8-1-2012

# Genetic Characteristics Of An Ancient Nomadic Group In Northern China

Haijing Wang
*Jilin University, People's Republic of China*

Lu Chen
*Jilin University, People's Republic of China*

Binwen Ge
*Jilin University, People's Republic of China*

Ye Zhang
*Jilin University, People's Republic of China*

Hong Zhu
*Jilin University, People's Republic of China*

***See next page for additional authors***

# Genetic characteristics of an ancient nomadic group in northern China

Haijing Wang[1, 2, 3] Lu Chen[4, &] Binwen Ge[4] Ye Zhang[4] Hong Zhu[2] Hui Zhou[1, 2, 4,*]

***1 Key Laboratory for Evolution of Past Life and Environment in Northeast Asia, Jilin University, Ministry of Education, Changchun 130023, People's Republic of China***

***2 Ancient DNA Laboratory, Research Center for Chinese Frontier Archaeology, Jilin University, Changchun 130012, People's Republic of China***

***3 College of Chemistry, Jilin University, Changchun 130012, People's Republic of China***

***4 College of Life Sciences, Jilin University, Changchun 130012, People's Republic of China***

**[&]Co-first authors**: Haijing Wang and Lu Chen contributed equally to this work.

***Correspondence to**: Hui Zhou, Ancient DNA Laboratory, Research Center for Chinese Frontier Archaeology, Jilin University, Changchun 130012, People's Republic of China. E-mail: zhouhui@mail.jlu.edu.cn

**Key words**: ancient DNA; mitochondrial DNA; Y-chromosome single nucleotide polymorphisms; nomads; Donghu

**Running title**: Genetic characteristics of an ancient nomadic group

## Abstract

Nomadic populations have played a significant role in the history of not only China but also in many nations worldwide. Because they had no written language, an important aspect in the study of these people is the discovery of their tombs. It has been generally accepted that Xiongnu was the first empire created by nomadic tribe in the 3rd century B.C. However, little population genetic information is available concerning the Donghu, another flourishing nomadic tribe at the same period because of the restriction of materials until Jinggouzi site was excavated. In order to test the genetic characteristics of ancient people in this site and explore the relationship between Jinggouzis and Donghus, two uniparentally inherited markers were analyzed from 42 human remains in this site, which located in northern China, dated approximately 2,500 years ago. With ancient DNA technology, four mtDNA haplogroups (D, G, C and M10) and one Y chromosome haplogroup (C) were identified using mitochondrial DNA and Y-chromosome single nucleotide polymorphisms (Y-SNPs). Those haplogroups are common in North Asia and East Asia. And the Jinggouzi people were genetically closest to the Xianbeis in ancient populations and to the Oroqens among extant populations, who were all pastoralists. This might indicate that ancient Jinggouzi people were nomads. Meanwhile, according to the genetic data and the evidences in archaeology, we inferred that Jinggouzi people were associated with Donghu. It is of much value to trace the history of Donghu tribe and might show some insight into the ancient nomadic society.

## Introduction

Many significant nomadic groups have roamed the eastern regions of the Eurasian steppes in both prehistorical and historical times. As ancient nomads migrated frequently due to the natural environment and geographical conditions, their burial sites are found more rarely than those of farming populations. It is acceptable by the worldwide that the Xiongnu was the first empire created by nomadic tribe in the 3rd century B.C. in history (Christine et al. 2003). Nonetheless few knew that there was another ancient nomadic tribal union in the same period—the Donghu, located in the eastern of Xiongnus in the record of some ancient Chinese literature (Lin 1989), had been flourishing in North China for many years. After being defeated by Xiongnus in early western Han dynasty (about 206 BC), the underlings broke up to two new nomadic tribes-- the Wuhuan and Xianbei (Lin 1989). And some later important tribes or ethnic groups such as the Khitan, Shiwei, Mongol, Daur and Xibo, among others, also belonged to the Donghu lineage (Lin 1989). Because the Donghu tribe had neither writings nor buildings, and the record on it was scarcity, so that it couldn't be known to the world. Some scholars try to find the remains of the Donghu to testify its existence and study the ancient tribe, but they didn't obtain valuable clues for many years until the Jinggouzi site was excavated.

