

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

Yong-Gang Yao,^{*1} Qing-Peng Kong,^{*†‡1} 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 (Uyгур, 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 Uyгур (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. 2002; Kong et al. 2003b) 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

¹ These authors contributed equally to this work.

Key words: mtDNA, haplogroup, Silk Road, ethnic group, migration, admixture.

E-mail: zhangyp@public.km.yn.cn.

Mol. Biol. Evol. 21(12):2265–2280. 2004

doi:10.1093/molbev/msh238

Advance Access publication August 18, 2004



FIG. 1.—Geographic locations of the samples considered in this study. The populations are abbreviated as follows: Uygur: Uyg, Uzbek: Uzb, Kazakh: KAZ, Mongolian: Mg, Hui (this study): Hui, Han (Yao et al. 2002a): Han, Mongolian from Inner Mongolia, China (Kong et al. 2003a): IM-Mg, Kirghiz: KIR, Kirghiz (Sary-Tash): KIR, and Kirghiz (Talas; Comas et al. 1998): KIT.

populations 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 Huis 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 Huis from Xinjiang (with HVS-I and II and coding region information available; Yao et al. 2002a); 55 Uighurs and 55 Kazaks from Kazakhstan; 47 Sary-Tash Kirghiz and 48 Talas Kirghiz from Kirgizistan (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 (-5176*AluI*), to confirm the D status. Site 4833 (+4831*HhaI*) 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 (Hui98; 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 Hui98 had mutations 199, 9824 (recognizable by +9820*HinfI*), 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 baptize 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/tRNA^{Lys} 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 ($\geq 40\%$; 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 1
Continued

Sample ^a	Haplogroup	HVS-I ^b (16000+)	HVS-II ^b	9-bp ^c	10171-10659 (10000+)	10171-10659						Other o Polymorphisms	
						e	e	f	a	a	g		
Uzb59	D5	189 223 362 (519)	73 150 263 309+C 315+C	2									1 1 1
Hui51	D5a	164 172 182C 183C 189 223 266 362		2									3 4 5 7 9 2 2 2 6 3 8 1 0 8 3 3 4 6 9 3 7 2 2 0 7 0
Uyg38	D5a	092 164 182C 183C 189 223 266 362		2									3 1 1 6 5 0 8 2 6
Hui106	G	223 362	73 263 309+C 315+C	2	398 400			+	+				o
Kaz45	G	223 299 362 (519)	73 152 263 309+C 315+C	2				+	+				
Mg58	G	192A 193 223 260 362	73 182 204 260 263 309+C 315+C	2				+	+				
Mg65	G	192A 193 223 260 286 362		2				+	+				
Uzb33	G	093 223 362 (519)	73 263 315+C	2				+	+				
Hui53	G1a	223 325 362		2				+	+				
Kaz26	G1a	223 325 362		2				+	+				
Mg36	G2	093 223 278 362		2				+	+				
Mg19	G2	223 234 278 362		2				-	+	+			
Mg57	G2	223 278 362		2					+	+			
Uzb44	G2a	223 227 278 362	73 152 263 315+C	2					+	+			
Hui109	G2a	223 227 278 362		2					+	+			
Uyg42	G2a	223 227 278 362		2					+				
Uyg25	G2a	223 227 278 293C 362		2					+				
Hui41	G2a	223 227 234 278 362		2					+	+			
Mg24	G2a	223 227 234 278 362		2					+	+			
Mg51	G2a	223 227 234 278 362		2					+	+			
Uyg1	G2a	169 223 227 262 278 294 318 362		2					+				
Uyg10	G2a	093 223 227 278 362		2					+				
Kaz43	G2a	093 223 227 234 278 309 362	73 152 195 263 315+C	2					+	+			
Uyg11	G2a	093 223 227 234 278 309 362		2					+				
Uzb53	G2a	093 223 227 234 278 289 309 362	73 152 263 315+C	2					+	+			
Mg63	G2a	093 189 209 223 227 234 278 309 362 463		2					+	+			
Uzb6	G3	189 223 265C 274 362	73 143 152 195 263 315+C	2	398 400			-	+	+	+		A G
Uyg8	G3	156 223 274 362 390		2					+	+			
Hui110	M7b	176 223 297		2									+
Hui104	M7b1	129 192 223 297 309 319		2									+
Uyg3	M7b1	129 189 192 223 297		2									+
Hui98	M7b3	086 297 324 399	73 199 263 315+C	2	398 400 497			-	+	+			12405
Hui43	M7c	093 223	73 146 199 263 315+C	2	398 400				-	+	+		
Hui105	M7c	172 223 295		2									+
Mg61	M7c	145 223 231 295 304		2					-	+			
Uyg29	M7c	223 254 295 362		2									+

