement of Polynesia have been examined from various perspectives, and two different models of population movements are

published data (14) on Southeast Asian populations have been incorporated in this study for purposes of comparison. The details of the 19 biallelic markers, PCR amplification protocols, haplotype construction, and nomenclature are given in Su et al. (14). Primer sequences and protocols for PCR amplification and analysis of these markers are available on request. Haplotype diversity was calculated following the method of Nei (15), and genetic distances were computed following Nei (15) and Reynolds et al. (16).

Results and Discussion

Using 19 biallelic markers, we identified 15 haplotypes in the total sample of 551 Y chromosomes. The haplotype frequencies in various populations are shown in Table 1. Earlier we had presented a parsimonious phylogenetic tree of the haplotypes based on the 19 markers (14), in which H1 was considered as the ancestral haplotype because of its appearance in chimpanzees. Among the other haplotypes, H2 is also relatively ancient with its occurrence in both African and non-African populations and, H5 appeared as the common ancestor of all other non-African haplotypes, which are regionally distributed. The Southeast
Examples of our Current Work 3

Three Hypotheses on Modern Human Origin

![Diagram showing timelines of human evolution with labels for Homo erectus, Homo sapiens, and H. s. sapiens, along with regions of Africa, Europe, Asia, and Australia.]
The Formation Of Modern Chinese Populations from microsatellites evidence

*Research materials*

- 28 Chinese populations
- 15 foreign groups as control
Basic methods

- Microsatellites
- Genotyping and Genescan
- Neighbor-joining method
What is the difference from the past researchers?

1. Strict in samples collecting
2. Marks and analysis methods
3. Compared with foreign samples
Main conclusions:

- 1. There are some genetic differences between nationalities in north and south China, but the relations are quite complicated.
Main conclusions:

- 2. Chinese people’s migration might be from the south to the north.
Main conclusions:

- Modern humans originating in Africa constitute the majority of the current gene pool in East Asia.
To test the hypotheses of modern human origin in East Asia, Yuehai Ke and Li Jin sampled 12,127 male individuals from 163 populations across Southeast Asia, Oceania, East Asia, Siberia, and Central Asia and typed for three Y chromosome biallelic markers (YAP, M89, and M130).
Study on 12,000 Y Chromosomes

- All the individuals carried a mutation coalesce to another mutation (M168T), which originated in Africa about 35,000 to 89,000 years ago.
- Therefore, the data do not support even a minimal in situ hominid contribution in the origin of anatomically modern humans in East Asia.
Conclusion

The phylogeny based on microsatellites and Y chromosomes revealed a clear distinction between southern and northern Chinese populations. In both phylogenies with different loci and populations, populations from East Asia always derived from a single lineage indicating the single origin of those populations.
The current analysis suggests that the southern populations in East Asia may be derived from the populations in Southeast Asia that originally migrated from Africa and the northern populations were under strong genetic influence from Altaic populations from the north. It is now probably safe to conclude that modern human originated in Africa constitutes majority of the current gene pool in East Asia.
Genetic relationship of populations in China

J. Y. CHU\textsuperscript{a,b}, W. HUANG\textsuperscript{b,c}, S. Q. KUANG\textsuperscript{c}, J. M. WANG\textsuperscript{c}, J. J. XU\textsuperscript{d}, Z. T. CHU\textsuperscript{a}, Z. Q. YANG\textsuperscript{a}, K. C. CHEN\textsuperscript{a}, M. WU\textsuperscript{e}, Z. C. GENG\textsuperscript{f}, C. C. TAN\textsuperscript{g}, R. F. DU\textsuperscript{d}, AND L. JING\textsuperscript{b,h,i}

\textsuperscript{a}Institute of Medical Biology, Chinese Academy of Medical Sciences, Kunming, People's Republic of China; \textsuperscript{b}Rui-Jin Hospital, Shanghai Jiao Tong University, Shanghai, People's Republic of China; \textsuperscript{c}Institute of Genetics, Chinese Academy of Sciences, Beijing, People's Republic of China; \textsuperscript{d}Institute of Cancer Research, Chinese Academy of Medical Sciences, Shanghai, People's Republic of China; \textsuperscript{e}Institute of Genetics, Fudan University, Shanghai, People's Republic of China; and \textsuperscript{f}Human Genetics Center, University of Texas, Houston, TX 77225

Contributed by Jiazheng Tan, June 26, 1998

ABSTRACT Despite the fact that the continuity of morphology of fossil specimens of modern humans found in China has repeatedly challenged the Out-of-Africa hypothesis, Chinese populations are underrepresented in genetic studies. Genetic profiles of 28 populations sampled in China supported the distinction between southern and northern populations, while the latter are biphyletic. Linguistic boundaries are often transgressed across language families studied, reflecting substantial gene flow between populations. Nevertheless, genetic evidence does not support an independent origin of \textit{Homo sapiens} in China. The phylogeny also suggested that it is more likely that ancestors of the populations currently residing in East Asia entered from Southeast Asia.

