Different Matrilineral Contributions to Genetic Structure of Ethnic Groups in the Silk Road Region in China

Yong-Gang Yao,† Qing-Peng Kong,‡†‡‡, Cheng-Ye Wang,‡‡ Chun-Ling Zhu,‡* and Ya-Ping Zhang‡†

*Laboratory of Cellular and Molecular Evolution, and Molecular Biology of Domestic Animals, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, China; †Laboratory for Conservation and Utilization of Bio-resource, Yunnan University, Kunming, China; and ‡Graduate School of the Chinese Academy of Sciences, Beijing, China

Previous studies have shown that there were extensive genetic admixtures in the Silk Road region. In the present study, we analyzed 252 mtDNAs of five ethnic groups (Uygur, Uzbek, Kazak, Mongolian, and Hui) from Xinjiang Province, China (through which the Silk Road once ran) together with some reported data from the adjacent regions in Central Asia. In a simple way, we classified the mtDNAs into different haplogroups (monophyletic clades in the rooted mtDNA tree) according to the available phylogenetic information and compared their frequencies to show the differences among the matrilineal genetic structures of these populations with different demographic histories. With the exception of eight unassigned M*, N*, and R* mtDNAs, all the mtDNA types identified here belonged to defined subhaplogroups of haplogroups M and N (including R) and consisted of subsets of both the eastern and western Eurasian pools, thus providing direct evidence supporting the suggestion that Central Asia is the location of genetic admixture of the East and the West. Although our samples were from the same geographic location, a decreasing tendency of the western Eurasian-specific haplogroup frequency was observed, with the highest frequency present in Uygur (42.6%) and Uzbek (41.4%) samples, followed by Kazak (30.2%), Mongolian (14.3%), and Hui (6.7%). No western Eurasian type was found in Han Chinese samples from the same place. The frequencies of the eastern Eurasian-specific haplogroups also varied in these samples. Combined with the historical records, ethno-origin, migratory history, and marriage customs might play different roles in shaping the matrilineal genetic structure of different ethnic populations residing in this region.

Introduction

Central Asia is located in an intermediate region of the Eurasian continent and has undergone unceasing migrations in both prehistorical and historical times. Hitherto, two competing hypotheses concerning the origin of Central Asians have been proposed on the basis of the classical genetic markers and uniparental genetic data. The genetic admixture hypothesis suggested that Central Asians were the results of admixture of the eastern and western Eurasians (Zhao and Lee 1989; Cavalli-Sforza, Menozzi, and Piazza 1994; Comas et al. 1998; Yao et al. 2000). Under this hypothesis, one would expect that the genetic components of Central Asians consisted of a subset of those typical of the East and the West. On the contrary, the Central Asian heartland hypothesis proposed that Central Asia was a heartland and reservoir for the genetic diversity of Europe, America, and India (Wells et al. 2001). Under this hypothesis, Central Asia should harbor not only the founder types of most eastern and western Eurasian haplogroups but also some unique genetic components. Evidence from mtDNA supported the genetic admixture hypothesis (Comas et al. 1998; Yao et al. 2000); however, both studies (Comas et al. 1998; Yao et al. 2000) only employed a comparison of population genetic parameters, such as haplotype diversity, nucleotide diversity, and pairwise differences, together with a Neighbor-Joining tree, to show the clustering pattern of the Central Asian populations among the Eurasian samples. None of them had performed a detailed comparison at a lineage level for the populations in this region. Although Comas et al. (1998) and Yao et al. (2000) tried to assign the mtDNAs to the eastern and western Eurasian pools, the strategies they adopted were problematic from the mtDNA phylogenetic point of view (Yao et al. 2002a), thus hindering them from better understanding the difference in matrilineal genetic structures of these samples. Fortunately, the recently available well resolved mtDNA phylogeny based on complete sequences in Europeans (Finnilä, Lehtonen, and Majamaa 2001; Herrnstadt et al. 2002) and East Asians (Kivisild et al. 2001; Kong et al. 2003) makes it possible for us to dissect the lineages into respective haplogroups and thus gain insights into their matrilineal genetic structures. Such a strategy has proved to be very useful in reconstructing the past (Richards et al. 2000; Yao et al. 2002a, 2003; Yao and Zhang 2002; Kong et al. 2003a). In recent reports, Comas et al. (2004) used this strategy and analyzed 232 mtDNAs from 12 Central Asian populations. They found that nearly all of the samples could be assigned to a subset of either the eastern or the western Eurasian pools, thus supporting the admixture hypothesis. Yao and Zhang (2003) came to a similar result by reanalyzing the ancient DNAs from Central Asia. Quintana-Murci et al. (2004) evaluated the genetic structure of populations from the Iranian plateau, the Indus Valley, and Central Asia and constructed a well defined yet complex mtDNA landscape in this region. All these studies, evidently, are helpful in understanding the mtDNA variation in Central Asia and in uncovering the veil of the “Central Asian corridor.”

Xinjiang Province, located in northwest China, is in the vicinity of and sometimes has been regarded as part of Central Asia. This region is famous for being home to part of the ancient Silk Road and has undergone unceasing migration events through present day, maintaining various cultures (Ge, Wu, and Chao 1997; Chen 1999). However, few mtDNA studies have been carried out for the ethnic
population residing in this region (Yao et al. 2000). A finer dissection of matrilineal components for more samples from this region will undoubtedly be helpful in better understanding the origin of Central Asians and in providing a clearer insight into the genetic structure of the ethnic populations that have undergone different demographic histories; thus serving as a good model for understanding whether the matrilineal genetic structure of the populations reflect the recent demographic episodes that have occurred and/or the different cultures they practiced.

In this study, we compared the matrilineal genetic components of six ethnic populations (Uygur, Uzbek, Kazak, Mongolian, Hui, and Han) from Xinjiang, China, and the reported samples from the adjacent regions (Comas et al. 1998, 2004), as well as a Mongolian population from northeast China (Kong et al. 2003a). Our results showed that the extent of genetic admixtures of the eastern and western Eurasian lineages varied among these populations from Xinjiang. There was no central heartland in mtDNA distribution pattern in this region.

Materials and Methods

Subjects

The whole blood of 45 Huis, 49 Mongolians, 58 Uzbeks, and 23 Kazaks was collected from various villages in concentrated communities of each ethnic population in Yili and Kashen, Xinjiang, China (fig. 1). Informed consent was obtained from all individuals involved. In addition, the reported 30 Kazaks, 47 Uygurs, and 47 Hans from Xinjiang (Yao et al. 2000, 2002a) were included for comparison.

Among the six populations from Xinjiang considered here, the Han was recorded to migrate to this region mainly in the 1950’s from other parts of China (cf. Yao et al. 2002a). The Mongolian, Hui, Uzbek, and Kazak populations underwent more recent migration to this region than Uygur, but earlier than the Han people (Du and Yip 1993; Ge, Wu, and Chao 1997; Chen 1999). It should be mentioned that although the Uzbeks and Kazaks in Xinjiang are included in recent migration histories, both of them could trace their origins to the populations residing in Central Asia more than 2,000 years ago (Chen 1999), Hans and Huis speak Chinese (which belongs to the Sino-Tibetan language family), whereas Uygurs, Uzbeks, and Kazaks speak languages belonging to the Turkic language group of the Altaic language family. The Mongolians in Xinjiang speak a Uirad dialect of the Mongolian language that belongs to the Mongolian language group of the Altaic language family (Du and Yip 1993). The Uygurs, Huis, Uzbeks, and Kazaks are Muslims whereas the Mongolians generally believe in the Yellow sect of Lamaism (Du and Yip 1993).

Data Collection

Genomic DNA was isolated from the whole blood by standard phenol/chloroform method. The mtDNA control region hypervariable segment I (HVS-I) sequences of Uzbek (N=58), Hui (N=45), Kazak (N=23), and Mongolian (N=49) samples were amplified and sequenced as described in our previous studies (Yao et al. 2002a, 2003; Kong et al. 2003a). Some samples were further selected for the hypervariable segment II (HVS-II) region sequencing and/or coding region motif(s) typing by using the same strategy and condition described in our previous studies (Yao et al. 2002a, 2003; Kong et al. 2003a). The reported Uygur (N=47) and Kazak (N=30) samples with only HVS-I information available (Yao et al. 2000) were also subjected to haplogroup-specific HVS-II and/or coding region polymorphisms typing in this study. Other reported mtDNA data sets considered here included 47 Hans from Xinjiang (with HVS-I and II and coding region information available; Yao et al. 2002a); 55 Uighurs and 55 Kazakhs from Kazakhstan; 47 Sary-Tash Kirghizs and 48 Talas Kirghizs from Kirgizia (with HVS-I information; Comas et al. 1998); and 232 mtDNAs from 12 Central Asian populations (with HVS-I and II information; Comas et al. 2004). Because the sample size of each population in Comas et al. (2004) was relatively small (≤20), we aggregated all the samples together as one population in our analysis. We also included 48 Mongolians from Inner Mongolia (Kong et al. 2003a) to compare with the Mongolian sample from Xinjiang. The sequence data of the mtDNA control region in this study can be retrieved from DDBJ/EMBL/GenBank by accession numbers AY677733-AY678071.
MtDNA Haplogroup Classification

We used the strategy described in recent studies (Yao et al. 2002a, 2003; Kong et al. 2003a, 2004) to classify the mtDNAs into haplogroups, and we followed the haplogroup notation system described by Richards et al. (2000), Yao et al. (2002a, 2003), and Kong et al. (2003b). Briefly, by a first round of haplogroup-specific HVS-I motif searching and (near-)matching with the published data set with coding region information available, we tentatively assigned each mtDNA to respective haplogroups. Then, the haplogroup-specific HVS-II and/or coding region motif(s) of the mtDNAs were further tested to confirm the predicted haplogroup status of each mtDNA. If the predicted HVS-II or coding region motifs of the mtDNA were not observed, additional typing efforts were employed to further allocate the mtDNA into the nested haplogroup it belongs to. For example, one mtDNA HVS-I sequence with mutations at sites 16223 and 16362 could be D, G, or another type in M or N. We firstly tested for haplogroup D-specific polymorphism, 5178A (−5176A(A)), to confirm the D status. Site 4833 (+4831Hحح) would be needed to test for G status if 5178A was found to be absent in that sample. If both mutations were absent, we would start a further round of typing in other coding-region fragment(s) to determine the status of the sample. Through this discrimination process, each mtDNA could be easily identified without exhaustively typing for all haplogroups’ characteristic mutations. For the reported mtDNAs in Comas et al. (1998, 2004), we tentatively assigned them to respective haplogroups according to information provided by HVS-I and II and/or by (near-)matching with the published data sets with coding region information as well as the mtDNAs analyzed in current study.

