Science in China Series C: Life Sciences

www.scichina.com
life.scichina.com
www.springerlink.com

Gene admixture in ethnic populations in upper part of Silk Road revealed by mtDNA polymorphism

YANG LiuQi, TAN SiJie, YU HaiJing, ZHENG BingRong, QIAO EnFa, DONG YongLi, ZAN RuiGuang & XIAO ChunJie†

Key Laboratory of Bioresources Conservation and Utilization & Human Genetics Center, Yunnan University, Kunming 650091, China

To evaluate the gene admixture on the current genetic landscape in Gansu Corridor (GC) in China, the upper part of the ancient Silk Road which connects the Eastern and Central Asia, we examined mitochondrial DNA (mtDNA) polymorphisms of five ethnic populations in this study. Using PCR-RFLP and sequencing, we analyzed mtDNA haplotypes in 242 unrelated samples in three ethnic populations from the GC region and two ethnic populations from the adjacent Xinjiang Uygur Autonomous Region of China. We analyzed the data in comparison with the previously reported data from Eastern, Central and Western Asia and Europe. We found that both European-specific haplogroups and Eastern Asian-specific haplogroups exist in the Gansu Corridor populations, while a modest matrilineal gene flow from Europeans to this region was revealed. The Gansu Corridor populations are genetically located between Eastern Asians and Central Asians, both of who contributed significantly to the maternal lineages of the GC populations. This study made the landscape of the gene flow and admixture along the Silk Road from Europe, through Central Asia, to the upper part of the Silk Road more complete.

polymorphism, mtDNA, gene admixture, Gansu Corridor

In the 2nd century B.C., the Chinese established a trade route (the Silk Road) from the ancient ChangAn city (now Xi'an city, Shanxi province of China), passing through Gansu Corridor, Xinjiang Uygur Autonomous Region of China, Central and Western Asia, to the Mediterranean Sea, which later became one of the world's oldest and most historically important trade routes. It has not only influenced the culture of China, Central Asia and the West, but also played an important role in human population migration, movement and the gene admixtures of Eastern Asians and Europeans.

Genetic structure studies on the populations along the Silk Road have provided extensive data on gene admixture in neighborhood of the Road. The matrilineal genetic structures of many populations living in Western Asia, Central Asia and Xinjiang Uygur Autonomous Region of China showed evidence that Eastern Asian and European gene lineages flowed to Central Asia and gene admixture had arisen between Europeans and Eastern Asians in Central Asia[1−5].

The Gansu Corridor (GC) is the upper part of the Silk Road which includes part of Gansu and Qinghai provinces in China (Figure 1). It is a corridor connecting Central Asia and Eastern China. Human population migrations such as the ancient Huihus (a population lived in northeast of Asia) and Mongol migrations from east to west along the upper part of Silk Road[6] had been recorded and the gene admixture of Central Asia and Eastern China has been inferred. However, the genetic admixture of human populations in this region has never been estimated despite of the important geographic location of the GC. Therefore, we hypothesize that significant genetic exchanges have occurred between Eastern Asians

Received March 31, 2007; accepted January 7, 2008
doi: 10.1007/s11427-008-0056-2
†Corresponding author (email: cjxiao@public.km.yn.cn)
Supported by the State Key Basic Research Project (973 Project) of China (Grant No. 2006CB708502) and National Natural Science Foundation of China (Grant No. 30660076)

and Central Asians in this region.

In order to test this hypothesis, in this study we collected samples of three ethnic populations from the GC region and analyzed the mtDNA sequence variations by comparing them with two ethnic populations from Xinjiang Uygur Autonomous Region of China (Figure 1) as well as the reported samples from Asia and Europe[1,2,7 – 10] as reference populations. This study documents and interprets the genetic structure of the ethnic populations in this region, which, together with other studies, will make the mtDNA landscape of the Silk Road region more complete.

1 Material and methods

1.1 Population Samples

Samples from five ethnic populations in Gansu, Qinghai, and Xinjiang Uygur Autonomous Region of China were collected. These samples were 50 Baoans from Jishishan County and 64 Dongxiangs from Dongxiang County of Gansu province, 48 Tus from Huzu County of Qinghai province, which belong to the Gansu Corridor region, and 50 Uygurs from Shufu County and 30 Uzbeks from Yining City of Xinjiang Uygur Autonomous Region (Figure 1). All individuals were confirmed to be unrelated through at least three generations, and grandmothers' birthplaces were in the same province. All participates gave informed consent. Human sample handling was approved by the Yunnan University scientific review board.

To set our data in an appropriate comparative framework, we included the data from four Han populations from Yao et al.[7], one Uygur (designated as Uygur1) and one Kazak (Kazak1) from Yao et al.[5], one Uygur (Uygur2), one Kazak (Kazak2) and two Kirkiz (KIT and KIR) from Comas et al.[2], one Tuscan population from Torroni et al.[8], one Sicilian population from Francalacci et al.[9] and one Slovenian population from Malyarchuk et al.[10] in our analyses.