Table 1
Continued

Sample ^a	Haplogroup	HVS-I ^b (16000+)	HVS-II ^b	9-bp ^c	10171–10659 (10000+)										Other Polymorphisms	
					e	e	f	a	a	g	o	o	o	o		o
Kaz10	Z	129 185 223 224 260 298	73 151 152 249d 263 309+C 315+C	2												15784
Kaz24	Z	129 185 223 224 260 298	73 151 152 249d 263 309+C 315+C	2												15784
Hui37	M8a	223 298 319	73 263 309+C 315+C	2												14470
Uyg15	M8a	223 298 319	73 263 315+C	2	398	400						+				14470
Kaz41	M8a	114A 223 248 298 319 357	(16T) 73 263 309+C 315+C	2												14470
Uyg18	M8a	129 184 189 223 298 319 470 471 473 476T		2												14470
Uyg35	M8a	086 134 184 223 298 319		2												14470
Hui114	M9a	129 223 234 316 362		2								+				
Hui42	M9a	223 234 291 316 362		2								+				
Mg33	M9a	145 148 223 234 294 316		2								+				
Kaz13	M9a	093 223 234 316 357		2								–				
Hui50	M11	223	73 200 215 263 309+C 315+C 318 326	2	398	400						–	+			
Mg28	M11	223	73 198 263 315+C 318 326	2	398	400						–	+			
Hui36	M11	051 173 223	73 146 198 200 204 213 215 263 309+C 315+C 318 326	2	398	400						–	+			
Uzb19	M11	223 292A	73 198 200 215 263 309+C 315+C 318 326	2	235	398	400									
Mg50	M13a	145 148 172 188 189 223 381	73 152 263 315+C	2	398	400	411					–	+			
Hui45	M*	051 209 223 362	73 153 263 315+C	2	398	400						–	+	–		6338 – 6700 = CRS
Kaz17	M*	214 223 274 311	73 146 199 263 315+C	2	400							–	+			
Kaz38	M*	129 223 265C (519)	73 263 315+C	2	398	400						+				
Uyg5	M*	223 234 260 293C	73 150 152 203 204 263 309+C 315+C	1	398	400						–	+	–		15510 15520 15629 15721
Uzb57	M*	169 223 260 295 (519)	65G 73 150 203 204 263 309+CC 315+C	1	199	398	400					+	–			
Hui113	A	093 126 223 234 235 290 319		2								+				
Mg42	A	066 223 234 290 293C 319		2								+	–			
Mg29	A	189 223 290 319		2								+				
Mg32	A	189 223 290 319		2								+				
Kaz52	A	223 290 319 (519)	73 207 235 263 309+C 315+C 318	2								+				
Uyg33	A	223 290 311 319		2								+				
Uzb5	A	223 249 290 319 362	73 152 235 263 309+C 315+C	2								+				
Uzb29	A	223 290 292A 319 325 362	73 152 235 263 298A 309+CC 315+C	2								+				
Hui103	A	223 274 290 319 362		2								+				