Data Analyses

The frequencies of mtDNA haplogroups in each population were estimated. We focused on the comparison of the haplogroup distribution patterns between the samples with different ethno-origins, demographic histories, and religions. Fisher exact tests were performed to quantify the difference between the frequencies of certain haplogroup(s) in the samples. Furthermore, a principal component (PC) map was constructed on the basis of the haplogroup frequency matrix to show the clustering pattern of the samples.

Results

MtDNA Haplogroup Profile

Sequence variation of 252 mtDNAs under study is listed in table 1. With the exception of eight M*, N*, and R* types that could not be further assigned, each mtDNA can be allocated to the smallest branches in the haplogroup classification tree (fig. 2) based on the HVS-I/II and certain coding-region information. Several specific lineages, which seemed to be prevalent in East Asians, are identified and characterized here. For instance, one Mongolian sample (Mg50) harbors a motif (16145−16188−16223) that was also present in five Tibetans (cf. Yao and Zhang 2002) and one Korean (Kor92; Kong et al. 2003a). Further information from region 10171−10659 reveals that Tibetan76 (Yao and Zhang 2002), Kor92, and Mg50 shared mutations at sites 10398, 10400, and 10411. We therefore designate this lineage (defined by a motif 10398−10400−10411−16145−16188−16223) as haplogroup M13, and its subclade, which is further characterized by mutations 16148 and 16189 in HVS-I, is named M13a. Extensive literature and our own unpublished data set searching reveals that haplogroup M13 is also present in five Yakuts (Fedorova et al. 2003; Pakendorf et al. 2003) and two samples from Yunnan Province, southwest China (unpublished data). One Hui sample (Hu98; with motif 16086−16297−16324) obviously belongs to an unidentified haplogroup that was prevalent in the aboriginal Taiwanese mtDNAs reported by Tajima et al. (2003). Additional typing efforts showed that Hu98 had mutations 199, 9824 (recognizable by +9820Hحح), and 12405, thus suggesting that this lineage is a subclade of haplogroup M7b, and it is named M7b3 here. The control-region motif of the haplogroup B type, Hui59, matches a complete sequence (SD10313) reported in our recent study (Kong et al. 2003b). Further analysis revealed that Hui59 had mutations 13928C and 14305; we thus designate the clade (characterized by motif 13928C−14305−16093−16179−16189) harboring Hui59 and SD10313 as B6. It is noteworthy that three samples, Mg46 (F2b), Uzb57 (M*) and, Uyg5 (M*), bear the 9-bp deletion in the COII/tRNALys intergenic region (one of the specific motifs of haplogroup B), thus, constituting additional cases for multiple origins of the 9-bp deletion in Chinese (Yao, Watkins, and Zhang 2000). In total, 75 defined (sub)haplogroups are identified in our samples (fig. 2). It is obvious that nearly all the mtDNAs identified here belonged to a subset of either the defined western (Richards et al. 1998, 2000; Macaulay et al. 1999) or eastern (Yao et al. 2002a; Kong et al. 2003b) Eurasian pool. No Central Asian–specific haplogroup was found in the present study.

Haplogroup Distribution Frequency

Table 2 shows the mtDNA haplogroup frequencies in the six populations sampled in Xinjiang. By comparing them with the reported mtDNA data sets (Comas et al. 1998, 2004; Kong et al. 2003b), two main features could be discerned. First, the western Eurasian-specific haplogroups occupy different percentages of the matrilineal components in these populations. The Han Chinese samples contain no western Eurasian-specific haplogroups, i.e., all haplotypes could be assigned to the eastern Eurasian mtDNA haplogroups (Yao et al. 2002a; Kong et al. 2003b), while 6.7% and 14.3% of the lineages in Hui and Mongolian samples, respectively, belong to the western Eurasian-specific haplogroups. In the other samples from Xinjiang and its proximal regions, the percentages of the western Eurasian lineages are all higher than 27%. The frequencies of the western Eurasian-specific haplogroups in Uygur and Uzbek from Xinjiang are approximately equal to those of Kazakh, Uighur, and Sary-Tash Kirghiz from Central Asia (≥40%; Comas et al. 1998). Second, among the six populations from the same province, large differences are observed in the distribution frequencies of each haplogroup. For instance, haplogroup B presents a high frequency in the Hui sample (17.8%), followed by
<table>
<thead>
<tr>
<th>Sample</th>
<th>Haplogroup</th>
<th>HVS-I (16000+)</th>
<th>HVS-II (10000+)</th>
<th>9-bp Polymorphisms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uzb25</td>
<td>D</td>
<td>223 362</td>
<td>73 263 309 + CC 315 + C</td>
<td>2</td>
</tr>
<tr>
<td>Uzb23</td>
<td>D</td>
<td>223 362</td>
<td>73 263 309 + C 315 + C</td>
<td>2</td>
</tr>
<tr>
<td>Hui39</td>
<td>D</td>
<td>223 362</td>
<td>73 263 309 + C 315 + C</td>
<td>2</td>
</tr>
<tr>
<td>Mg41</td>
<td>D</td>
<td>223 362</td>
<td>73 263 309 + C 315 + C</td>
<td>2</td>
</tr>
<tr>
<td>Mg53</td>
<td>D</td>
<td>223 362</td>
<td>73 263 309 + C 315 + C</td>
<td>2</td>
</tr>
<tr>
<td>Mg55</td>
<td>D</td>
<td>223 362</td>
<td>73 263 309 + C 315 + C</td>
<td>2</td>
</tr>
<tr>
<td>Kaz12</td>
<td>D</td>
<td>223 362</td>
<td>73 263 315 + C</td>
<td>2</td>
</tr>
<tr>
<td>Uyg28</td>
<td>D</td>
<td>223 362</td>
<td>73 263 315 + C</td>
<td>2</td>
</tr>
<tr>
<td>Uyg44</td>
<td>D</td>
<td>223 362</td>
<td>73 263 315 + C</td>
<td>2</td>
</tr>
<tr>
<td>Hui112</td>
<td>D</td>
<td>223 319 362</td>
<td>73 263 315 + C</td>
<td>2</td>
</tr>
<tr>
<td>Mg27</td>
<td>D</td>
<td>223 294 362</td>
<td>73 263 315 + C</td>
<td>2</td>
</tr>
<tr>
<td>Kaz59</td>
<td>D</td>
<td>223 287 362</td>
<td>73 263 315 + C</td>
<td>2</td>
</tr>
<tr>
<td>Uzb31</td>
<td>D</td>
<td>223 274 362</td>
<td>73 263 309 + CC (519) 315 + C</td>
<td>2</td>
</tr>
<tr>
<td>Kaz53</td>
<td>D</td>
<td>223 274 362</td>
<td>73 146 152 263 309 + CC 315 + C</td>
<td>2</td>
</tr>
<tr>
<td>Uzb15</td>
<td>D</td>
<td>223 261 278 362 (519)</td>
<td>73 194 263 315 + C</td>
<td>2</td>
</tr>
<tr>
<td>Uzb30</td>
<td>D</td>
<td>223 256 362</td>
<td>73 263 309 + C 315 + C</td>
<td>2</td>
</tr>
<tr>
<td>Mg59</td>
<td>D</td>
<td>223 245 362 368</td>
<td>73 263 309 + C 315 + C</td>
<td>2</td>
</tr>
<tr>
<td>Mg62</td>
<td>D</td>
<td>223 245 362 368</td>
<td>73 263 309 + C 315 + C</td>
<td>2</td>
</tr>
<tr>
<td>Uzb14</td>
<td>D</td>
<td>224 245 292 362 (519)</td>
<td>73 263 309 + C 315 + C</td>
<td>2</td>
</tr>
<tr>
<td>Hui61</td>
<td>D</td>
<td>209 223 362 368 399</td>
<td>73 263 309 + C 315 + C</td>
<td>2</td>
</tr>
<tr>
<td>Uzb13</td>
<td>D</td>
<td>188 223 231 264 362 (519)</td>
<td>73 146 263 309 + C 315 + C</td>
<td>2</td>
</tr>
<tr>
<td>Hui58</td>
<td>D</td>
<td>186 245 362</td>
<td>73 263 315 + C</td>
<td>2</td>
</tr>
<tr>
<td>Hui102</td>
<td>D</td>
<td>184 231 311 362</td>
<td>73 195 263 315 + C 316</td>
<td>2</td>
</tr>
<tr>
<td>Uzb1</td>
<td>D</td>
<td>176 183 223 274 390 319</td>
<td>73 195 263 315 + C 316</td>
<td>2</td>
</tr>
<tr>
<td>Kaz47</td>
<td>D</td>
<td>174 223 362</td>
<td>73 263 309 + C 315 + C</td>
<td>2</td>
</tr>
<tr>
<td>Mg37</td>
<td>D</td>
<td>171 223 311 362</td>
<td>73 263 309 + C 315 + C</td>
<td>2</td>
</tr>
<tr>
<td>Uzb56</td>
<td>D</td>
<td>167 172 174 223 287 362</td>
<td>73 263 309 + C 315 + C</td>
<td>2</td>
</tr>
<tr>
<td>Uyg36</td>
<td>D</td>
<td>129 223 362</td>
<td>73 263 309 + C 315 + C</td>
<td>2</td>
</tr>
<tr>
<td>Kaz15</td>
<td>D</td>
<td>129 223 256 362</td>
<td>73 263 309 + C 315 + C</td>
<td>2</td>
</tr>
<tr>
<td>Uzb4</td>
<td>D</td>
<td>093 223 290 362 (519)</td>
<td>73 146 263 315 + C</td>
<td>2</td>
</tr>
<tr>
<td>Hui57</td>
<td>D</td>
<td>092 223 362</td>
<td>73 263 309 + C 315 + C</td>
<td>2</td>
</tr>
<tr>
<td>Kaz4</td>
<td>D</td>
<td>086 223 362</td>
<td>73 263 309 + C 315 + C</td>
<td>2</td>
</tr>
<tr>
<td>Mg31</td>
<td>D</td>
<td>082 223 362</td>
<td>73 263 309 + C 315 + C</td>
<td>2</td>
</tr>
<tr>
<td>Kaz60</td>
<td>D</td>
<td>042 214 223 362</td>
<td>73 195 263 315 + C</td>
<td>2</td>
</tr>
<tr>
<td>Mg16</td>
<td>D</td>
<td>172 171 215 223 319 362</td>
<td>73 195 263 315 + C</td>
<td>2</td>
</tr>
<tr>
<td>Uzb36</td>
<td>D</td>
<td>093 172 173 215 223 319</td>
<td>73 195 263 309 + C 315 + C</td>
<td>2</td>
</tr>
<tr>
<td>Mg45</td>
<td>D</td>
<td>182 183C 189 223 362</td>
<td>73 195 263 309 + C 315 + C</td>
<td>2</td>
</tr>
<tr>
<td>Uzb20</td>
<td>D</td>
<td>172 183C 189 223 294 362</td>
<td>73 195 263 309 + C 315 + C</td>
<td>2</td>
</tr>
<tr>
<td>Uyg20</td>
<td>D5</td>
<td>189 223 362</td>
<td>73 195 263 309 + C 315 + C</td>
<td>2</td>
</tr>
</tbody>
</table>