1.2 mtDNA Haplotype Determination

Total DNA was extracted from whole blood by the standard phenol/chloroform methods. Both the hypervariable segment 1 (HVS-1) motif and the coding region variations in mitochondrial DNA (mtDNA) were used to infer haplogroups. The HVS-1 region was amplified by primers L15996 and H16498[7]. The purified PCR products were sequenced using the ABI PRISM BigDye Terminator V3.1 Sequencing Kit and an ABI 377 genetic analyzer (Applied Biosystems). Sixteen sites containing haplogroup diagnostic polymorphisms in the coding region (listed in Table 1) were amplified with primers according to Kivisild et al.[11], Macaulay et al.[12] and Yao et al.[7]. The polymorphisms in these regions were then determined with restriction enzymes as listed in Table 1. The major Eurasian mtDNA haplogroups identified by these polymorphisms were listed in Table 2.

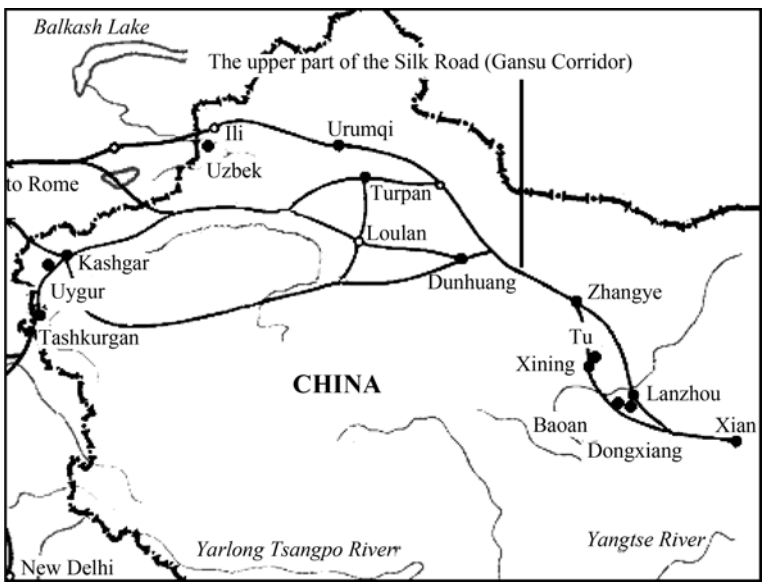

Figure 1 Sketch map of the upper part of the Silk Road and locations of population sampled.

Table 1 PCR-RFLP primer sequence, T_m and the definition of polymorphism.

Site	Primer (sense and antisense)	T_m	Polymorphism
10394/10397	AAACAACTAACCTGCCACTA	60	10394*Dde*I$^+$
	GAAGTGAGATGGTAAATGCT		10397*Alu*I$^+$
5176	AACTTAAACTCCAGCACCAC	55	*Alu*I$^-$
	TAGGTAGGAGTAGCGTGGTA		
4831	AACATGCTAGCTTTTATTCCAG	55	*Hha*I$^+$
	AGAAGAAGCAGTCCGGATGT		
3391	CCAACCTCCTACTCCTCATTG	53	*Hae*III$^+$
	TGGGACCTTTGCGTAGTTGTA		
4715	CAAGTATTTCCTCACGCAAGCA	57	*Hae*III$^+$
	GGTAGTATTGGTTATGGTTCATGG		
9824	CATTTCCGACGGCATCTAC	56	*Hin*fI$^+$
	GGTGGATTTTTCTATGTAGCC		
7598	AAAACCATTTCATAACTTTGTC	56	*Hha*I$^-$
	GTGTTAGGAAAAGGGCATAC		
663	AATTTTATCTTTTGGCGGTATG	53	*Hae*III$^+$
	TTGATGCTTGTCCCTTTTGA		
5417	ATCTCGCACCTGAAACAAGCTAAC	56	*Rsa*I$^-$
	TGGGGTGGGTTTTGTATGTTCGTA		
12406	AACCACCCTAACCCTGACTT	55	*Hpa*I$^-$
	AGTATGGTAATTAGGAAGATGAG		
10646	AGGAATAATACTATCGCTGTT	55	*Rsa*I$^+$
	TAGTCTAGGCCATATGTGTTG		
10310	TGAGCCCTACAAACAACTAACAT	54	*Nla*III$^-$
	ATACCAATTCGGTTCAGTCTAATC		
9 bp	ATGCTAAGTTAGCTTTACAG	57	
	ACAGTTTCATGCCCATCGTC		
14766	CATTATTCTCGCACGGACTAC	55	*Mse*I$^-$
	TGGTTAACTGATTTTATTAGGG		
7025	ACTCCACGGAAGCAATATGA	55	*Alu*I$^-$
	CTATGATGGCAAATACTGCTC		
4216	CTACTACAACCCTTCGCTGAC	55	*Nla*III$^+$
	GATAGGTGGCACGGAGAATT		

Note: The polymorphic sites and restriction endonucleases were based on Macaulay et al.[12], Kivisild et al.[11] and Yao et al.[7].