**Table 1
Continued**

Sample ^a	Haplogroup	HVS-I ^b (16000+)	HVS-II ^b	9-bp ^c	10171-10659 (10000+)										Other Polymorphisms												
						e	e	f	a	a	g	o	o	o		o											
Mg64	F1a	108 129 162 172 304	73 249d 263 315+C	2																							
Mg23	F1a	129 150 162 172 261 304	73 198 249d 263 309+C 315+C	2																							
Hui34	F1b	183C 189 304	73 152 249d 263 309+CC 315+C	2																							
Hui52	F1b	129 162 189 304	73 146 195 249d 263 315+C	2																							
Kaz3	F1b	092 172 189 304	73 146 249d 263 309+C 315+C	2																							
Mg34	F1b	183C 189 232A 249 304 311	73 249d 263 309+C 315+C	2																							
Mg39	F1b	183C 189 232A 249 304 311	73 249d 263 309+C 315+C	2																							
Kaz1	F1b	183C 189 232A 249 304 311	73 204 249d 263 315+C	2																							
Uyg27	F1b	183C 189 232A 249 304 311	73 249d 263 309+CC 315+C	2																							
Uzb60	F1b	172 183C 189 232A 249 304 311	73 152 204 249d 263 309+C 315+C	2																							
Uyg6	F1b	172 183C 189 232A 249 304 311	73 152 204 249d 263 315+C	2																							
Kaz58	F1b	172 182C 183C 189 232A 249 304 311	73 152 204 249d 263 309+CC 315+C	2																							
Mg43	F1b	172 179 183C 189 232A 249 304 311	73 152 204 249d 263 315+C	2																							
Hui35	F1c	111 129 266 304	73 152 249d 263 309+C 315+C	2																							
Hui56	F1c	111 129 266 304	73 152 249d 263 309+C 315+C	2																							
Mg49	F2	304	73 152 234 235 249d 263 309+CC 315+C	2	310	535	586																				
Mg52	F2	304	73 152 234 235 249d 263 309+C 315+C	2	310	535	586																				
Mg38	F2	304	73 185 249d 263 309+C 315+C	2	310	535	586																				
Kaz28 ^c	F2	304	73 207 249d 263 315+C	2	310	535	586																				
Uzb8	F2a3	093 203 291 304 (519)	73 249d 263 309+C 315+C	2	310	535	586																				
Hui38 ^c	F2a3	093 203 291 304	73 152 249d 263 315+C	2	310	535	586																				
Uzb62	F2a3	203 231 291 304 (519)	73 195 249d 263 309+CC 315+C	2	310	535	586																				
Mg46 ^c	F2b	203 304	73 210 249d 263 315+C	1	310	535	586																				
Uzb10	F4	126 207 304 362 399 456	73 146 152 249d 263 309+C 315+C 317A	2																							
Uzb11	F4	126 207 304 362 399 456	73 146 152 249d 263 309+C 315+C 317A	2																							
Kaz6	HV	129	263 309+CC 315+C	2																							
Uyg51	HV	311	263 309+C 315+C	3																							

**Table 1
Continued**

Sample ^a	Haplogroup	HVS-I ^b (16000+)	HVS-II ^b	9-bp ^c	10171–10659 (10000+)							Other Polymorphisms
					e	e	f	a	g	o	o	
Uzb21	HV	172 311	263 309+C 315+C	2						1 1 1		
Mg47	HV	172 311	263 309+CC 315+C	2						3 4 5 7 9 2 2 2		
Kaz19	HV1	067 355	150 263 309+C 315+C	2						6 3 8 1 0 8 3 3 4		
Kaz50	HV1	067 355	150 263 309+C 315+C	2						6 9 3 7 2 2 0 7 0		
Uyg21	HV2	217	73 152 263 315+C	2						3 1 1 6 5 0 8 2 6		
Mg20	H	CRS	263 309+CC 315+C	2								
Mg35	H	CRS	263 315+C	2								
Uzb12	H	CRS	93 263 315+C	2								
Uyg30	H	CRS	152 263 309+C 315+C	2								
Uzb42	H	(519)	263 315+C	2								
Uzb22	H	(519)	73 263 315+C	2	CRS						A G	
Kaz25	H	176	152 263 309+C 315+C	2								
Uzb55	H	304	263 315+C	2								
Uzb45	H	278 (519)	146 263 309+C 315+C	2								
Kaz56	H	362 482	239 263 309+CC 315+C	2								
Uyg16	H	362 482	152 239 263 309+CC 315+C	2							A G	
Kaz40	H	362 482 (519)	239 263 309+C 315+C	2								
Kaz16	H	356 362 482	152 239 263 309+CC 315+C	2							A G	
Uyg39	H	193 354	150 263 309+C 315+C	2								
Uyg41	H	145 304	93 263 315+C	2								
Uzb47	H	129 249	263 309+C 315+C	2								
Uyg23	H	056 266	263 309+CC 315+C	2								
Kaz46	V	298	72 263 309+CC 315+C	2								
Uzb27	T	126 294 (519)	73 152 194 263 309+CC 315+C	2	463							
Kaz55	T	126 183C 189 294 (519)	73 194 200 263 309+CC 315+C	2	463							
Hui44	T1	126 163 186 189 249 294		2	463							
Uzb48	T1	126 163 186 189 294	73 263 309+CC 315+C	2	463							
Hui49	T2	126 256 294 296		2	463							
Kaz2	T2	126 294 296 304		2	463							
Kaz5	T2	126 294 296 304		2	463							
Kaz29	T2	126 294 296 304 362		2	463							
Uyg13	T2	126 292 294 296 304	73 263 309+C 315+C 319	2	463							
Uyg34	J	063 069 126 303	73 263 295 309+C 315+C	2	398							
Uyg2	J1	069 126 145 185 261	73 263 271 295 315+C	2	398 410A							
Kaz23	J1b	069 126 145 222 261 311	73 200 263 295 315+C	2	398							
Uzb52	J1b	069 126 145 222 261 (519)	73 188 295 315+C	2	398							
Uzb17	U2	051 234 264	73 152 263 315+C	2	421						G A	
Uyg43	U2	051 086 259A 291 353	73 146 234 263 315+C	2							G A 12106 12507	
Uzb51	U2e	051 129C 183C 193+C 209 362 391 (519)	73 152 199 217 263 315+C	2							G A	