**Note:** The table lists Haplogroup variations in 252 samples from the Silk Road Region in China, focusing on HVS-I and HVS-II sequences with specific polymorphisms indicated.
<table>
<thead>
<tr>
<th>Sample</th>
<th>Haplogroup</th>
<th>HVS-I&lt;sup&gt;b&lt;/sup&gt;(16000+)</th>
<th>HVS-II&lt;sup&gt;b&lt;/sup&gt;</th>
<th>9-bp&lt;sup&gt;c&lt;/sup&gt; (10000+)</th>
<th>Other Polymorphisms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uzb59</td>
<td>D5</td>
<td>189 223 362 (519)</td>
<td>73 150 263 309+C 315+C</td>
<td>2</td>
<td>–</td>
</tr>
<tr>
<td>Hui51</td>
<td>D5a</td>
<td>164 172 182C 183C 189</td>
<td>223 266 362</td>
<td>2</td>
<td>–</td>
</tr>
<tr>
<td>Uyg38</td>
<td>D5a</td>
<td>092 164 182C 183C 189</td>
<td>223 266 362</td>
<td>2</td>
<td>–</td>
</tr>
<tr>
<td>Hui106</td>
<td>G</td>
<td>223 362</td>
<td>73 263 309+C 315+C</td>
<td>2</td>
<td>398 400</td>
</tr>
<tr>
<td>Kaz45</td>
<td>G</td>
<td>223 299 362 (519)</td>
<td>73 152 263 309+C 315+C</td>
<td>2</td>
<td>+ + −</td>
</tr>
<tr>
<td>Mg58</td>
<td>G</td>
<td>192A 193 223 260 362</td>
<td>73 182 204 260 263</td>
<td>2</td>
<td>+ +</td>
</tr>
<tr>
<td>Mg65</td>
<td>G</td>
<td>192A 193 223 260 286 362</td>
<td></td>
<td>2</td>
<td>+ +</td>
</tr>
<tr>
<td>Uzb33</td>
<td>G</td>
<td>093 223 362 (519)</td>
<td>73 263 315+C</td>
<td>2</td>
<td>+ +</td>
</tr>
<tr>
<td>Hui53</td>
<td>G1a</td>
<td>223 325 362</td>
<td></td>
<td>2</td>
<td>+ +</td>
</tr>
<tr>
<td>Kaz26</td>
<td>G1a</td>
<td>223 325 362</td>
<td></td>
<td>2</td>
<td>+ +</td>
</tr>
<tr>
<td>Mg36</td>
<td>G2</td>
<td>093 223 278 362</td>
<td></td>
<td>2</td>
<td>+ +</td>
</tr>
<tr>
<td>Mg19</td>
<td>G2</td>
<td>223 234 278 362</td>
<td></td>
<td>2</td>
<td>+ +</td>
</tr>
<tr>
<td>Mg57</td>
<td>G2</td>
<td>223 278 362</td>
<td></td>
<td>2</td>
<td>+ +</td>
</tr>
<tr>
<td>Uzb44</td>
<td>G2a</td>
<td>223 227 278 362</td>
<td></td>
<td>2</td>
<td>+ +</td>
</tr>
<tr>
<td>Hui109</td>
<td>G2a</td>
<td>223 227 278 362</td>
<td></td>
<td>2</td>
<td>+ +</td>
</tr>
<tr>
<td>Uyg42</td>
<td>G2a</td>
<td>223 227 278 362</td>
<td></td>
<td>2</td>
<td>+ +</td>
</tr>
<tr>
<td>Uyg25</td>
<td>G2a</td>
<td>223 227 278 293C 362</td>
<td></td>
<td>2</td>
<td>+ +</td>
</tr>
<tr>
<td>Hui41</td>
<td>G2a</td>
<td>223 227 234 278 362</td>
<td></td>
<td>2</td>
<td>+ +</td>
</tr>
<tr>
<td>Mg24</td>
<td>G2a</td>
<td>223 227 234 278 362</td>
<td></td>
<td>2</td>
<td>+ +</td>
</tr>
<tr>
<td>Mg51</td>
<td>G2a</td>
<td>223 227 234 278 362</td>
<td></td>
<td>2</td>
<td>+ +</td>
</tr>
<tr>
<td>Uyg1</td>
<td>G2a</td>
<td>169 223 272 262 278 294</td>
<td></td>
<td>2</td>
<td>+ +</td>
</tr>
<tr>
<td>Uyg10</td>
<td>G2a</td>
<td>093 223 227 278 362</td>
<td></td>
<td>2</td>
<td>+ +</td>
</tr>
<tr>
<td>Kaz43</td>
<td>G2a</td>
<td>093 223 227 234 278 309</td>
<td>73 152 195 263 315+C</td>
<td>2</td>
<td>+ +</td>
</tr>
<tr>
<td>Uyg11</td>
<td>G2a</td>
<td>093 223 227 234 278 309</td>
<td></td>
<td>2</td>
<td>+ +</td>
</tr>
<tr>
<td>Uzb53</td>
<td>G2a</td>
<td>093 223 227 234 278 289</td>
<td>73 152 263 315+C</td>
<td>2</td>
<td>+ +</td>
</tr>
<tr>
<td>Mg63</td>
<td>G2a</td>
<td>093 189 209 223 227 234</td>
<td>278 309 362 286 463</td>
<td>2</td>
<td>+ +</td>
</tr>
<tr>
<td>Uzb6</td>
<td>G3</td>
<td>189 223 263C 274 362</td>
<td>73 143 152 195 263 315+C</td>
<td>2</td>
<td>398 400</td>
</tr>
<tr>
<td>Uyg8</td>
<td>G3</td>
<td>156 223 274 362 390</td>
<td></td>
<td>2</td>
<td>+ +</td>
</tr>
<tr>
<td>Hui110</td>
<td>M7b</td>
<td>176 223 297</td>
<td></td>
<td>2</td>
<td>+ +</td>
</tr>
<tr>
<td>Hui104</td>
<td>M7b1</td>
<td>129 192 223 297 309 319</td>
<td></td>
<td>2</td>
<td>+ +</td>
</tr>
<tr>
<td>Uyg3</td>
<td>M7b1</td>
<td>129 189 192 223 297</td>
<td></td>
<td>2</td>
<td>+ +</td>
</tr>
<tr>
<td>Hui98</td>
<td>M7b3</td>
<td>086 297 324 399</td>
<td>73 199 263 315+C</td>
<td>2</td>
<td>398 400 497 − + +  12405</td>
</tr>
<tr>
<td>Hui43</td>
<td>M7c</td>
<td>093 223</td>
<td>73 146 199 263 315+C</td>
<td>2</td>
<td>398 400 − + +</td>
</tr>
<tr>
<td>Hui105</td>
<td>M7c</td>
<td>172 223 295</td>
<td></td>
<td>2</td>
<td>− +</td>
</tr>
<tr>
<td>Mg61</td>
<td>M7c</td>
<td>145 223 231 295 304</td>
<td></td>
<td>2</td>
<td>− +</td>
</tr>
<tr>
<td>Uyg29</td>
<td>M7c</td>
<td>223 254 295 362</td>
<td></td>
<td>2</td>
<td>+ +</td>
</tr>
<tr>
<td>Sample</td>
<td>Haplogroup</td>
<td>HVS-I (16000+)</td>
<td>HVS-II</td>
<td>9-bp</td>
<td>Other Polymorphisms</td>
</tr>
<tr>
<td>--------</td>
<td>------------</td>
<td>----------------</td>
<td>---------</td>
<td>-------</td>
<td>---------------------</td>
</tr>
<tr>
<td>Uyg17</td>
<td>M7c</td>
<td>223 294 295</td>
<td>73 146 199 263 315+C</td>
<td>2</td>
<td>+ 3882 4071 4769 4850</td>
</tr>
<tr>
<td>Kaz48</td>
<td>M7c</td>
<td>051 183C 189 234 294 295 (519)</td>
<td>73 152 199 309+CC 315+C</td>
<td>2</td>
<td>+</td>
</tr>
<tr>
<td>Kaz57</td>
<td>C</td>
<td>223 298 327 (519)</td>
<td>73 152 195 249d 263 315+C</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Kaz21</td>
<td>C</td>
<td>223 298 327</td>
<td>73 151 152 249d 263 315+C</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Mg26</td>
<td>C</td>
<td>223 298 318 327 357</td>
<td>73 195 249d 263 309+C 315+C</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Mg25</td>
<td>C</td>
<td>223 298 311 327</td>
<td>73 249d 263 309+C 315+C</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Kaz11</td>
<td>C</td>
<td>223 261 288 298</td>
<td>73 249d 263 309+C 315+C</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Mg60</td>
<td>C</td>
<td>223 259+A 298 327</td>
<td>73 146 249d 263 309+C 315+C</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Hui46</td>
<td>C</td>
<td>189 223 298 327 357</td>
<td>73 249d 263 309+C 315+C</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Mg56</td>
<td>C</td>
<td>171 223 298 327 344 357</td>
<td>73 249d 263 309+C 315+C</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Kaz8</td>
<td>C</td>
<td>148 164 223 288 298 327</td>
<td>73 249d 263 315+C 315+C</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Kaz39</td>
<td>C</td>
<td>129 223 298 327 (519)</td>
<td>73 195 249d 263 315+C</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Uyg24</td>
<td>C</td>
<td>129 223 298 327</td>
<td>73 195 249d 263 315+C</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Mg48</td>
<td>C</td>
<td>129 223 298 319 327</td>
<td>73 195 249d 263 309+CC</td>
<td>315+C</td>
<td></td>
</tr>
<tr>
<td>Uzb7</td>
<td>C</td>
<td>129 150 223 298 327 (519)</td>
<td>73 195 249d 263</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Uyg19</td>
<td>C</td>
<td>126 223 298 327</td>
<td>73 249d 263 309+CC 315+C</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Kaz20</td>
<td>C</td>
<td>093 129 223 298 327</td>
<td>73 195 249d 263 309+C</td>
<td>315+C</td>
<td></td>
</tr>
<tr>
<td>Uyg37</td>
<td>C</td>
<td>093 129 223 298 327</td>
<td>73 249d 263 309+C</td>
<td>315+C</td>
<td></td>
</tr>
<tr>
<td>Kaz18</td>
<td>C</td>
<td>093 129 213 223 298 327</td>
<td>73 195 249d 315+C</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Hui100</td>
<td>Z</td>
<td>185 223 260 298</td>
<td>73 152 249d 263 315+C</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Kaz14</td>
<td>Z</td>
<td>185 223 260 298</td>
<td>73 151 152 249d 263</td>
<td>315+C</td>
<td></td>
</tr>
<tr>
<td>Hui101</td>
<td>Z</td>
<td>185 223 260 298 302</td>
<td>73 151 152 249d 263 315+C</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Kaz27</td>
<td>Z</td>
<td>185 223 260 298 302 399</td>
<td>73 151 152 249d 263 309+C</td>
<td>315+C</td>
<td></td>
</tr>
<tr>
<td>Kaz30</td>
<td>Z</td>
<td>185 223 260 298 302 399</td>
<td>73 151 152 249d 263 309+C</td>
<td>315+C</td>
<td></td>
</tr>
<tr>
<td>Kaz7</td>
<td>Z</td>
<td>185 215 223 260 298 355</td>
<td>73 249d 263 279 315+C</td>
<td>2</td>
<td>15784</td>
</tr>
<tr>
<td>Uzb2</td>
<td>Z</td>
<td>129 219 223 224 260 298 (519)</td>
<td>73 151 152 249d 263 315+C</td>
<td>2</td>
<td></td>
</tr>
</tbody>
</table>
Table 1
Continued