Table2 RFLP polymorphisms used to identify major Eurasian mtDNA haplogroups

Haplogroup	Characteristic restriction site
Eastern Asian	
M	10394*Dde*I$^+$, 10397*Alu*I$^+$
A	663*Hae*III$^+$
B	9-bp deletion
D	10394*Dde*I$^+$, 10397*Alu*I$^+$, 5176*Alu*I$^-$
G2a	10394*Dde*I$^+$, 10397*Alu*I$^+$, 7598*Hha*I$^-$
F	10310*Nla*III$^-$, 12406*Hpa*I$^-$
G	10394*Dde*I$^+$, 10397*Alu*I$^+$, 4831*Hha*I$^+$
M7	10394*Dde*I$^+$, 10397*Alu*I$^+$, 9824*Hin*fI$^+$
M8	10394*Dde*I$^+$, 10397*Alu*I$^+$, 4715*Hae*III$^+$
M9	10394*Dde*I$^+$, 10397*Alu*I$^+$, 3391*Hae*III$^+$
M10	10394*Dde*I$^+$, 10397*Alu*I$^+$, 10646*Rsa*I$^+$
R9	10310*Nla*III$^-$
N9	5417*Rsa*I$^-$
European	
HV	14766*Mse*I$^-$
H	14766*Mse*I$^-$, 7025*Alu*I$^-$
JT	4216*Nla*III$^+$

Note: The polymorphic identification was based on Yao et al.[7] and Malyarchuk et al.[10].

If a haplogroup could not be defined into either macrohaplogroup M or subgroups A, R, N9, HV or JT of macrohaplogroup N by these site were further amplified by primers L12188 (5′-CTTACGACCCCTTATTTACCG-3′) and H12472 (5′-ATAAAGGTGGATGCGACAATG-3′) for diagnosing polymorphism 12308 and sequenced. Furthermore, if a haplogroup could not be inferred by the sites in HVS-1 motif and the coding region, the HVS-2 region of mtDNA was amplified by primers L29 and H408 [7] and sequenced to determine haplogroups. PCR and RFLP experiment conditions followed Kivisild et al.[11], Macaulay et al.[12] and Yao et al.[7].

1.3 Statistical analyses

DNA sequences were edited and aligned by DNASTAR software (DNASTAR, Inc.) and were compared with the revised Cambridge reference sequence (CRS)[13]. Principal-component analysis (PCA) was conducted using mtDNA haplogroup frequencies and SPSS11.5 software (SPSS). Results of PCA are presented by the maps of the first two PCs. Pairwise F_{ST} values were computed in ARLEQUIN version 2.000[14], using Tamura/Nei distance[15]. The Neighbour-Joining tree analysis (NJA) based on the F_{ST} distance was constructed using Phylip 3.62 software. Admix 2.0 software[16] was used to estimate the degree of admixture of Eastern Asians and Central Asians in the five ethnic populations (Baoan, Dongxiang, Tu, Uygur and Uzbek), using the method of Roberts and Hiorns (hereafter referred to as "RH")[17].

Since selection of reference populations is critical for appropriate estimation of admixture proportions[18,19], we selected reference populations with extant geographic coverage in Eastern Asia and Central Asia. The average haplogroup frequencies of four Han populations (Guangdong, Yunnan, Shandong and Liaoning[7]) were used as the eastern parental populations (Eastern Asians [EAs]), and the average haplogroup frequencies of Uygur, Kazak[5], Kazak, Uygur and two Kirkiz[2] as the western parental populations (Central Asians [CAs]).

2 Results

2.1 Haplogroup profile distribution

Twenty-seven haplogroups were identified for 242 individuals in this study and listed in Table 3. All of them belong to macrohaplogroup M (M*) and N (N*). The haplogroups in the five populations (Baoan, Tu, Dongxiang, Uygur and Uzbek) are shaped mainly by the East Asian and European gene lineages (Table 3; Figure 2). Both West Eurasian (i.e., European) dominant haplogroups[20,21] and East Eurasian (i.e., Eastern Asian) dominant haplogroups[7,12,22] were observed in each of the five ethnic populations. The Eastern Asian-specific Haplogroups A, B, D, F and M8 were the predominant haplogroups in the three GC populations: Baoan, Dongxiang and Tu. Both Eastern Asian-specific A, D, M8 and European-specific H, J and U were found with high frequencies in Uygur and Uzbek samples. In contrast, among the reference populations, only a few European types were reported in two Han populations and no Eastern Asian type has been found in the three European populations (Tuscan, Slovenian and Sicilian). The frequencies of Eastern Asian and European mtDNA types in the three GC populations (Baoan, Dongxiang and Tu) are between those of EAs (four Hans) and Central Asians (Uygur and Uzbek). In addition, haplogroup A presents in Tu with a high frequency (43.8%) and haplogroup T presents in Uygur with a relatively high frequency (24.0%).

2.2 PC maps for mtDNA data

In order to reveal the genetic relationships among the populations, the principal component analysis on the total 18 populations was carried out. The PC map in Figure 3 for the first two principal components account for 69.28% of the total genetic variations. The four Han populations representing EAs locate in the bottom right corner in the PC map in which the two southern Hans (Guangdong and Yunnan) are close to each other, and the two northern Hans (Shandong and Liaoning) group together. The three European populations (Tuscan, Slovenian and Sicilian) occupy the top left corner. The three Uygurs, two Kirkiz and Uzbek representing Cas are located between EAs and Europeans (EUs), and the populations from GC are distributed between the Cas and EAs, whereas Baoan and Dongxing are close to northern Hans. Tu locates relatively independent in the vicinity of Han, KIT, KIR and Kazak. Hence, the PC map indicates that the three GC populations are between CAs and EAs.