Table 1
Continued

Sample ^a	Haplogroup	HVS-I ^b (16000+)	HVS-II ^b	9-bp ^c	10171–10659 (10000+)							Other Polymorphisms
						e	e	f	a	a	g	
Uyg40	U2e	051 092 129C 183C 189 362		2							1 1 1	
Uyg45	U2e	051 129C 183C 362		2							3 4 5 7 9 2 2 2	
Mg18	U4	356		2							6 3 8 1 0 8 3 3 4	
Uyg22	U4	356		2							6 9 3 7 2 2 0 7 0	
Uyg26	U4	356		2							3 1 1 6 5 0 8 2 6	
Uzb50	U4	356 (519)	73 195 263 309+CC 315+C	2								
Uzb39	U4	134 356 (519)	73 195 263 315+C	2								
Uzb41	U4	134 356 (519)	72 73 152 195 263 315+C	2								
Uzb28	U4	093 134 356 (519)	73 152 189 195 263 309+C 315+C	2								
Kaz9	U5a	256 270 399		2								
Uzb32	U5a	256 270 399	73 263 309+C 315+C	2								12432
Mg22	U5b	129 189 270 465		2								
Mg40	U5b	129 189 270 465		2								
Mg44	U5b	129 189 270 465		2								
Kaz54	U5b	182C 183C 189 232 270 (519)	73 150 263 315+C	2								
Uzb18	U7	069 227 318C 359 (519)	73 144A 151 152 203C 263 309+C 315+C	2			–					12216
Uyg12	U7	318C 390 400		2								
Uzb61	K	224 311 320 (519)	73 146 195 263 315+C	2								
Uyg14	K	172 224 301 311		2								
Uzb16	R2	071 320 362 (519)	73 152 263 309+C 315+C	2								
Uzb43	R*	223	103 235 249d 263 309+C 315+C	2	CRS				+	+		12335 – 12800 = CRS
Hui111	R*	131 213 311	73 263 295A 309+C 315+C	2	658			–		+		

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, “d” indicates deletions, and “+” indicates insertions. Indels are recorded at the last possible site.

^c “1” indicates the deletion of the 9-bp (CCCCCTCTA); “2” 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).