The Jinggouzi site consisted of diverse burial patterns (21 single burials, 12 double and 22 multiple burials). Many animal bones, such as those of horse, cattle, sheep, donkey, mule and dog, were found at this archaeological site, while no farming tools and farming products were discovered. This observation implied that stockbreeding was a dominant activity for their livelihood (Wang et al. 2010). The Jinggouzi site was estimated to have

been used in the late Spring-Autumn period (770–476 BC) and the early War States period (475–221 BC) based on the associated funeral material and was corroborated by reasonable radiocarbon ($^{14}C$) measurements (2485 ± 45 B.P.) (Wang et al. 2010).

Genetic tools can be applied to assess origins and their genetic relationships with other present populations (Zhao et al. 2011, Ewen 2011). In this study, we analyzed mtDNA HVSI sequence variability, and genotyped 5 biallelic markers on the Y chromosome in 42 Jinggouzis. Meanwhile, we compared the Jinggouzs with ancient and contemporary populations. Through these genetic analyses, we studied the genetic characteristics of the ancient remains and explored the relationships between the Jinggouzi people and Donghus.

## Materials and methods

### Sampling

The Jinggouzi burial site lies in the sloping loess hillock of a wadi in the northern bank of the upper Xilamulun River, Chifeng city, Inner Mongolia, China (located at position N43°23.169', E118°14.288') (Figure 1). The altitude of this site is approximately 976 m, and the average annual temperature is 4.2°C (Wang et al. 2010). The Jinggouzi site is divided into two areas: the Western area, from which our samples were obtained, and the Eastern area. Teeth and bones of 42 individuals were collected, including 6 tooth samples (sample codes 9, 11, 13, 18, 19, 20), 21 bone samples (sample code 1, 3, 5, 6, 7, 8, 10, 12, 14, 15, 17, 22, 23, 29, 31, 33, 34, 35, 36, 38, 41) and 15 paired tooth and bone samples (sample codes 2, 4, 16, 21, 24, 25, 26, 27, 28, 30, 32, 37, 39, 40, 42). All the skeletal remains were chosen based on their macroscopic preservation of having good

outward appearance and no fractures.

**mtDNA extraction, sequencing and haplogroup assignment**

The samples were excavated by professional researchers. And strict precautions were taken to avoid the contamination of samples with modern DNA during the process of excavating. And we have excluded the sequences that were the same with those of researchers and people who contacted these samples when we analyzed the data.

DNA was extracted from 57 samples corresponding to the remains of 42 individuals (15 individuals were typed from both tooth and bone samples). In order to eliminate surface contamination, nearly 2–3 mm of the outer surface of the bones was removed with a sanding machine. A 3 cm fragment was cut from the bone using a sterilized hacksaw. The sample (bone fragment or tooth) was immersed in 5% sodium hypochlorite solution for 20 min, rinsed with ultrapure water and 100% ethanol and then UV-irradiated for 30 min on each side. Subsequently, each sample was ground into a powder in a 6850 Freezer Mill (Spex SamplePrep, Metuchen, NJ, USA). 500mg powder samples were incubated at 50℃, 220 r.p.m.min$^{-1}$ for 24 hours in 3mL digestion mix consisting 0.45 moL·L$^{-1}$ EDTA , 0.5% SDS, 0.7 g·L$^{-1}$ protease K. DNA was carefully extracted from the powdered samples using the QIAquick PCR Purification Kit (Qiagen, Germany) according to the manufacturer's instructions. Amplification of the mtDNA was performed on two sub-regions (A 16017-16251, B 16201-16409) of the first hypervariable segment I (HVSI) of the control region using the primers listed in Table 1. PCR amplification was carried out in 25 μL of a reaction mixture containing 2 mL extract, 1.5 reaction buffer (Fermentas, Burlington, Canada), 1U of Taq polymerase (Fermentas), 2.5mM $MgCl_2$

(Fermentas), 0.2mM dNTP Mix (Promega, Madison, WI, USA), 0.8mg.mL$^{-1}$ BSA (Takara, Da Lian, China) and 0.2 mM of each primer (Sangon, Shanghai, China). Cycling parameters were 94ºC for 4 min, followed by 33 cycles with 94ºC for 30s, 45 s at 55ºC, elongation for 30s at 72ºC, extension 10 min at 72ºC and storage at 4ºC. Negative controls containing all of the reagents but without the ancient DNA template were included in each DNA extraction/ PCR run. Amplification products were checked on a 2% agarose gel and purified with the QIAEX II Gel Extraction Kit (Qiagen Germany). PCR products were sequenced using the ABI 310 Terminator Sequencing kit (PE Applied Biosystems) and were analyzed on the ABI PRISM 310 automatic sequencer (PE Applied Biosystems).