Microsatellites have been widely used to study the relationship among human populations. Traditional assignments (8–10). Simulation results indicate that microsatellite loci generally provide a more reliable picture of the relationship among closely related populations than other markers such as blood groups (11). Microsatellite loci also reflect closely related populations by reflecting the geographical distribution of the populations in the study, which may result in major changes in allele frequencies (12). In turn, the reliability of such changes in the presence of genetic admixture is therefore compromised (17). Nevertheless, the ease of analysis of satellite alleles and the availability of large sets of highly informative loci across the human genome make microsatellites an ideal tool for population studies.
African Origin of Modern Humans in East Asia: A Tale of 12,000 Y Chromosomes

Yuehai Ke,1,4 Bing Su,2,1,3,4 Xiufeng Song,1 Daru Lu,1 Lifeng Chen,5 Hongyu Li,6 Chunlian Qi,1 Sangkot Marzuki,4 Ranjan Deka,1 Peter Underhill,4 Chunjie Xiao,6 Mark Shriver,6 Jeff Lelliott,9 Douglas Wallace,6 Robert Spencer Wells,10 Mark Sisestra,10 Peter Oefner,10 Dingliang Zhu,7 Jianzhong Jin,10 Wei Huang,10 Ranjit Chakraborty,3 Zhu Chen,5,6 Li Jin,1,5,9,11

To test the hypotheses of modern human origin in East Asia, we sampled 52,172 male individuals from 163 populations and typed for three Y chromosome biallelic markers (YAP, M89, and M130). All the individuals carried a mutation at one of the three sites. These three mutations (YAP +, M89T, and M130T) coalesce to another mutation (M168T), which originated in Africa about 35,000 to 89,000 years ago. Therefore, the data do not support even a minimal level of recent natural contribution in the origin of the modern human populations in East Asia.

The "Out of Africa" hypothesis suggests that anatomically modern humans originated in Africa about 100,000 years ago and spread outward and completely replaced local archaic populations outside Africa (1, 2). This proposition has been supported by genetic and archaeological findings (3–9). The replacement in Europe was supported by recent ancient DNA analyses, which ruled out the contribution of Neandertal to modern Europeans (10, 11). However, it has been argued that the abundant hominid fossils found in China and other regions in East Asia (e.g., Peking man and Java man) demonstrate continuity, not only in morphological characters but also in spatial and temporal distributions (12–14). In this report, we test the competing hypotheses of modern human origins using Y chromosome polymorphisms.

We sampled 12,172 male individuals from 163 populations across Southeast Asia, Oceania, East Asia, Siberia, and Central Asia and typed for three Y chromosome biallelic markers (YAP, M89, and M130) (17, 18) (Table 1). Being a single-locus multiplex (i.e., haplotype) system, the Y chromosome is one of the most powerful molecular tools for tracing human evolutionary history (5, 9, 19–21). In previous Y chromosome studies, an extreme geographic structure was revealed in global populations in which the oldest clade represents Africans and the younger ones represent Asians and all non-African populations (21). One Y chromosome polymorphism (C to T mutation) at the M168 locus is shared by all non-African populations and was originally derived from Africa on the basis of a study of 1062 globally representative male individuals (21). The age of M168 was estimated to be 44,000 years (95% confidence interval: 35,000 to 89,000 years), marking the recent Out of Africa migrations (21). Under the M168T lineage, there are three major derived sublineages defined by polymorphisms at loci YAP (Aa insertion) (2), M89 (C to T mutation), and M130 (C to T mutation, also called RPS4Y) (Fig. 1) (22, 23). Therefore, these three markers can be used to test the completeness of the replacement of modern humans of African origin in East Asia. An observation of a male individual not carrying any of the three polymorphisms would be indicative of a potential ancient origin and could possibly lead to the rejection of such completeness.