<table>
<thead>
<tr>
<th>Sample</th>
<th>Haplogroup</th>
<th>HVS-I (16000+)</th>
<th>HVS-II (10000+)</th>
<th>9-bp</th>
<th>Other Polymorphisms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kaz10</td>
<td>Z</td>
<td>129 185 223 224 260 298</td>
<td>309-C 315-C</td>
<td>2</td>
<td>15784</td>
</tr>
<tr>
<td>Kaz24</td>
<td>Z</td>
<td>129 185 223 224 260 298</td>
<td>309-C 315-C</td>
<td>2</td>
<td>15784</td>
</tr>
<tr>
<td>Hui37</td>
<td>M8a</td>
<td>223 298 319</td>
<td>2</td>
<td>14470</td>
<td></td>
</tr>
<tr>
<td>Uyg15</td>
<td>M8a</td>
<td>223 298 319</td>
<td>2</td>
<td>14470</td>
<td></td>
</tr>
<tr>
<td>Kaz41</td>
<td>M8a</td>
<td>1414A 223 248 298 319 357</td>
<td>2</td>
<td>14470</td>
<td></td>
</tr>
<tr>
<td>Uyg18</td>
<td>M8a</td>
<td>129 184 189 223 298 319 470</td>
<td>2</td>
<td>14470</td>
<td></td>
</tr>
<tr>
<td>Uyg35</td>
<td>M8a</td>
<td>086 134 184 223 298 319</td>
<td>2</td>
<td>14470</td>
<td></td>
</tr>
<tr>
<td>Hui114</td>
<td>M9a</td>
<td>129 223 234 316 362</td>
<td>2</td>
<td>14470</td>
<td></td>
</tr>
<tr>
<td>Hui42</td>
<td>M9a</td>
<td>223 234 291 316 362</td>
<td>2</td>
<td>14470</td>
<td></td>
</tr>
<tr>
<td>Mg33</td>
<td>M9a</td>
<td>145 148 223 234 294 316</td>
<td>2</td>
<td>14470</td>
<td></td>
</tr>
<tr>
<td>Kaz13</td>
<td>M9a</td>
<td>093 223 234 316 357</td>
<td>2</td>
<td>14470</td>
<td></td>
</tr>
<tr>
<td>Hui50</td>
<td>M11</td>
<td>223</td>
<td>2</td>
<td>14470</td>
<td></td>
</tr>
<tr>
<td>Mg28</td>
<td>M11</td>
<td>223</td>
<td>2</td>
<td>14470</td>
<td></td>
</tr>
<tr>
<td>Hui36</td>
<td>M11</td>
<td>051 173 223</td>
<td>2</td>
<td>14470</td>
<td></td>
</tr>
<tr>
<td>Uzb19</td>
<td>M11</td>
<td>223 292A</td>
<td>2</td>
<td>14470</td>
<td></td>
</tr>
<tr>
<td>Mg50</td>
<td>M13a</td>
<td>145 148 172 188 189 223 381</td>
<td>2</td>
<td>14470</td>
<td></td>
</tr>
<tr>
<td>Hui45</td>
<td>M*</td>
<td>051 209 223 362</td>
<td>2</td>
<td>14470</td>
<td></td>
</tr>
<tr>
<td>Kaz17</td>
<td>M*</td>
<td>214 223 274 311</td>
<td>2</td>
<td>14470</td>
<td></td>
</tr>
<tr>
<td>Kaz38</td>
<td>M*</td>
<td>129 223 265C (519)</td>
<td>2</td>
<td>14470</td>
<td></td>
</tr>
<tr>
<td>Uyg5</td>
<td>M*</td>
<td>223 234 260 293C</td>
<td>2</td>
<td>14470</td>
<td></td>
</tr>
<tr>
<td>Uzb57</td>
<td>M*</td>
<td>169 223 260 295 (519)</td>
<td>2</td>
<td>14470</td>
<td></td>
</tr>
<tr>
<td>Hui113</td>
<td>A</td>
<td>093 126 223 234 235 290 319</td>
<td>2</td>
<td>14470</td>
<td></td>
</tr>
<tr>
<td>Mg24</td>
<td>A</td>
<td>066 223 234 290 293C 319</td>
<td>2</td>
<td>14470</td>
<td></td>
</tr>
<tr>
<td>Mg29</td>
<td>A</td>
<td>189 223 290 319</td>
<td>2</td>
<td>14470</td>
<td></td>
</tr>
<tr>
<td>Mg32</td>
<td>A</td>
<td>223 290 290 319</td>
<td>2</td>
<td>14470</td>
<td></td>
</tr>
<tr>
<td>Kaz52</td>
<td>A</td>
<td>223 290 290 319 (519)</td>
<td>2</td>
<td>14470</td>
<td></td>
</tr>
<tr>
<td>Uyg33</td>
<td>A</td>
<td>223 290 311 319</td>
<td>2</td>
<td>14470</td>
<td></td>
</tr>
<tr>
<td>Uzb5</td>
<td>A</td>
<td>223 249 290 319 362</td>
<td>2</td>
<td>14470</td>
<td></td>
</tr>
<tr>
<td>Uzb29</td>
<td>A</td>
<td>223 290 292A 319 325 362</td>
<td>2</td>
<td>15510 15520</td>
<td></td>
</tr>
<tr>
<td>Hui103</td>
<td>A</td>
<td>223 274 290 319 362</td>
<td>2</td>
<td>15510 15520</td>
<td></td>
</tr>
<tr>
<td>Sample</td>
<td>Haplogroup</td>
<td>HVS-I (16000+)</td>
<td>HVS-II</td>
<td>9-bp^c (10000+)</td>
<td>Other Polymorphisms</td>
</tr>
<tr>
<td>---------</td>
<td>------------</td>
<td>----------------</td>
<td>--------</td>
<td>-----------------</td>
<td>--------------------</td>
</tr>
<tr>
<td>Hu54</td>
<td>A</td>
<td>223 290 319 362</td>
<td></td>
<td>73 150 195 263 315 +C</td>
<td>2</td>
</tr>
<tr>
<td>Uyg32</td>
<td>A</td>
<td>223 290 319 362</td>
<td></td>
<td>73 150 195 263 315 +C</td>
<td>2</td>
</tr>
<tr>
<td>Mg30</td>
<td>A</td>
<td>086 223 290 319 356 362</td>
<td></td>
<td>73 150 195 263 315 +C</td>
<td>2</td>
</tr>
<tr>
<td>Kaz22</td>
<td>A</td>
<td>223 290 294 311 319 362</td>
<td></td>
<td>73 150 195 263 315 +C</td>
<td>2</td>
</tr>
<tr>
<td>Mg21</td>
<td>N9a</td>
<td>129 223 257A 261</td>
<td>73 150 195 263 315 +C</td>
<td>2</td>
<td>5417</td>
</tr>
<tr>
<td>Kaz44</td>
<td>N9a</td>
<td>129 223 257A 261</td>
<td>73 150 195 263 315 +C</td>
<td>2</td>
<td>5417</td>
</tr>
<tr>
<td>Uzb3</td>
<td>N9a</td>
<td>051 172 223 257A 261 325</td>
<td>73 150 195 263 315 +C</td>
<td>2</td>
<td>5417</td>
</tr>
<tr>
<td>Uzb46</td>
<td>N9a</td>
<td>111 129 223 257A 261</td>
<td>73 150 195 263 315 +C</td>
<td>2</td>
<td>5417</td>
</tr>
<tr>
<td>Uzb54</td>
<td>Y1</td>
<td>126 231 266 (519)</td>
<td>73 146 263 309 +C 315 +C</td>
<td>2</td>
<td>5417</td>
</tr>
<tr>
<td>Kaz37</td>
<td>I1</td>
<td>129 172 223 311 391 (519)</td>
<td>73 199 203 204 250 263</td>
<td>2</td>
<td>14766 15043</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>15326 15924</td>
<td></td>
</tr>
<tr>
<td>Uyg4</td>
<td>W</td>
<td>223 292</td>
<td>73 119 146 152 189 195</td>
<td>2</td>
<td>11674 11719 11947</td>
</tr>
<tr>
<td>Uyg7</td>
<td>W</td>
<td>223 292</td>
<td>73 189 194 195 199 204</td>
<td>2</td>
<td>11674 11719 11947</td>
</tr>
<tr>
<td>Uyg9</td>
<td>W</td>
<td>223 292</td>
<td>73 189 194 195 199 204</td>
<td>2</td>
<td>11674 11719 11947</td>
</tr>
<tr>
<td>Uzb40</td>
<td>X</td>
<td>093 189 223 278 (519)</td>
<td>73 153 195 263 315 +C</td>
<td>2</td>
<td>13966 14447</td>
</tr>
<tr>
<td>Uzb9</td>
<td>X</td>
<td>189 261 278 (519)</td>
<td>73 153 189 195 225 263</td>
<td>2</td>
<td>13966 14447</td>
</tr>
<tr>
<td>Uzb26</td>
<td>N*</td>
<td>111 172 183C 189 209 223</td>
<td>73 185 189 209 234 263</td>
<td>2</td>
<td>12705</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>309 +C 315 +C 315 +C</td>
<td></td>
</tr>
<tr>
<td>Uyg49</td>
<td>B</td>
<td>183C 189 278</td>