Moreover, a set of coding-region mtDNA polymorphisms were typed to determine the major classifications of the mtDNA tree. Haplogroups G (4833A) and M10 (10646A) were identified by direct sequencing, and haplogroups M (10400T), C (14318C), D (5178A) and D4 (3010A) were examined by using the amplified product-length polymorphism (APLP) method (Shinoda et al. 2006), and the primers listed in Table 1. The constitution of the PCR reaction mixture was the same as described above. The direct sequencing conditions were the same as those for the mitochondrial HVSI amplification. The APLP PCR conditions were as follows: incubation at 95 ºC for 4 min; 40 cycles at 94 ºC for 10 sec, 52 ºC–54 ºC for 10 sec, and 72 ºC for 5 sec; and final extension at 72 ºC for 1 min. A 2 μL aliquot of the PCR product was separated by electrophoresis in an 8-cm native polyacrylamide gel (10%T, 5% C) containing 1x TBE buffer (pH 8.0) with running buffer (1x TBE, pH 8.0). DNA bands were detected by ultraviolet irradiation after

staining with ethidium bromide.

**Sex identification and Y-chromosome single nucleotide polymorphism (SNP) analysis**

Sex determination was performed in all samples using the amelogenin fragment. The male samples were then chosen for further analysis. We screened all male samples with five bi-allelic markers (M216-C, M8-C1, M38-C2, PK2-C3 and M356-C5) which define the major branches on the Eurasian haplogroup tree (Karafet 2008, Hammer 2002). All primers are listed in Table 1 and the PCR reaction conditions were the same as those for the mitochondrial HVSI amplifications but the length of the PCR products was all around 100–150 bp.

**Cloning of PCR products**

To minimize the risk of contamination, 4 of 42 samples were cloned at random. The mtDNA HVSI products were cloned using the pGEM-T Easy Vector System I (Promega) following the instructions of the supplier. Six to ten clones from two independent amplifications were selected for automated DNA sequencing, using vector M13 primers.

**Measures taken to ensure authenticity**

Ancient DNA can be easily contaminated by exogenous DNA. In order to prevent this situation, generally accepted guidelines for DNA studies were followed (Pääbo et al. 2004; Cooper et al. 2000). Three separate rooms were used for specimen handling, pre- and post-PCR analyses, and multiple reagent blanks and negative control reactions were performed with all amplifications. All tools applied for the research were properly decontaminated (UV exposure/autoclaving), and every researcher was required to wear

full body protective clothing, facemasks and gloves. Four samples were sent to the State Key Laboratory for Agrobiotechnology and College of Biological Science, China Agricultural University in Beijing, China, for independent confirmation (Table 2). For mtDNA analysis, four samples were cloned using the Promaga Cloning kit (Table 2). In order to prevent potential Y-chromosome DNA contamination, only female researchers conducted all procedures for the Y-SNP analysis. In addition, the mitochondrial hypervariable I sequences of all researchers were obtained and compared with those of the samples (Table 2).

**Data analysis**

Published data on mtDNA haplogroups and Y-SNP haplogroups of ancient and extant populations were retrieved, and the source of which is provided in Fig.1. DNA sequences were analyzed using Cluster X 1.83 (http: //www.clustal. org/download/1.X/ftp-igbmc.u-strasbg.fr/pub/ClustalX/). MtDNA haplogroups were used to multidimensional analysis (MDS). Pairwise $F_{ST}$ values among populations were worked out in ARLEQUIN 3.1. MDS based on $F_{ST}$ values were performed using SPSS 13.0. The MtDNA haplogroups and Y-SNP haplogroups were also used for frequency analysis.

## Results

### Authenticity of results

Extremely thorough laboratory precautions and systematic controls to prevent contamination from modern DNA were employed (see Materials and Methods). The sequences of tooth and bone from the same individual were compared and proved to be the same. Four samples re-tested at the State Key Laboratories AgriBiotechnology in China Agricultural University yielded consistent results (Table 2). Four samples were chosen randomly to be cloned, and the results were the same as that of direct sequencing (Table 2). All researchers and persons who came into contact with these samples submitted DNA for sequencing, and the results showed that none of them matched the samples (Table 2). In this study, all blank and negative results were negative. Collectively, 42 mtDNA fragments were obtained and considered to be authentic.