Each of the 12,172 samples typed carried one of the three polymorphisms (YAP+, M89T, or M130T) (Table 1). In other words, they all fall into the lineage of M168T that was originally derived from Africa. Hence, no ancient non-African Y chromosome was found in the extant East Asian populations (P = 5.4 × 10−10 assuming a frequency of 1/1000 of local contribution in the extant populations), suggesting an absence of either an independent origin or a 1,000,000-year shared global evolution. This result indicates that modern humans of African origin completely replaced earlier populations in East Asia.
Examples of our Current Work 4

GENOME DIVERSITY AND DISEASE

*Human Origins/Evolution*

*Disease susceptibility*

“*Disease is Part of an Evolutionary Process*”
Genome diversity research related to diseases

G6PD (glucose-6-phosphate dehydrogenase)

Deficiency gene rate

- Dai 0.1610
- Lahu 0.1136
- Deang 0.2014
- Wa 0.0821
- Jingpo 0.0450
- Hani 0.0370
**Incidence rate of abnormal hemoglobin**

<table>
<thead>
<tr>
<th>Ethnicity</th>
<th>Incidence Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Achang</td>
<td>0.4293</td>
</tr>
<tr>
<td>Jino</td>
<td>0.1625</td>
</tr>
<tr>
<td>Deang</td>
<td>0.1399</td>
</tr>
<tr>
<td>Jingpo</td>
<td>0.1225</td>
</tr>
<tr>
<td>Dai</td>
<td>0.1212</td>
</tr>
<tr>
<td>Lisu</td>
<td>0.1044</td>
</tr>
<tr>
<td>Aini</td>
<td>0.1180</td>
</tr>
<tr>
<td>Bai</td>
<td>0.0479</td>
</tr>
<tr>
<td>Lahu</td>
<td>0.0299</td>
</tr>
<tr>
<td>Hani</td>
<td>0.0182</td>
</tr>
<tr>
<td>Wa</td>
<td>0.0182</td>
</tr>
</tbody>
</table>
System lupus erythematosus is an autoimmune disease, and a complex polygenic disorder

**Problems for complex traits:**

- Incomplete penetrance
- Phenocopy
- Genetic heterogeneity
- Polygenic inheritance
- High frequency of mutant alleles
- Mitochondrial inheritance
- Anticipation from trinucleotide repeats
- Imprinting
Some conclusions of SLE

- TNFa₁ (p=0.0281), TNFc₂ (p=0.0000) allele frequencies were significantly increased in the SLE group compared to controls.

- But TNFa₂ (p=0.0476), c₃ (p=0.0043), c₄ (p=0.0013), d₆ (p=0.0448) allele frequencies were decreased in the SLE group compared to controls.

- Meanwhile, we have found the possible new genotypes in Yunnan Han population: TNFc₃ (size of fragment is 161bp), TNFc₄ (163bp), TNFd₇ (136bp), TNFd₈ (138bp), TNFe₁ (99bp).
Parkinson’s disease

- Parkinson’s disease (PD) is the second most common neurodegenerative disorder after Alzheimer’s disease, affecting people seriously.
- Although many factors, including circumstance, neurotoxin etc, are regarded as probable reasons, genetics analysis in recent years revealed that PD was associated with some genetic factor.
Linkage mapping analysis of this PD family

- Using 400 microsatellite markers to perform whole genome genescanning, and position the rough location of the candidate gene
- Design a series of primers to position the exact location of gene
- Detect the mutation of candidate gene and find the genetic reason of PD
Long-lived people in Bama

- Bama in Guangxi province is one of famous long-lived people counties in the world
- The fifth Chinese census (in 2000)
  - the number of longevous people in Bama whose age over 100 is up to 74 from 66 in 1990
  - the proportion of long-lived people, 0.031%, is the highest in the whole world
Beautiful Bama
The old people of Bama
Two Pedigrees of Bama Old People

Huang Buhan, male, 105

Huang Mayou, female, 104
Call for international cooperation

- Data base and analysis
- Compared with foreign samples.
- Exchange DNA samples and cell lines.
- Cooperation on Diversity and Disease Gene research.
Acknowledgements

- Chinese Natural Science foundation
- Ministry of Science and Technology

- Chen Zhu, Chinese Academy of Sciences,
- Qiang Boqin, Chinese Academy of Medical Sciences
- Jin Li, Cincinaty University, USA,
- Li Pu, Harbin Medical University,
- Xu Jiujin, Institute of Genetics, CAS,
- Huang Wei, Shanghai 2ed Medical University,
- Qian Yaping, Chu Zhengtao, Yang Zhaoqing,
- Yu Jiankun, Institute of Medical Biology, CAMS