<td>73 119 146 152 189 195</td>
<td>2</td>
<td>11674 11719 11947</td>
</tr>
<tr>
<td>Hu108</td>
<td>B + a</td>
<td>051 129 182C 183C 189 261</td>
<td>73 185 189 209 234 263</td>
<td>2</td>
<td>12705</td>
</tr>
<tr>
<td>Hu62</td>
<td>B + a</td>
<td>129 182C 183C 189 217 261</td>
<td>73 185 189 209 234 263</td>
<td>2</td>
<td>12705</td>
</tr>
<tr>
<td>Hu63</td>
<td>B + a</td>
<td>129 182C 183C 189 217 261</td>
<td>73 185 189 209 234 263</td>
<td>2</td>
<td>12705</td>
</tr>
<tr>
<td>Hu40</td>
<td>B + a</td>
<td>129 145 182C 183C 189 217</td>
<td>73 185 189 209 234 263</td>
<td>2</td>
<td>12705</td>
</tr>
<tr>
<td>Kaz42</td>
<td>B + a</td>
<td>129 145 182C 183C 189 217</td>
<td>73 185 189 209 234 263</td>
<td>2</td>
<td>12705</td>
</tr>
<tr>
<td>Kaz51</td>
<td>B + a</td>
<td>086 182C 183C 189 217 261</td>
<td>73 185 189 209 234 263</td>
<td>2</td>
<td>12705</td>
</tr>
<tr>
<td>Uzb49</td>
<td>B + b1</td>
<td>136 183C 189 217 261 (519)</td>
<td>73 263 309 +C 315 +C</td>
<td>1</td>
<td>5417</td>
</tr>
<tr>
<td>Uzb24</td>
<td>B + b1</td>
<td>136 183C 189 217 261 (519)</td>
<td>73 263 309 +C 315 +C</td>
<td>1</td>
<td>5417</td>
</tr>
<tr>
<td>Hu47</td>
<td>B + b1</td>
<td>136 183C 189 217 284</td>
<td>73 263 309 +C 315 +C</td>
<td>1</td>
<td>5417</td>
</tr>
<tr>
<td>Mg54</td>
<td>B + b1</td>
<td>086 136 182C 183C 189 217</td>
<td>73 263 309 +C 315 +C</td>
<td>1</td>
<td>5417</td>
</tr>
<tr>
<td>Uzb58</td>
<td>B + c</td>
<td>126 140 182C 183C 189 217</td>
<td>73 150 263 309 +C 315 +C</td>
<td>1</td>
<td>CRS</td>
</tr>
<tr>
<td>Hu107</td>
<td>B + a</td>
<td>140 182C 183C 189 266A 362</td>
<td>73 185 189 209 234 263</td>
<td>2</td>
<td>12705</td>
</tr>
<tr>
<td>Hu60</td>
<td>B + a</td>
<td>140 182C 183C 189 234 266A</td>
<td>73 185 189 209 234 263</td>
<td>2</td>
<td>12705</td>
</tr>
<tr>
<td>Hu59</td>
<td>B + a</td>
<td>093 179 182C 183C 189</td>
<td>73 150 152 263 315 +C</td>
<td>1</td>
<td>13928C 14305</td>
</tr>
<tr>
<td>Hu55</td>
<td>F + a</td>
<td>129 162 172 304</td>
<td>73 150 152 263 315 +C</td>
<td>1</td>
<td>13928C 14305</td>
</tr>
<tr>
<td>Uyg31</td>
<td>F + a</td>
<td>129 162 172 304</td>
<td>73 150 152 263 315 +C</td>
<td>1</td>
<td>13928C 14305</td>
</tr>
<tr>
<td>Sample</td>
<td>Haplogroup</td>
<td>HVS-I (16000+)</td>
<td>HVS-II</td>
<td>9-bp (10000+)</td>
<td>Other Polymorphisms</td>
</tr>
<tr>
<td>-------</td>
<td>------------</td>
<td>----------------</td>
<td>---------</td>
<td>--------------</td>
<td>---------------------</td>
</tr>
<tr>
<td>Mg64</td>
<td>F1a</td>
<td>108 129 162 172 304</td>
<td>73 249d 263 315+C</td>
<td>2</td>
<td>–</td>
</tr>
<tr>
<td>Mg23</td>
<td>F1a</td>
<td>129 150 162 172 261 304</td>
<td>73 198 249d 263 309+C</td>
<td>2</td>
<td>–</td>
</tr>
<tr>
<td>Hui34</td>
<td>F1b</td>
<td>183C 189 304</td>
<td>73 152 249d 263 309+CC 315+C</td>
<td>2</td>
<td>–</td>
</tr>
<tr>
<td>Hui52</td>
<td>F1b</td>
<td>129 162 189 304</td>
<td>73 146 195 249d 263 315+C</td>
<td>2</td>
<td>–</td>
</tr>
<tr>
<td>Kaz3</td>
<td>F1b</td>
<td>092 172 189 304</td>
<td>73 146 249d 263 309+C</td>
<td>2</td>
<td>–</td>
</tr>
<tr>
<td>Mg34</td>
<td>F1b</td>
<td>183C 189 232A 249 304 311</td>
<td>73 249d 263 309+C 315+C</td>
<td>2</td>
<td>–</td>
</tr>
<tr>
<td>Mg39</td>
<td>F1b</td>
<td>183C 189 232A 249 304 311</td>
<td>73 249d 263 309+C 315+C</td>
<td>2</td>
<td>–</td>
</tr>
<tr>
<td>Kaz1</td>
<td>F1b</td>
<td>183C 189 232A 249 304 311</td>
<td>73 204 249d 263 315+C</td>
<td>2</td>
<td>–</td>
</tr>
<tr>
<td>Uyg27</td>
<td>F1b</td>
<td>183C 189 232A 249 304 311</td>
<td>73 249d 263 309+CC</td>
<td>2</td>
<td>–</td>
</tr>
<tr>
<td>Uyg6</td>
<td>F1b</td>
<td>172 183C 189 232A 249 304 311</td>
<td>73 152 204 249d 263</td>
<td>2</td>
<td>–</td>
</tr>
<tr>
<td>Kaz58</td>
<td>F1b</td>
<td>172 182C 183C 189 232A 249 304 311</td>
<td>73 152 204 249d 263</td>
<td>2</td>
<td>–</td>
</tr>
<tr>
<td>Mg43</td>
<td>F1b</td>
<td>172 179 183C 189 232A 249 304 311</td>
<td>73 152 204 249d 263</td>
<td>2</td>
<td>–</td>
</tr>
<tr>
<td>Hui35</td>
<td>F1c</td>
<td>111 129 266 304</td>
<td>73 152 249d 263 309+C 315+C</td>
<td>2</td>
<td>–</td>
</tr>
<tr>
<td>Hui56</td>
<td>F1c</td>
<td>111 129 266 304</td>
<td>73 152 249d 263 309+C 315+C</td>
<td>2</td>
<td>–</td>
</tr>
<tr>
<td>Mg49</td>
<td>F2</td>
<td>304</td>
<td>73 152 234 235 249d 263</td>
<td>309+CC 315+C</td>
<td>2 310 535 586</td>
</tr>
<tr>
<td>Mg52</td>
<td>F2</td>
<td>304</td>
<td>73 152 234 235 249d 263</td>
<td>309+C 315+C</td>
<td>2 310 535 586</td>
</tr>
<tr>
<td>Mg38</td>
<td>F2</td>
<td>304</td>
<td>73 185 249d 263 309+C 315+C</td>
<td>2 310 535 586</td>
<td>+</td>
</tr>
<tr>
<td>Kaz28</td>
<td>F2</td>
<td>304</td>
<td>73 207 249d 263 315+C</td>
<td>315+C</td>
<td>2 310 535 586</td>
</tr>
<tr>
<td>Uzb8</td>
<td>F2a3</td>
<td>093 203 291 304 (519)</td>
<td>73 249d 263 309+C 315+C</td>
<td>2 310 535 586</td>
<td>+</td>
</tr>
<tr>
<td>Hui38</td>
<td>F2a3</td>
<td>093 203 291 304</td>
<td>73 152 249d 263 315+C</td>
<td>2 310 535 586</td>
<td>+ 12338</td>
</tr>
<tr>
<td>Uzb62</td>
<td>F2a3</td>
<td>203 231 291 304 (519)</td>
<td>73 195 249d 263 309+CC 315+C</td>
<td>2 310 535 586</td>
<td>+</td>
</tr>
<tr>
<td>Mg46</td>
<td>F2b</td>
<td>203 304</td>
<td>73 210 249d 263 315+C</td>
<td>1 310 535 586</td>
<td>+ 1005 1824 7828 10810 12338 13708</td>
</tr>
<tr>
<td>Uzb10</td>
<td>F4</td>
<td>126 207 304 362 399 456</td>
<td>73 146 152 249d 263</td>
<td>2</td>
<td>+</td>
</tr>
<tr>
<td>Uzb11</td>
<td>F4</td>
<td>126 207 304 362 399 456</td>
<td>73 146 152 249d 263</td>
<td>2</td>
<td>+</td>
</tr>
<tr>
<td>Kaz6</td>
<td>HV</td>
<td>129</td>
<td>263 309+CC 315+C</td>
<td>2</td>
<td>+</td>
</tr>
<tr>
<td>Uyg51</td>
<td>HV</td>
<td>311</td>
<td>263 309+C 315+C</td>
<td>3</td>
<td>+</td>
</tr>
</tbody>
</table>
Table 1  
Continued