2.3 Phylogenetic tree for mtDNA data

An unrooted Neighbor-Joining (NJ) tree of the five populations and 13 reference populations was constructed based on the mtDNA data using Phylip 3.62 software (Figure 4). The genetic affinities are reflected

Table 3 Haplogroup frequency in this study and those of Eurasian populations from previous studies for comparison

Populations (Haplogroups)	BA	Tu	DX	Uzb	Uy	GH[a)]	YH[a)]	SH[a)]	LH[a)]	Sic[b)]	Tus[c)]	Slo[d)]	Uyg1[e)]	Kaz1[e)]	Kaz2[f)]	Uyg2[f)]	Kit2[f)]	Kir2[f)]
	(50)	(48)	(64)	(30)	(50)	(30)	(43)	(50)	(51)	(49)	(48)	(104)	(47)	(53)	(55)	(55)	(48)	(47)
A	8.0	43.8	17.2	3.3	6.0	0.0	4.7	4.0	5.9	0.0	0.0	0.0	4.3	3.8	9.1	7.3	6.3	2.1
B	14.0	2.1	6.3	0.0	2.0	20.0	18.6	12.0	17.6	0.0	0.0	0.0	2.1	3.8	5.4	7.3	4.2	10.6
D	20.0	14.6	20.3	13.3	10.0	16.7	14.0	36.0	25.5	0.0	0.0	0.0	10.6	13.2	18.2	16.4	14.6	25.5
F3	4.0	0.0	0.0	3.3	0.0	0.0	2.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
F	4.0	4.2	7.8	0.0	0.0	23.3	23.3	12.0	5.9	0.0	0.0	0.0	6.4	7.6	3.6	7.2	4.2	2.1
G2a	2.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
G	4.0	10.4	3.1	10.0	0.0	0.0	0.0	6.0	7.8	0.0	0.0	0.0	12.7	5.7	5.5	1.8	14.6	2.1
M7	14.0	0.0	4.7	0.0	2.0	3.3	18.6	4.0	7.8	0.0	0.0	0.0	6.4	1.9	0.0	5.4	0.0	0.0
M8	4.0	8.3	12.5	3.3	8.0	3.3	4.7	8.0	11.8	0.0	0.0	0.0	6.4	26.4	9.1	1.8	12.5	14.9
M9	4.0	6.3	3.1	0.0	2.0	3.3	0.0	4.0	3.9	0.0	0.0	0.0	0.0	1.9	1.8	1.8	4.2	0.0
M10	0.0	0.0	0.0	0.0	2.0	0.0	2.3	2.0	5.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
M11	2.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
N9a	2.0	0.0	6.3	0.0	2.0	6.7	7.0	6.0	2.0	0.0	0.0	0.0	0.0	1.9	0.0	1.8	6.3	0.0
M*	0.0	4.2	3.1	0.0	0.0	23.3	2.3	2.0	0.0	0.0	0.0	0.0	2.1	3.8	5.5	3.6	6.3	4.3
Y	0.0	0.0	0.0	0.0	2.0	0.0	0.0	2.0	2.0	0.0	0.0	0.0	0.0	0.0	1.8	0.0	0.0	0.0
R9b	0.0	0.0	0.0	6.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
R	0.0	0.0	1.6	3.3	0.0	0.0	0.0	2.0	2.0	0.0	0.0	0.0	0.0	0.0	0.0	12.7	4.2	2.1
HV	0.0	0.0	0.0	20.0	12.0	0.0	2.3	0.0	0.0	0.0	0.0	6.7	4.2	5.7	21.8	0.0	12.5	25.5
H	8.0	2.1	6.3	13.3	6.0	0.0	0.0	0.0	0.0	52.0	41.7	47.1	10.6	7.5	0.0	10.9	0.0	0.0
I	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.0	4.2	1.9	0.0	1.9	0.0	0.0	0.0	0.0
J	0.0	2.1	1.6	6.7	2.0	0.0	0.0	0.0	0.0	14.0	14.6	9.6	4.2	1.9	0.0	0.0	4.2	6.4
T	0.0	0.0	0.0	0.0	24.0	0.0	0.0	0.0	2.0	14.0	10.4	5.8	2.1	7.6	7.3	1.8	2.1	0.0
U	6.0	2.1	1.6	16.7	14.0	0.0	0.0	0.0	0.0	8.0	16.7	23.1	14.9	3.8	7.3	20.1	4.2	4.3
V	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.0	0.0	0.0	0.0	1.9	0.0	0.0	0.0	0.0
W	0.0	0.0	0.0	0.0	6.0	0.0	0.0	0.0	0.0	2.0	2.1	4.8	6.4	0.0	3.6	0.0	0.0	0.0
X	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4.0	8.3	1.0	0.0	0.0	0.0	0.0	0.0	0.0
N*	4.0	0.0	4.7	0.0	0.0	0.0	0.0	0.0	0.0	2.0	2.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Note: Haplogroup D includes subgroups D4 and D5; Haplogroup M8 includes subgroups M8a, C and Z; Haplogroup U includes subgroups U1—U8 and K. Numbers in parenthesis show the number of examined individuals. Asterisks indicate haplotypes that do not belong to any of the subgroups of the given haplogroup. Populations are coded as BA (Baoan), Tu (Tu), DX (Dongxiang), Uy (Uygur), Uzb (Uzbek), GH (Guangdong Han), YH (Yunnan Han), SH (Shandong Han), LH (Liaoning Han), Sic (Sicilian), Tus (Tuscan) and Slo (Slovenian). a) Yao et al.[7]; b) Francalacci et al.[9]; c) Torroni et al.[8]; d) Malyarchuk et al.[10]; e) Yao et al.[5]; f) Comas et al.[2].