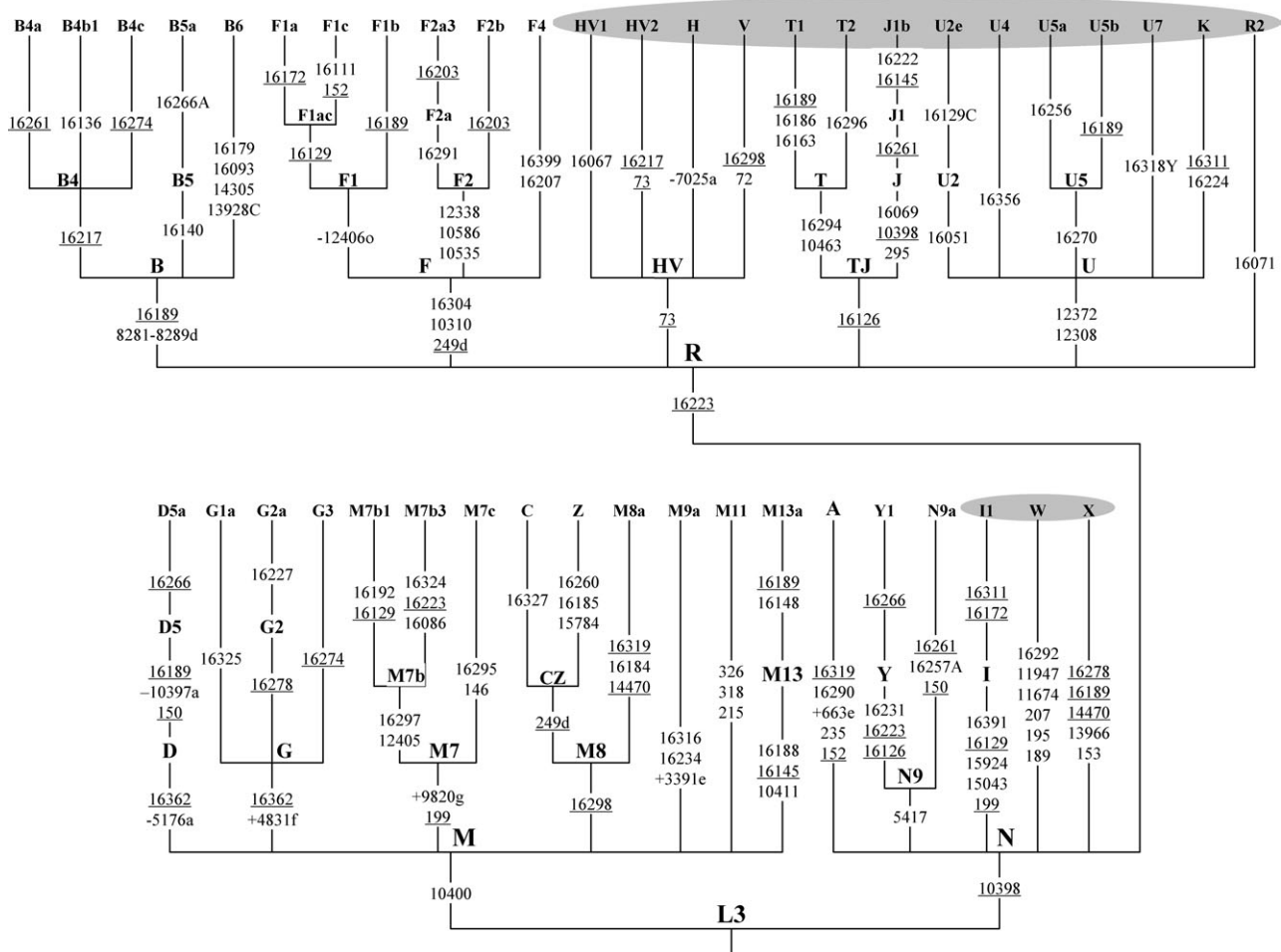


FIG. 2.—Schematic profile of the mtDNA haplogroups in 252 samples from Xinjiang, China. The characteristic mutations (relative to the revised CRS; Andrews et al. 1999) considered here are indicated on the branches in an arbitrary order. The suffix indicates a transversion, “d” indicates deletion, and recurrent mutations are underlined. The restriction enzymes are designated by the following sign-letter codes: a = *AluI*, e = *HaeIII*, f = *HhaI*, g = *HinI*, o = *HincII*. “-” and “+” denote the absence and presence of the restriction site, respectively.

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 belonging 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