<table>
<thead>
<tr>
<th>Sample</th>
<th>Haplogroup</th>
<th>HVS-I*(16000+)</th>
<th>HVS-II*</th>
<th>9-bp* (10000+)</th>
<th>Other Polymorphisms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uzb21</td>
<td>HV</td>
<td>172 311</td>
<td>263 309+C 315+C</td>
<td>2</td>
<td>+</td>
</tr>
<tr>
<td>Mg47</td>
<td>HV</td>
<td>172 311</td>
<td>263 309+CC 315+C</td>
<td>2</td>
<td>+</td>
</tr>
<tr>
<td>Kaz49</td>
<td>HV1</td>
<td>067 355</td>
<td>150 263 309+C 315+C</td>
<td>2</td>
<td>+</td>
</tr>
<tr>
<td>Kaz50</td>
<td>HV1</td>
<td>067 355</td>
<td>150 263 309+C 315+C</td>
<td>2</td>
<td>+</td>
</tr>
<tr>
<td>Uyg21</td>
<td>HV2</td>
<td>217</td>
<td>73 152 263 315+C</td>
<td>2</td>
<td>+ A G</td>
</tr>
<tr>
<td>Mag20</td>
<td>H</td>
<td>CRS</td>
<td>263 309+CC 315+C</td>
<td>2</td>
<td>-</td>
</tr>
<tr>
<td>Mag35</td>
<td>H</td>
<td>CRS</td>
<td>263 315+C</td>
<td>2</td>
<td>-</td>
</tr>
<tr>
<td>Uzb12</td>
<td>H</td>
<td>CRS</td>
<td>263 315+C</td>
<td>2</td>
<td>-</td>
</tr>
<tr>
<td>Uyg30</td>
<td>H</td>
<td>CRS</td>
<td>152 263 309+C 315+C</td>
<td>2</td>
<td>+</td>
</tr>
<tr>
<td>Uzb42</td>
<td>H</td>
<td>(519)</td>
<td>263 315+C</td>
<td>2</td>
<td>-</td>
</tr>
<tr>
<td>Uzb22</td>
<td>H</td>
<td>(519)</td>
<td>73 263 315+C</td>
<td>2</td>
<td>CRS A G</td>
</tr>
<tr>
<td>Kaz25</td>
<td>H</td>
<td>176</td>
<td>152 263 309+C 315+C</td>
<td>2</td>
<td>-</td>
</tr>
<tr>
<td>Uzb55</td>
<td>H</td>
<td>304</td>
<td>263 315+C</td>
<td>2</td>
<td>-</td>
</tr>
<tr>
<td>Uzb47</td>
<td>H</td>
<td>278 (519)</td>
<td>146 263 309+C 315+C</td>
<td>2</td>
<td>-</td>
</tr>
<tr>
<td>Kaz56</td>
<td>H</td>
<td>362 482</td>
<td>239 263 309+CC 315+C</td>
<td>2</td>
<td>-</td>
</tr>
<tr>
<td>Uyg16</td>
<td>H</td>
<td>362 482</td>
<td>152 239 263 309+CC 315+C</td>
<td>2</td>
<td>- A G</td>
</tr>
<tr>
<td>Kaz40</td>
<td>H</td>
<td>362 482 (519)</td>
<td>239 263 309+C 315+C</td>
<td>2</td>
<td>-</td>
</tr>
<tr>
<td>Kaz16</td>
<td>H</td>
<td>356 362 482</td>
<td>152 239 263 309+CC 315+C</td>
<td>2</td>
<td>- A G</td>
</tr>
<tr>
<td>Uyg39</td>
<td>H</td>
<td>193 354</td>
<td>150 263 309+C 315+C</td>
<td>2</td>
<td>-</td>
</tr>
<tr>
<td>Uyg41</td>
<td>H</td>
<td>145 304</td>
<td>93 263 315+C</td>
<td>2</td>
<td>-</td>
</tr>
<tr>
<td>Uzb47</td>
<td>H</td>
<td>129 249</td>
<td>263 309+C 315+C</td>
<td>2</td>
<td>-</td>
</tr>
<tr>
<td>Uyg23</td>
<td>H</td>
<td>056 266</td>
<td>263 309+CC 315+C</td>
<td>2</td>
<td>-</td>
</tr>
<tr>
<td>Kaz46</td>
<td>V</td>
<td>298</td>
<td>72 263 309+CC 315+C</td>
<td>2</td>
<td>-</td>
</tr>
<tr>
<td>Uzb27</td>
<td>T</td>
<td>126 294 (519)</td>
<td>73 152 194 263 309+CC 315+C</td>
<td>2</td>
<td>463</td>
</tr>
<tr>
<td>Kaz55</td>
<td>T</td>
<td>126 183C 189 294 (519)</td>
<td>73 194 200 263 309+CC 315+C</td>
<td>2</td>
<td>463</td>
</tr>
<tr>
<td>Hui44</td>
<td>T1</td>
<td>126 163 186 189 249 294</td>
<td>73 263 309+CC 315+C</td>
<td>2</td>
<td>463</td>
</tr>
<tr>
<td>Uzb48</td>
<td>T1</td>
<td>126 163 186 189 294</td>
<td>73 263 309+CC 315+C</td>
<td>2</td>
<td>463</td>
</tr>
<tr>
<td>Hui49</td>
<td>T2</td>
<td>126 256 294 296</td>
<td>263 309+C 315+C</td>
<td>2</td>
<td>463</td>
</tr>
<tr>
<td>Kaz2</td>
<td>T2</td>
<td>126 294 296 304</td>
<td>263 309+C 315+C</td>
<td>2</td>
<td>463</td>
</tr>
<tr>
<td>Kaz5</td>
<td>T2</td>
<td>126 294 296 304</td>
<td>263 309+C 315+C</td>
<td>2</td>
<td>463</td>
</tr>
<tr>
<td>Uyg13</td>
<td>T2</td>
<td>126 292 294 296 304</td>
<td>73 263 309+C 315+C 319</td>
<td>2</td>
<td>463</td>
</tr>
<tr>
<td>Uyg34</td>
<td>J</td>
<td>063 069 126 303</td>
<td>73 263 295 309+CC 315+C</td>
<td>2</td>
<td>398</td>
</tr>
<tr>
<td>Uyg32</td>
<td>J1</td>
<td>069 126 145 185 261</td>
<td>73 263 295 309+CC 315+C</td>
<td>2</td>
<td>398</td>
</tr>
<tr>
<td>Uyg32</td>
<td>J1b</td>
<td>069 126 145 222 261 311</td>
<td>73 200 263 295 315+C</td>
<td>2</td>
<td>398</td>
</tr>
<tr>
<td>Uzb52</td>
<td>J1b</td>
<td>069 126 145 222 261 (519)</td>
<td>73 188 295 315+C</td>
<td>2</td>
<td>398</td>
</tr>
<tr>
<td>Uzb17</td>
<td>U2</td>
<td>051 234 264</td>
<td>73 152 263 315+C</td>
<td>2</td>
<td>421</td>
</tr>
<tr>
<td>Uyg43</td>
<td>U2</td>
<td>051 086 259A 291 353</td>
<td>73 146 234 263 315+C</td>
<td>2</td>
<td>+ G A</td>
</tr>
<tr>
<td>Uzb51</td>
<td>U2e</td>
<td>051 129C 183C 193+C 209 309</td>
<td>73 152 199 217 263 315+C</td>
<td>2</td>
<td>+ G A</td>
</tr>
</tbody>
</table>