in the NJ tree based on the pairwise F_{ST} values. The four Han populations form a cluster, and Baoan and Dongxiang are close to them. Tu is at a relatively long branch which connects with Dongxiang and then converges with Baoan. Three European populations (Sicilian, Slovenian and Tuscan) form a distinct cluster, which has the longest distance to EAs. The Uygur and Uzbek group with the CAs including two Uygur, two Kazak and two Kirkiz populations which are between the Hans-GCs group and the European populations. This result shows that the GC populations are genetically far for EUs but close to EAs and CAs. The genetic distances based on mtDNA haplogroups among the populations in the unrooted NJ tree correlate to some extent with their longitude locations over the Silk Road and are consistent with the result of PCA.

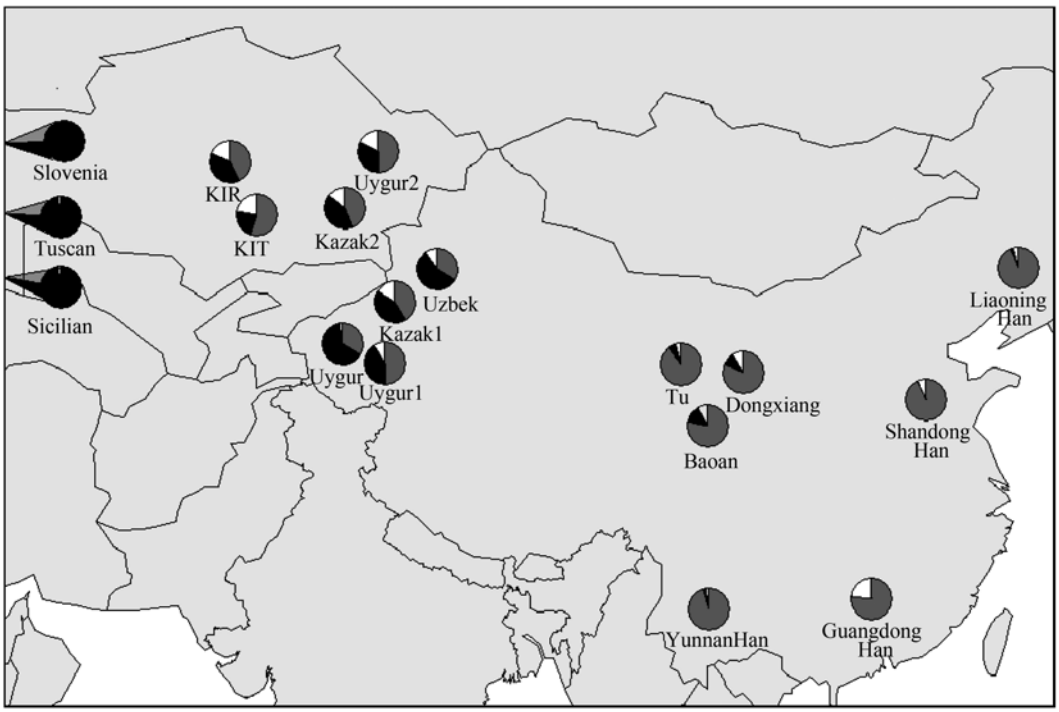

Figure 2 Map of the Eurasia, showing the samples analyzed in the present study (referenced populations included). The pie charts indicate the distribution of the main mtDNA haplogroups in the populations studied. The darkly shaded sections reflect the frequency of European haplogroup clusters (HV, H, V, J, T, U, I, W and X), the intermediately shaded sections show the Eastern Asian (M7, M8, M9, M10, A, B, D, G, F, F3 and N9a) and the white sections show the others (R, R9b, M* and N*).