Haplogroup	Hui (N = 45)	Mg (N = 49)	Kaz (N = 53)	Uzb (N = 58)	Uyg (N = 47)	Han ^a (N = 47)	UIG ^b (N = 55)	KIT ^b (N = 48)	KIR ^b (N = 47)	KAZ ^b (N = 55)	IM-Mg ^c (N = 48)	Central Asian ^d (N = 232)
D*	13.3	18.4	13.2	19.0	6.4	19.1	16.4	14.6	25.5	16.4	33.3	13.4
D5*		2.0		3.4	2.1	2.1				1.8	4.2	0.4
D5a	2.2				2.1	4.3					2.1	
G*	2.2	4.1	1.9	1.7								
G1a	2.2		1.9									
G2*		6.1					1.8				4.2	
G2a	4.4	6.1	1.9	3.4	10.6	2.1		14.6	2.1	5.5		3.4
G3				1.7	2.1							
M7*						2.1						
M7b*	2.2											1.3
M7b1	2.2				2.1	2.1						
M7b2						4.3						0.4
M7b3	2.2						1.8					
M7c	4.4	2.0	1.9		4.3	2.1	3.6				2.1	0.9
C	2.2	10.2	13.2	1.7	6.4	6.4	1.8	12.5	12.8	7.3	6.3	7.8
Z	4.4		11.3	1.7		2.1			2.1	1.8	4.2	0.9
M8a	2.2		1.9		6.4	4.3						1.3
M9a	4.4	2.0	1.9			4.3	1.8	4.2		1.8	4.2	0.4
M10								2.1			2.1	0.4
M11	4.4	2.0		1.7		2.1						
M13a		2.0										
M*	2.2		3.8	1.7	2.1		3.6	4.2	4.3	5.5		0.9
A	6.7	8.2	3.8	3.4	4.3	10.6	7.3	6.3	2.1	9.1	8.3	3.4
N9a		2.0	1.9	3.4			1.8	6.3			6.3	0.9
Y*												1.3
Y1				1.7		2.1				1.8		
W					6.4					3.6		1.3
X				3.4								
N1a												0.4
N1b												1.7
I*												1.7
I1			1.9									
N*				3.4								
B*					2.1							
B4*							1.8	2.1	2.1	1.8		0.9
B4a	8.9		3.8	$\frac{0}{3.4}$					2.1	1.8	2.1	1.3
B4b1	2.2	2.0				2.1					6.3	0.4
B4c				1.7								
B5*												0.4
B5a	4.4					4.3	5.5		4.3		2.1	
B5b						2.1		2.1		1.8		
B6	2.2											
F1a	2.2	4.1			2.1	4.3	1.8	2.1			4.2	
F1b	4.4	6.1	5.7	1.7	4.3	2.1	3.6	2.1		3.6		2.6
F1c	4.4					2.1	1.8				2.1	
F2*		6.1	1.9									
F2a2						2.1						0.4
F2a3	2.2			3.4		2.1					2.1	0.4
F2b		2.0										
F3						4.3						0.4
F4				3.4		2.1						
R9b						2.1	1.8				2.1	0.4
HV*		2.0	1.9	1.7	2.1			2.1	6.4	5.5		1.3
HV1			3.8						2.1			0.4
HV2					2.1							1.3
H		4.1	7.5	10.3	10.6		10.9	8.3	21.3	16.4		16.8
V			1.9									0.9
T*			1.9	1.7			1.8					
T1	2.2			1.7				2.1		5.5	2.1	0.9
T2	2.2		5.7		2.1					1.8		6.5
J*					2.1			2.1	2.1			0.4
J1*					2.1							2.2
J1b			1.9	1.7				2.1				1.7
J2									4.3			0.9
U*												0.4
U1							1.8	2.1	4.3			3.4
U2*				1.7	2.1		3.6			1.8		1.7
U2e				1.7	4.3							0.9

Table 2
Continued

Haplogroup	Hui (N = 45)	Mg (N = 49)	Kaz (N = 53)	Uzb (N = 58)	Uyg (N = 47)	Han ^a (N = 47)	UIG ^b (N = 55)	KIT ^b (N = 48)	KIR ^b (N = 47)	KAZ ^b (N = 55)	IM-Mg ^c (N = 48)	Central Asian ^d (N = 232)
U3												0.9
U4		2.0		6.9	4.3		5.5			1.8		1.3
U5*		6.1						2.1				
U5a			1.9	1.7						1.8		2.2
U5b			1.9									
U7				1.7	2.1		5.5					0.4
K				1.7	2.1		3.6					3.0
R2				1.7			1.8			1.8		0.4
R5												0.9
R11												0.9
R*	2.2						9.1	4.2	2.1			0.9

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

^a Data from Yao et al. (2002a).

^b Data from Comas et al. (1998). Kazakh, Uighur, Kirghiz (Sary-Tash), and Kirghiz (Talas) are abbreviated as KAZ, UIG, KIR, KIT, respectively.

^c Mongolian from Inner Mongolia, data from Kong et al. (2003a).

^d Aggregated samples reported by Comas et al. (2004).