Note: * denotes haplogroup and 9-bp polymorphisms.
### Table 1

**Continued**

<table>
<thead>
<tr>
<th>Sample*</th>
<th>Haplogroup</th>
<th>HVS-I(16000+)</th>
<th>HVS-IIb</th>
<th>Other Polymorphisms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uyg40</td>
<td>U2e</td>
<td>051 092 129C 183C 189 362</td>
<td>2</td>
<td>G A 12354</td>
</tr>
<tr>
<td>Uyg45</td>
<td>U2e</td>
<td>051 129C 183C 362</td>
<td>2</td>
<td>G A</td>
</tr>
<tr>
<td>Mg18</td>
<td>U4</td>
<td>356</td>
<td>2</td>
<td>G A</td>
</tr>
<tr>
<td>Uyg22</td>
<td>U4</td>
<td>356</td>
<td>2</td>
<td>G A</td>
</tr>
<tr>
<td>Uyg26</td>
<td>U4</td>
<td>356</td>
<td>2</td>
<td>G A</td>
</tr>
<tr>
<td>Uzb50</td>
<td>U4</td>
<td>356 (519)</td>
<td>2</td>
<td>G A</td>
</tr>
<tr>
<td>Uzb39</td>
<td>U4</td>
<td>134 356 (519)</td>
<td>315+C</td>
<td></td>
</tr>
<tr>
<td>Uzb41</td>
<td>U4</td>
<td>134 356 (519)</td>
<td>2</td>
<td>G A</td>
</tr>
<tr>
<td>Uzb28</td>
<td>U4</td>
<td>093 134 356 (519)</td>
<td>2</td>
<td>G A</td>
</tr>
<tr>
<td>Kaz9</td>
<td>U5a</td>
<td>256 270 399</td>
<td>2</td>
<td>G A</td>
</tr>
<tr>
<td>Uzb32</td>
<td>U5a</td>
<td>256 270 399</td>
<td>2</td>
<td>G A 12432</td>
</tr>
<tr>
<td>Mg42</td>
<td>U5b</td>
<td>129 189 270 465</td>
<td>2</td>
<td>G A</td>
</tr>
<tr>
<td>Mg40</td>
<td>U5b</td>
<td>129 189 270 465</td>
<td>2</td>
<td>G A</td>
</tr>
<tr>
<td>Mg44</td>
<td>U5b</td>
<td>129 189 270 465</td>
<td>2</td>
<td>G A</td>
</tr>
<tr>
<td>Kaz54</td>
<td>U5b</td>
<td>182C 183C 189 232 270</td>
<td>2</td>
<td>G A</td>
</tr>
<tr>
<td>Uzb18</td>
<td>U7</td>
<td>069 227 318C 359 (519)</td>
<td>2</td>
<td>G A 12216</td>
</tr>
<tr>
<td>Uyg12</td>
<td>U7</td>
<td>318C 390 400</td>
<td>2</td>
<td>G A</td>
</tr>
<tr>
<td>Uzb61</td>
<td>K</td>
<td>224 311 320 (519)</td>
<td>2</td>
<td>G A</td>
</tr>
<tr>
<td>Uyg14</td>
<td>K</td>
<td>172 224 301 311</td>
<td>2</td>
<td>G A</td>
</tr>
<tr>
<td>Uzb16</td>
<td>R2</td>
<td>071 320 362 (519)</td>
<td>2</td>
<td>CRS + + 12335 – 12800 = CRS</td>
</tr>
<tr>
<td>Uzb43</td>
<td>R*</td>
<td>223</td>
<td>2</td>
<td>CRS + + 12335 – 12800 = CRS</td>
</tr>
<tr>
<td>Hui111</td>
<td>R*</td>
<td>131 213 311</td>
<td>2</td>
<td>CRS + + 12335 – 12800 = CRS</td>
</tr>
</tbody>
</table>

**NOTE.**—Positions are numbered according to the revised Cambridge reference sequence (rCRS) of Andrews et al. (1999); the mtDNAs that have no mutations in a sequenced region compared with the reference sequence are labeled as CRS. All individuals have been sequenced for region 16001–16497 and/or region 30–407, and any mutations spanning outside the region(s) are listed in parentheses. The restriction enzymes used in the present study are designated by the following sign-letter codes: a = AluI, e = HaeIII, f = HhaI, g = HinfI, o = HincII. ‘ ’ and ‘ ’ denote the absence and presence of the restriction site, respectively. When sequence information was not available, items have been left blank.

a The populations Uygur, Uzbek, Kazak, Mongolian, and Hui are abbreviated as Uyg, Uzb, Kaz, Mg, and Hui, respectively.

b Suffixes A, G, C, and T indicate transversions, ‘ ’ indicates deletions, and ‘ ‘ indicates insertions. Indels are recorded at the last possible site.

c ‘ ’ indicates the deletion of the 9-bp (CCCCCTCTA); ‘ ‘ indicates nondeletion.

d The absence of mutation 315+C in HVS-II of Uzb7 was confirmed by two independent experiments.

e These three samples have been reported in Kong et al. (2004).
the Han (8.5%), whereas in the Uzbek, Uygur, Kazak, and Mongolian samples the frequency was less than 5.2%. The differences between the Hui and each of the three samples (Uygur, Kazak, and Mongolian) are statistically significant ($P < 0.05$). The frequencies of haplogroup F are approximately equal in the Hui, Han, and Mongolian samples and are slightly higher than those of the Kazak and Uygur samples. Haplogroup M8 (including M8a, C, and Z) presents the highest frequency in Kazak (26.4%), whereas in the other five samples from the same region, its frequency is lower than 13% (Kazak vs. Uzbek, $P < 0.05$). Haplogroups U and H, the two major western Eurasian haplogroups, reach the highest frequencies in Uzbek (15.5% and 10.3%) and Uygur (14.9% and 10.6%), followed by Mongolian (8.2% and 4.1%) and Kazak (3.8% and 7.5%). However, both of these two haplogroups are absent in the Hui and Han samples. Consequently, the matrilineal genetic structures of these populations are different, more or less, at the lineage level.

Principal Component Map

Figure 3 presents the clustering pattern of 12 populations based on the basal haplogroup frequency matrix that was calculated according to table 2. We aggregated haplogroups Y and N9a together as N9 and haplogroups I, N1a, and N1b as N1 according to phylogenetic information based on mtDNA complete sequences (Kong et al. 2003b; unpublished data). The first two principal components account for 44.4% of the total variation. The Xinjiang Han and Hui samples are clearly separated from the other samples by the first PC, and this distinction was mainly contributed by East Asian prevalent haplogroups F, B, M7, etc. (fig. 3b). The two Mongolian populations occupy the intermediate positions in the PC map, whereas the remaining eight populations are close together on the first PC map. In general, samples with the same ethnic group nomination tend to be closer in the PC map, although there are some variations. For instance, the two Kazak samples (one from Xinjiang and the other from Kazakhstan) are distinct from each other by the second principal component; the Uygur sample from Xinjiang and the Uighur sample from Kazakhstan are also separated from each other, though the latter was said to have migrated from Xinjiang in recent decades (Comas et al. 1998). It is evident that samples believing in the same religion, for example, Uzbek and Hui, did not show a generally
Table 2  
The Haplogroup Distribution Frequencies (%) in the 12 Populations Considered in the Present Study

<table>
<thead>
<tr>
<th>Haplogroup</th>
<th>Hui (N = 45)</th>
<th>Mg (N = 49)</th>
<th>Kaz (N = 53)</th>
<th>Uzb (N = 58)</th>
<th>Uyg (N = 47)</th>
<th>Han (N = 47)</th>
<th>UIG (N = 55)</th>
<th>KIT (N = 48)</th>
<th>KIR (N = 47)</th>
<th>KAZ (N = 55)</th>
<th>IM-Mg (N = 48)</th>
<th>Central Asian (N = 232)</th>
</tr>
</thead>
<tbody>
<tr>
<td>D*</td>
<td>13.3</td>
<td>18.4</td>
<td>13.2</td>
<td>19.0</td>
<td>6.4</td>
<td>19.1</td>
<td>16.4</td>
<td>14.6</td>
<td>25.5</td>
<td>16.4</td>
<td>33.3</td>
<td>13.4</td>
</tr>
<tr>
<td>D5*</td>
<td>2.0</td>
<td>3.4</td>
<td>2.1</td>
<td>2.1</td>
<td>1.8</td>
<td>4.2</td>
<td>0.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D5a</td>
<td>2.2</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>G*</td>
<td>2.2</td>
<td>4.1</td>
<td>1.9</td>
<td>1.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1a</td>
<td>2.2</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>G2*</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>G2a</td>
<td>4.4</td>
<td>6.1</td>
<td>1.9</td>
<td>3.4</td>
<td>10.6</td>
<td>2.1</td>
<td>14.6</td>
<td>2.1</td>
<td>5.5</td>
<td>3.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>G3</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>M7*</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>M7b*</td>
<td>2.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.3</td>
</tr>
<tr>
<td>M7b1</td>
<td>2.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.4</td>
</tr>
<tr>
<td>M7b2</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>M7b3</td>
<td>2.2</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>M7c</td>
<td>4.4</td>
<td>2.0</td>
<td>1.9</td>
<td></td>
<td>4.3</td>
<td>2.1</td>
<td>3.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M8a</td>
<td>2.2</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>M9a</td>
<td>4.4</td>
<td>2.0</td>
<td>1.9</td>
<td></td>
<td>4.3</td>
<td>1.8</td>
<td>4.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M10</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>M11</td>
<td>4.4</td>
<td>2.0</td>
<td></td>
<td></td>
<td>1.7</td>
<td>2.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M11a</td>
<td>2.0</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>M13a</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>M*</td>
<td>2.2</td>
<td>3.8</td>
<td>1.7</td>
<td>2.1</td>
<td>3.6</td>
<td>4.2</td>
<td>4.3</td>
<td>5.5</td>
<td></td>
<td></td>
<td></td>
<td>0.9</td>
</tr>
<tr>
<td>A</td>
<td>6.7</td>
<td>8.2</td>
<td>3.8</td>
<td>3.4</td>
<td>10.6</td>
<td>7.3</td>
<td>6.3</td>
<td>9.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N9a</td>
<td>2.0</td>
<td>1.9</td>
<td>3.4</td>
<td></td>
<td></td>
<td>1.8</td>
<td>6.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.3</td>
</tr>
<tr>
<td>Y1</td>
<td>1.7</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>W</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>X</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>N1a</td>
<td>0.4</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>N1b</td>
<td>1.7</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>I*</td>
<td>1.7</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>I1</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>N*</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>B*</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>B4*</td>
<td>8.9</td>
<td>3.8</td>
<td>1.7</td>
<td></td>
<td></td>
<td>1.8</td>
<td>2.1</td>
<td>2.1</td>
<td>1.8</td>
<td></td>
<td></td>
<td>0.9</td>
</tr>
<tr>
<td>B4a</td>
<td>2.2</td>
<td>2.0</td>
<td>1.7</td>
<td></td>
<td></td>
<td>2.1</td>
<td>1.8</td>
<td>2.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B4b1</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>B4c</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>B5*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.4</td>
</tr>
<tr>
<td>B5a</td>
<td>4.4</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>B5b</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>B6</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>F1a</td>
<td>2.2</td>
<td>4.1</td>
<td>2.1</td>
<td></td>
<td></td>
<td>4.3</td>
<td>1.8</td>
<td>2.1</td>
<td></td>
<td></td>
<td></td>
<td>4.2</td>
</tr>
<tr>
<td>F1b</td>
<td>4.4</td>
<td>6.1</td>
<td>5.7</td>
<td>1.7</td>
<td>4.3</td>
<td>2.1</td>
<td>3.6</td>
<td>2.1</td>
<td></td>
<td></td>
<td></td>
<td>2.6</td>
</tr>
<tr>
<td>F1c</td>
<td>4.4</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>F2*</td>
<td>6.1</td>
<td>1.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.4</td>
</tr>
<tr>
<td>F2a2</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>F2a3</td>
<td>2.2</td>
<td>3.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.4</td>
</tr>
<tr>
<td>F2b</td>
<td>2.0</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>F3</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>F4</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>R9b</td>
<td>2.0</td>
<td>1.9</td>
<td>1.7</td>
<td></td>
<td></td>
<td>2.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.4</td>
</tr>
<tr>
<td>HV*</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>HV1</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>HV2</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>H</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>V</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>T*</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>T1</td>
<td>2.2</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>T2</td>
<td>2.2</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>J*</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>J1*</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>J1b</td>
<td>1.9</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>J2</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>U*</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>U1</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>U2*</td>
<td>1.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.7</td>
</tr>
<tr>
<td>U2e</td>
<td>1.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.9</td>
</tr>
</tbody>
</table>
concordant feature in the PC map. We also constructed a PC map on the basis of the frequency matrix shown in table 2 and obtained a similar clustering pattern (data not shown).