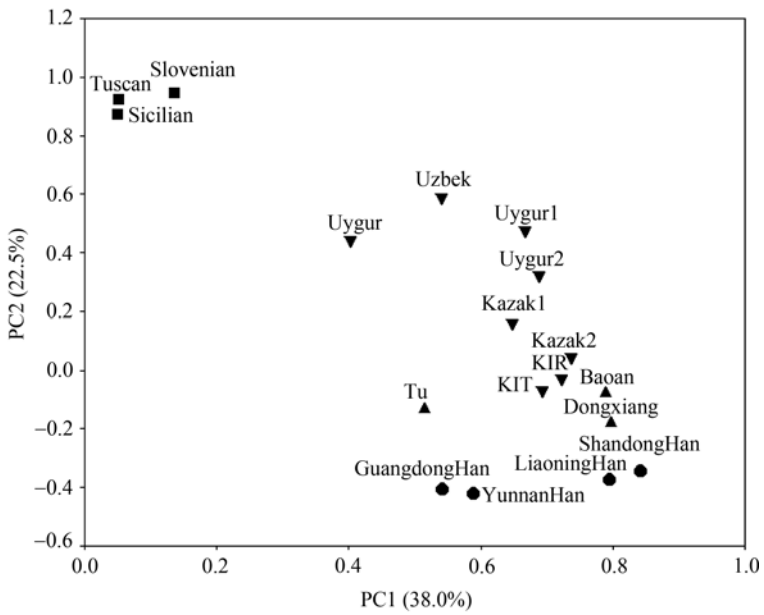

Figure 3 PC map based on mtDNA sequence of the five ethnic populations in this study and the thirteen Eurasian reference populations.

2.4 Estimation of gene admixture

The proportions of haplotypes contributed by EAs or CAs in the five ethnic populations estimated by the method of Roberts and Hiorns[16] are presented in Table 4. On the basis of mtDNA variations, the contributions of the Eastern Asian and European mtDNA pool (M_{RH}) to these populations fall in two groups. Uygur and Uzbek were 100 percents contributed by CAs as expected. The

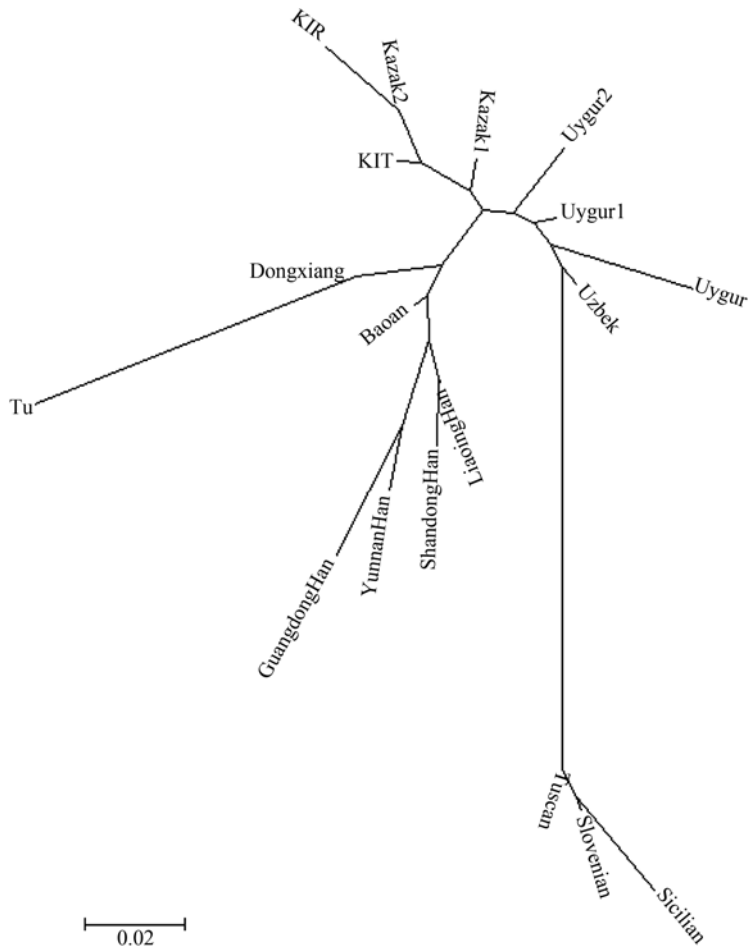

Figure 4 Neighbor-Joining tree based on mtDNA sequence of the five ethnic populations in this study and the thirteen Eurasian reference populations.

Table 4 Admixture of Eastern Asians and Central Asians in five ethnic populations

Populations	M_{RH} [a)]	
	Eastern Asians	Central Asians
Tu	0.6696	0.3304
Dongxiang	0.2868	0.7132
Baoan	0.6695	0.3305
Uygur	0.0000	1.0000
Uzbek	0.0000	1.0000

a) Eastern and Central Asians admixture proportion (M) is estimated by the method of Roberts and Hiorns (M_{RH}) [17].

GC population pool has contributions from both the CAs and EAs. Tu and Baoan have more maternal genetic contributions from EAs, while Dongxiang has more contributions from CAs. We displayed the frequency of mtDNA haplogroups identified as Eastern Asian or European in 26 European, Western Asian, Central Asian, Gansu Corridor and Eastern Asian populations in Table 5.