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 (Finnilä, 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

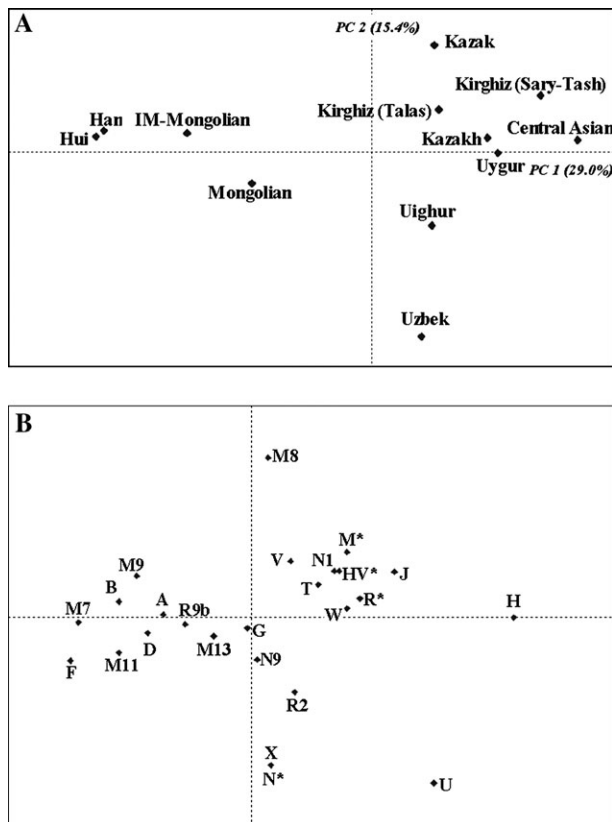


FIG. 3.—(a) Principal component (PC) map of the populations in the Silk Road region based on the basal haplogroup frequency matrix according to table 2, and (b) haplogroup contribution to the population clustering pattern observed in (a).

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 HVS-I and HVS-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

- Andrews, R. M., I. Kubacka, P. F. Chinnery, R. N. Lightowlers, D. M. Turnbull, and N. Howell. 1999. Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat. Genet.* **23**:147.
- Cavalli-Sforza, L. L., P. Menozzi, and A. Piazza. 1994. *The history and geography of human genes*. Princeton University Press, Princeton, New Jersey.
- Chen, L. K. 1999. *Zhongguo minzushi gangyao* (The compendium of Chinese nationality histories). China Financial & Economic Press, Beijing, China.
- Comas, D., F. Calafell, E. Mateu et al. (12 co-authors). 1998. Trading genes along the silk road: mtDNA sequences and the origin of central Asian populations. *Am. J. Hum. Genet.* **63**:1824–1838.
- Comas, D., S. Plaza, R. S. Wells, N. Yuldaseva, O. Lao, F. Calafell, and J. Bertranpetit. 2004. Admixture, migrations, and dispersals in Central Asia: evidence from maternal DNA lineages. *Eur. J. Hum. Genet.* **12**:495–504.