Discussion

Previous studies have attempted to dissect the mtDNAs from Xinjiang and its proximal regions into eastern and western Eurasian pools (Comas et al. 1998; Yao et al. 2000). However, lack of knowledge of the mtDNA phylogeny hindered those efforts to some extent. Our current dissection of mtDNAs from this region according to the well defined mtDNA haplogroup phylogeny (Finnila¨, Lehtonen, and Majamaa 2001; Herrnstadt et al. 2002; Kivisild et al. 2002; Yao et al. 2002a; Kong et al. 2003b) is informative in comparing the matrilineal genetic components of the populations at a detailed lineage level and discerning the origin of the mtDNAs. As shown in tables 1 and 2, nearly all mtDNA lineages identified in samples from Xinjiang, China, consisted of a subset of the haplogroups specific to eastern and western Eurasia. This result was in agreement with the suggestion of an extensively genetic admixture in this region, based on classic genetic data and less resolved mtDNA data (Zhao and Lee 1989; Cavalli-Sforza, Menozzi, and Piazza 1994; Comas et al. 1998, 2004; Yao et al. 2000). The central heartland hypothesis, which is based on an interpretation of Y-chromosome evidence (Wells et al. 2001) and suggests that most of the modern Eurasian lineages are descended from the Central Asian pool, would conflict with the distribution pattern of mtDNA haplogroups in this region and could not account for the absence of some of the eastern or western Eurasian-specific deep-rooting haplogroups in the Central Asian mtDNA pool.

A prominent feature that should be mentioned is the extent to which the admixture varied among these populations, although our comparison of the well resolved matrilineal components of populations from Xinjiang supported the genetic admixture hypothesis. A decreasing tendency of the western Eurasian-specific haplogroup frequency was observed in these populations from the same region, with the highest frequencies present in the earlier inhabitants, such as the Uygur, Uzbek, and Kazak, followed by Mongolian and Hui; the latter two populations have relatively recent migration histories. This distribution pattern was, coincidentally, well in agreement with the historical ethno-origins of these populations (Ge, Wu and Chao 1997; Chen 1999) and thus suggested that the matrilineal genetic structures of the populations bear the

### Table 2

<table>
<thead>
<tr>
<th>Haplogroup</th>
<th>Hui (N = 45)</th>
<th>Mg (N = 49)</th>
<th>Kaz (N = 53)</th>
<th>Uzb (N = 47)</th>
<th>Uyg (N = 47)</th>
<th>Han (N = 55)</th>
<th>UIG (N = 48)</th>
<th>KIT (N = 47)</th>
<th>KIR (N = 47)</th>
<th>KAZ (N = 48)</th>
<th>IM-Mg (N = 232)</th>
</tr>
</thead>
<tbody>
<tr>
<td>U3</td>
<td>2.0</td>
<td>6.9</td>
<td>4.3</td>
<td>5.5</td>
<td>1.8</td>
<td>1.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>U4</td>
<td>1.9</td>
<td>1.7</td>
<td></td>
<td>2.1</td>
<td>1.8</td>
<td>2.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>U5a</td>
<td></td>
<td>1.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>U5b</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>U7</td>
<td>1.7</td>
<td>2.1</td>
<td></td>
<td>5.5</td>
<td>1.8</td>
<td>0.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>K</td>
<td>1.7</td>
<td>2.1</td>
<td></td>
<td>3.6</td>
<td>1.8</td>
<td>0.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R2</td>
<td>1.7</td>
<td>1.8</td>
<td></td>
<td>1.8</td>
<td></td>
<td>0.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.9</td>
</tr>
<tr>
<td>R11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.9</td>
</tr>
<tr>
<td>R*</td>
<td>2.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.9</td>
</tr>
</tbody>
</table>

**Note.**—The populations Uyghur, Uzbek, Kazak, Mongolian, and Hui are abbreviated as Uyg, Uzb, Kaz, Mg, and Hui, respectively.

- Data from Yao et al. (2000a).
- Data from Comas et al. (1998). Kazakh, Uighur, Kirghiz (Sary-Tash), and Kirghiz (Talas) are abbreviated as KAZ, UIG, KIR, KIT, respectively.
- Mongolian from Inner Mongolia, data from Kong et al. (2003a).
- Aggregated samples reported by Comas et al. (2004).
imprints of their migratory histories (Yao et al. 2002a, 2002b; Yao and Zhang 2002). As a result, it is not strange that no western Eurasian lineage was found in the Xinjiang Han sample—a sample with the most recent migration history (Yao et al. 2002a). Intriguingly, the haplogroups that were mainly found in southern Pakistan, India, the Near East/Caucasus region, the Iranian plateau, and the Arabian Peninsula, such as HV2, R2, and U7 (Quintana-Murci et al. 2004), were only present in Uygur and Uzbek, which harbored an approximately equal amount of western Eurasian types. Haplogroup U2e, which is European-specific (cf. Quintana-Murci et al. 2004), was also restricted to these two samples. It thus seems that the Uygur and Uzbek populations were more open to accept external genes than another earlier inhabitant, the Kazaks, and this also explains why the Kazak populations harbored fewer western Eurasian types compared with Uygurs and Uzbeks.

Insights into the ethno-origins of these populations could be further solidified by focusing on the eastern Eurasian-specific components in these samples. The high frequencies of the north-prevalent haplogroups, such as M8, D, and G (Yao et al. 2002a; Kong et al. 2003a), found in Mongolian (46.9%), Kazak (45.3%), Uygur (36.2%), and Uzbek (32.8%) samples were in accordance with their northern origin as well as their current geographic distributions. As demonstrated in figure 3, the two Mongolian samples from Inner Mongolia (Kong et al. 2003a) and Xinjiang presented a close genetic relationship, notwithstanding the long geographic distance between their sampling locations. It thus seems that the Mongolians in Xinjiang, who heavily migrated to this region during the Yuan Dynasty (A.D.1,206–1,368), maintained most of their ancestral matrilineal components while assimilating genes from other ethnic groups.

A similar condition would also apply to other regional samples belonging to the same ethnic group, for instance the Uygur and Uighur. Although these two samples were separated from each other in the PC map (fig. 3a), the percentages of the total eastern Eurasian types (Uygur, 57.4%; Uighur, 56.4%) and western Eurasian types (Uygur, 42.6%; Uighur, 43.6%) were approximately equal. The difference between Uygur and Uighur was significantly smaller than the differences between each of them and the Han or Hui.

Our results constructed a good story about the unique ethnohistory of the Hui. According to historical records (Du and Yip 1993; Ge, Wu, and Chao 1997; Chen 1999), Huis traced their origin to the Persians and Arabs, who migrated to the southeast coast of China (e.g., Fujian and Guangdong Provinces) in the 7th century A.D., and later to the Central Asians, Persians, and Arabs who migrated to China during the Yuan Dynasty for trade and war reasons. At that time, most of the immigrants did not bring their family members and mainly married Han women. The intermarriages between the Hui and Han were further reinforced by the imperial edict in the Ming Dynasty (A.D. 1,368–1,644; Du and Yip 1993; Chen 1999). The prevalent presence of the eastern Eurasian-specific mtDNA haplogroups (93.3%) in the Hui compared with other Muslim samples, if not of recent gene flow from the Han people, would be attributed to the historical Han women contribution. The high frequencies of the south-prevalent haplogroups (e.g., B, 17.8%; F, 13.3%; and M7, 11.1%), the presence of an M7b3 type (Hui98), and other matched types across China in the Hui samples added further tallies to this suggestion and mirrored the trace of the earlier inhabitants along the southeast coast of China (Du and Yip 1993; Chen 1999). In addition, neither Hui nor Han samples harbored any H and U types that were highly present in other populations in Central Asia. At this point, the matrilineal gene pool of the Hui was shaped by their historically encouraged intermarriage with the Han people, and the religion did not have much influence on it.

In short, the peopling in Xinjiang is more complex than expected. Although the coexistence of the eastern and western Eurasian lineages in the ethnic groups from the Silk Road region generally supported the genetic admixture hypothesis, the extent of admixture varied and was more prone to be influenced by economic and political episodes (at least, this applied to the Han and Hui populations). The observed matrilineal genetic structure is thus a reservoir that aggregated all kinds of genetic and demographic effects.

Electronic Database Information

Accession numbers and URLs for the sequence data of mtDNA control region (including HV1 and HV2-II) in this article are as follows: GenBank (http://www.ncbi.nlm.nih.gov/web/Genbank; accession numbers: AY677733–AY678071).

Acknowledgments

We thank Chang Sun and Shi-Fang Wu for technical assistance, and we also thank two anonymous reviewers for helpful comments. The work was supported by grants from the Natural Science Foundation of China (NSFC, No. 30021004), the Chinese Academy of Sciences (KSCX2-SW-2010), the National Key Technologies R&D Program of China, and the Science and Technology Committee of Yunnan Province (2004BA901A07).

Literature Cited


Takashi Gojobori, Associate Editor

Accepted August 12, 2004