3 Discussions

Both Eastern Asian and European mtDNA lineages observed in the Uygur and Uzbek in Xinjiang populations with high frequencies (Table 3; Figure 2) supported that the extensive admixture existed in Central Asia as revealed by previous studies[2−5]. Both of the lineages were also observed in the three Gansu Corridor popula-

Table 5 Frequency of mtDNA haplogroups identified as Eastern Asian or European, in 26 European, Western Asian, Central Asian, Gansu Corridor and Eastern Asian populations

Region	Populations	Eastern Asian	European	Unassigned
Europe	Tuscan[a)]	0%	97.9%	2.1%
	Slovnian[b)]	0%	100%	0%
	Sicilian[c)]	0%	98%	2%
Western Asia	Turkish[d)]	8%	92%	0%
	Turkish(Azerbaijan)[d)]	2.5%	97.5%	0%
	Mazandarian[d)]	0%	95.2%	4.8%
	Gilaki[d)]	2.7%	94.6%	2.7%
	Iran[e)]	15%	80%	5%
Central Asia	Turkmen[d)]	36.5%	58.4%	5.1%
	Uzbek[d)]	30.9%	54.9%	14.2%
	Kazakh[f)]	56.4%	40.0%	3.6%
	Talas Kirghiz[f)]	72.9%	25.0%	2.1%
	Sary-Tash Kirghiz[f)]	59.6%	31.9%	8.5%
	Uighur[f)]	54.5%	34.5%	11%
	Kazak[g)]	66.2%	30.3%	3.5%
	Uygur[g)]	55.3%	42.4%	2.3%
	Uzbek[g)]	51.3%	37.6%	11.1%
	Uzbek	33.3%	56.7%	10%
	Uygur	36%	64%	0%
Gansu Corridor region in China	Baoan	82%	14%	4%
	Dongxiang	81.1%	9.5%	9.4%
	Tu	89.5%	6.3%	4.2%
Eastern Asia	Guangdong Han[h)]	76.7%	0%	23.3%
	Yunnan Han[h)]	95.4%	2.3%	2.3%
	Shandong Han[h)]	96%	0%	4%
	Liaoning Han[h)]	96%	2%	2%

Note: The assignment of mtDNA haplogroups was based on Macaulay et al.[12], Kivisild et al.[11] and Yao et al.[7]. a) Torroni et al.[8]; b) Malyarchuk et al.[10]; c) Francalacci et al.[9]; d) Quintana-Murci et al.[1]; e) Comas et al.[3]; f) Comas et al.[2]; g) Yao et al.[5]; h) Yao et al.[7].

tions Baoan, Dongxiang and Tu which demonstrated that the genetic admixture is also present in the GC region. Modest levels of European haplogroup frequencies were found in the Gansu Corridor populations, indicating that the spread of the European mtDNA lineages continued from Central Asia to the Gansu Corridor region along the Silk Road. A significant trend of the increase in the frequency of Eastern Asian haplogroupes and the decrease in the frequency of European haplogroups was shown through EUs, Western Asians, CAs, Gansu Corridor populations to EAs. The lower frequencies of European-specific haplogroups and higher frequencies of Eastern Asian-specific haplogroups found in the Gansu Corridor populations suggested less gene flow from European populations to the GC than to Central Asia.

The genetic position of the GC populations located between EAs and CAs with a closer genetic affinity to EAs than to CAs compared with Xinjiang populations, which is consistent with the observation of haplogroup frequency distributions on the whole that the Gansu Corridor populations harbored less European and more Eastern Asian matrilineal components (Table 5). It further confirmed the estimation of admixture of Eastern Asian and Central Asian lineages in the Gansu Corridor populations.

The historical studies showed that economic and cultural exchanges between the East and the West along the Silk Road were prevalent intermittently from the third century BC to the 17th century AD. Several historical events could possibly contribute to and shape the matrilineal genetic structure of populations residing in this

region[23]. Baoan and Dongxiang originated mainly from Mongol and Tu rooted in ancient Eastern Asia many centuries ago[24]. These populations settled down in the upper part of the Silk Road (GC) during the 7th-8th century. The Mongolian populations who were recorded as the major ancestors of Baoan and Dongxiang[24,25] conquered the northern part of Euraisa in the 13th century AD, when the local maternal contribution to Mongolian populations might have occurred. Members of other tribes from West such as Persian and Arab and from East joined Gansu Corridor populations during the formation of the nations[24]. Therefore, mtDNA of Europeans and Eastern Asians would be brought and imported into the gene pool of Gansu Corridor populations through marriage between East Eurasians and Europeans, migration and expansion of tribes along GC over many centuries.

Formerly mentioned studies[1−5] on the populations along the Silk Road living in Western Asia, Central Asia and Xinjiang Uygur Autonomous Region of China have provided extensive data on matrilineal genetic structures, which showed that Eastern Asian and European gene lineages flowed to Central Asia. Here, we presented the mtDNA polymorphisms data of the GC populations and found that compared with CAs, less European matrilineal components flooded the ethnic populations in the Gansu Corridor of China. It showed that the contribution of European mtDNA lineage was present but limited in the upper part of the Silk Road. The Gansu Corridor populations take a genetic position between Eastern Asians and Central Asians. It suggests that along the Silk Road connecting the West and East, in addition to significant gene admixture taken place in the low (west) and central part of the Road, modest gene admixture between Central Asians and Eastern Asians exists in the upper part of the Road.