- Du, R., and V. F. Yip. 1993. Ethnic groups in China. Science Press, Beijing, China.
- Fedorova, S. A., M. A. Bermisheva, R. Villems, N. R. Maksimova, and E. K. Khusnutdinova. 2003. Analysis of mitochondrial DNA haplotypes in yakut population. *Mol. Biol. (Mosk)* **37**:643–653.
- Finnilä, S., M. S. Lehtonen, and K. Majamaa. 2001. Phylogenetic network for European mtDNA. *Am. J. Hum. Genet.* **68**: 1475–1484.
- Ge, J. X., S. D. Wu, and S. J. Chao. 1997. *Zhongguo yimin shi* (The migration history of China). Fujian People's Press, Fuzhou, China.
- Herrnstadt, C., J. L. Elson, E. Fahy et al. (11 co-authors). 2002. Reduced-median-network analysis of complete mitochondrial DNA coding-region sequences for the major African, Asian, and European haplogroups. *Am. J. Hum. Genet.* **70**: 1152–1171.
- Kivisild, T., H. V. Tolk, J. Parik, Y. Wang, S. S. Papiha, H. J. Bandelt, and R. Villems. 2002. The emerging limbs and twigs of the East Asian mtDNA tree. *Mol. Biol. Evol.* **19**: 1737–1751.
- Kong, Q. P., Y. G. Yao, M. Liu, S. P. Shen, C. Chen, C. L. Zhu, M. G. Palanichamy, and Y. P. Zhang. 2003a. Mitochondrial DNA sequence polymorphisms of five ethnic populations from northern China. *Hum. Genet.* **113**:391–405.
- Kong, Q. P., Y. G. Yao, C. Sun, H. J. Bandelt, C. L. Zhu, and Y. P. Zhang. 2003b. Phylogeny of East Asian mitochondrial DNA lineages inferred from complete sequences. *Am. J. Hum. Genet.* **73**:671–676.
- Kong, Q. P., Y. G. Yao, C. Sun, C. L. Zhu, L. Zhong, C. Y. Wang, W. W. Cai, X. M. Xu, A. L. Xu, and Y. P. Zhang. 2004. Phylogeographic analysis of mitochondrial DNA haplogroup F2 in China reveals T12338C in the initiation codon of the ND5 gene not to be pathogenic. *J. Hum. Genet.* **49**:414–423.
- Macaulay, V., M. Richards, E. Hickey, E. Vega, F. Cruciani, V. Guida, R. Scozzari, B. Bonne-Tamir, B. Sykes, and A. Torroni. 1999. The emerging tree of West Eurasian mtDNAs: a synthesis of control-region sequences and RFLPs. *Am. J. Hum. Genet.* **64**:232–249.
- Pakendorf, B., V. Wiebe, L. A. Tarskaia, V. A. Spitsyn, H. Soodyall, A. Rodewald, and M. Stoneking. 2003. Mitochondrial DNA evidence for admixed origins of central Siberian populations. *Am. J. Phys. Anthropol.* **120**:211–224.
- Quintana-Murci, L., R. Chaix, R. S. Wells et al. (17 co-authors). 2004. Where west meets east: the complex mtDNA landscape of the southwest and Central Asian corridor. *Am. J. Hum. Genet.* **74**:827–845.
- Richards, M., V. Macaulay, E. Hickey et al. (37 co-authors). 2000. Tracing European founder lineages in the Near Eastern mtDNA pool. *Am. J. Hum. Genet.* **67**:1251–1276.
- Richards, M. B., V. A. Macaulay, H. J. Bandelt, and B. C. Sykes. 1998. Phylogeography of mitochondrial DNA in western Europe. *Ann. Hum. Genet.* **62**:241–260.
- Tajima, A., C. S. Sun, I. H. Pan, T. Ishida, N. Saitou, and S. Horai. 2003. Mitochondrial DNA polymorphisms in nine aboriginal groups of Taiwan: implications for the population history of aboriginal Taiwanese. *Hum. Genet.* **113**: 24–33.
- Wells, R. S., N. Yuldashewa, R. Ruzibakiev et al. (27 co-authors). 2001. The Eurasian heartland: a continental perspective on Y-chromosome diversity. *Proc. Natl. Acad. Sci. USA* **98**: 10244–10249.
- Yao, Y. G., Q. P. Kong, H. J. Bandelt, T. Kivisild, and Y. P. Zhang. 2002a. Phylogeographic differentiation of mitochondrial DNA in Han Chinese. *Am. J. Hum. Genet.* **70**: 635–651.
- Yao, Y. G., Q. P. Kong, X. Y. Man, H. J. Bandelt, and Y. P. Zhang. 2003. Reconstructing the evolutionary history of China: a caveat about inferences drawn from ancient DNA. *Mol. Biol. Evol.* **20**:214–219.
- Yao, Y. G., X. M. Lu, H. R. Luo, W. H. Li, and Y. P. Zhang. 2000. Gene admixture in the silk road region of China: evidence from mtDNA and melanocortin 1 receptor polymorphism. *Genes Genet. Syst.* **75**:173–178.
- Yao, Y. G., L. Nie, H. Harpending, Y. X. Fu, Z. G. Yuan, and Y. P. Zhang. 2002b. Genetic relationship of Chinese ethnic populations revealed by mtDNA sequence diversity. *Am. J. Phys. Anthropol.* **118**:63–76.
- Yao, Y. G., W. S. Watkins, and Y. P. Zhang. 2000. Evolutionary history of the mtDNA 9-bp deletion in Chinese populations and its relevance to the peopling of east and southeast Asia. *Hum. Genet.* **107**:504–512.
- Yao, Y. G., and Y. P. Zhang. 2002. Phylogeographic analysis of mtDNA variation in four ethnic populations from Yunnan Province: new data and a reappraisal. *J. Hum. Genet.* **47**:311–318.
- . 2003. Pitfalls in the analysis of ancient human mtDNA. *Chinese Sci. Bull.* **48**:826–830.
- Zhao, T. M., and T. D. Lee. 1989. Gm and Km allotypes in 74 Chinese populations: a hypothesis of the origin of the Chinese nation. *Hum. Genet.* **83**:101–110.

Takashi Gojobori, Associate Editor

Accepted August 12, 2004