In conclusion, we analyzed the mtDNA lineages of three ethnic populations (Baoan, Dongxiang and Tu) in the Gansu Corridor in the upper part of the Silk Road and detected matrilineal gene admixture between Eastern Asians and Central Asians in this region. These populations have a closer genetic relationship to Eastern Asians. This study revealed the gene flow from Central Asia to the upper part of the Silk Road, which makes the gene flow and gene admixture along the Silk Road more complete. Hence, the landscape of gene flow from Europe to the Central Asia then to the GC region becomes clearer.

We thank all of the donors for making this work possible and Dr. Bing Su for critical reading of the manuscript.

1 Quintana-Murci L, Chaix R, Wells R S, et al. Where west meets east: the complex mtDNA landscape of the southwest and Central Asian corridor. Am J Hum Genet, 2004, 74: 827−845 [DOI]

2 Comas D, Calafell F, Mateu E, et al. Trading genes along the Silk Road: mtDNA sequences and the origin of Central Asian populations. Am J Hum Genet, 1998, 63: 1824−1838 [DOI]

3 Comas D, Plaza S, Wells R S, et al. Admixture, migrations, and dispersals in Central Asia: Evidence from maternal DNA lineages. Eur J Hum Genet, 2004, 12: 495−504 [DOI]

4 Yao Y G, Lu X M, Luo H R, et al. Gene admixture in the silk road region of China: Evidence from mtDNA and melanocortin 1 receptor polymorphism. Genes Genet Syst, 2000, 75: 173−178 [DOI]

5 Yao Y G, Kong Q P, Wang C Y, et al. Different matrilineal contributions to genetic structure of ethnic groups in the silk road region in china. Mol Biol Evol, 2004, 21: 2265−2280 [DOI]

6 Song S H, Chen K J. Nations conspectus of China. Beijing: Central Nation University Press, 2001, 553−560

7 Yao Y G, Kong Q P, Bandelt H J, et al. Phylogeographic differentiation of mitochondrial DNA in Han Chinese. Am J Hum Genet, 2002, 70: 635−651 [DOI]

8 Torroni A, Huoponen K, Francalacci P, et al. Classification of European mtDNAs from an analysis of three European populations. Genetics, 1996, 144: 1835−1850

9 Francalacci P, Morelli L, Underhill P A, et al. Peopling of three Mediterranean islands (Corsica, Sardinia, and Sicily) Inferred by Y-Chromosome biallelic variability. Am J Phys Anthrop, 2003, 121: 270−279 [DOI]

10 Malyarchuk B A, Grzybowski T, Derenko M V, et al. Mitochondrial DNA variability in Bosnians and Slovenians. Ann Hum Genet, 2003, 67: 412−425 [DOI]

11 Kivisild T, Tolk H V, Parik J, et al. The emerging limbs and twigs of the East Asian mtDNA tree. Mol Biol Evol, 2002, 19: 1737−1751

12 Macaulay V, Richards M, Hickey E, et al. The emerging tree of west Eurasian mtDNA: A synthesis of control-region sequence and RFLPs. Am J Hum Genet, 1999, 64: 232−249 [DOI]

13 Andrews R M, Kubacka I, Chinnery P F, et al. Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. Nat Genet, 1999, 23: 147 [DOI]

14 Schneider S, Roessli D, Excoffier L. Arlequin version 2.000: A software for population genetics data analysis. Genetics and Biometry Laboratory, University of Geneva, Switzerland, 2000

15 Tamura K, Nei M. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol Biol Evol, 1993, 10: 512−526

16 Dupanloup I, Bertorelle G. Inferring admixture proportions from molecular data: Extension to any number of parental populations. Mol Biol Evol, 2001, 18: 672−675

17 Roberts D F, Hiorns R W. Methods of analysis of the genetic com-

position of a hybrid population. Hum Biol, 1965, 37: 38－43
18 Chakraborty R. Gene admixture in human populations: Models and predictions. Yearb Phys Anthropol, 1986, 29: 1－43 [DOI]
19 Sans M, Salzano F M, Chakraborty R. Historical genetics in Uruguay: Estimates of biological origins and their problems. Hum Biol, 1997, 69: 161－170
20 Finnila S, Lehtonen M S, Majamaa K. Phylogenetic network for European mtDNA. Am J Hum Genet, 2001, 68: 1475－1484 [DOI]
21 Herrnstadt C, Elson J L, Fahy E, et al. Reduced-median-network analysis of complete mitochondrial DNA coding-region sequences for the major African, Asian, and European haplogroups. Am J Hum Genet, 2002, 70: 1152－1171 [DOI]
22 Kong Q P, Yao Y G, Sun C, et al. Phylogeny of East Asian mitochondrial DNA lineages inferred from complete sequences. Am J Hum Genet, 2003, 73: 671－676 [DOI]
23 Cavalli-Sforza L L, Menozzi P, Piazza A. The history and geography of human genes. Princeton: Princeton University Press, 1994. 195－252
24 Bao E H, Jiang Y, Ya H Z, et al. Cyclopaedia of China-nations. Beijing: Cyclopaedia of China Press, 1986, 42. 97
25 Zhang F, Cao Y N, Ye X M, et al. General Hhistory of Mongolia Nation. Vol 1. Hohhot: Inner-mongolia University Press, 2002. 224－251