### MtDNA polymorphisms and haplogroup identification

42 mtDNA sequences were all successfully amplified from position 16017 to 16409 of the revised Cambridge Reference Sequence (rCRS). There were a total of 32 phylogenetically informative sites, and 26 mtDNA haplotypes were obtained, as presented in Table 2. A majority of the mtDNA of these specimens fully exhibited haplogroup motifs and therefore could be safely assigned to the relevant haplogroups, which were consistent with the APLP and sequencing results. The 26 different haplotypes were further assigned to 4 haplogroups: C, D, G and M10. Out of 42 obtained sequences, 6 were assigned to haplogroup C, 25 to haplogroup D, 10 to G and 1 to M10 (Table 2). The high frequency of the 16223-16362 motif was the main characteristic of the mtDNA sequences, and haplogroups D and G were dominant in the ancient Jinggouzi people.

Moreover, almost all of the samples belonging to the mtDNA haplogroup D were in the sub-haplogroup D4, except for one (sample code 29) which was assigned to haplogroup D5 based on the mutation 16189 (Yao et al. 2002a) and another one (sample code 14) which could not be further classified. All of the haplogroup G samples were further divided into haplogroup G2 based on the mutation 16278 (Yao et al. 2002a). All of these results are shown in Table 2.

**Sex identification and Y chromosome SNP analysis**

Sex identification was successfully achieved for 36 ancient Jinggouzi remains, and 6 samples failed to be amplified. 16 individuals were identified to be female, and 20 were determined as males and 12 were subsequently subjected to be assigned to haplogroup C of Y-SNP, which is defined by the marker M216 (C→T). In this study, a set of Y-SNP markers was investigated, including M8-derived (Hg C1), M38-derived (Hg C2), PK2-derived (Hg C3) and M356-derived (Hg C5) markers, none of which matched our samples. Considering that haplogroup C4 defined by the marker M347 and haplogroup C6 defined by the marker P55 are completely restricted to Australia and New Guinea, respectively (Hudjashov et al. 2007, Tatiana et al. 2008), far away from our sample collection site, these samples were assigned to the Hg C* group in this study.

**Phylogenetic analyses**

Some individuals shared the same haplotypes: (i) 13, 17, 24, and 26 (ii) 3, 5 and 6 (iii) 20, 28 and 42 (iv) 15, 22 and 31 (v) 23 and 35 (vi) 34 and 36 (vii) 11 and 25 (viii) 38 and 41 (ix) 33 and 39 (x) 30 and 40 (xi) 7 and 19. The individuals sharing the same sequence were concentrated in same or nearby burial. There may be the existence of

kinship on the maternal among them based on this fact. A correction should be carried out in order to reducing the bias of frequencies due to the kinship (Vernesi et al., 2004, Ainhoa et al., 2006, Zhao et al., 2010). Referring to the correction criterion applied by Ainhoa and Zhao, 26 sequences were selected from 42 remains for statistic analysis.

Pairwise $F_{ST}$ values are summarized in a multidimensionally scaled (MDS) graph (Fig. 2), in which the plotted points correspond to Asian populations and the genetic distances between them are proportional to the linear distances between the plotted points. In the MDS graph, Jinggouzi people showed the nearest distance with Xianbeis among ancient populations and with Oroqens among extant populations (Fig.2).

## Discussion

Jinggouzi cemeteries were composed of relatively well-preserved skeletons. Indeed, the climatic conditions (cold and dry) and the archaeological context encountered at this site had undoubtedly protected the most recovered specimens against DNA degradation. In the present article, strict precautions were taken to avoid the contamination of samples with modern DNA during sampling and laboratory analysis. With appropriate effort to exclude possible errors in the obtained sequences, genetic data were successfully obtained from a sample of 42 human skeletal remains in this site using maternal and paternal markers.

From the maternal standpoint, four mtDNA haplogroups (D, C, G and M10) were observed, which all belonging to the macrohaplogroup M and could be assigned to an Eastern Eurasian mtDNA gene pool. Some individuals sharing the same sequence (shown in Table 2) had close maternal genetic relationships (Vernesi et al. 2004, Ainhoa et al.

2006, Zhao et al. 2011). Regarding the paternal inheritance, only one Y-SNP haplogroup (Hg C*) was detected. It indicated that ancient Jinggouzi people had simplex paternal genetic structure.

According to MDS, the results indicated some degree of genetic similarity between Jinggouzis and nomadic populations (Xiangbeis, Xiongnus, Oroqens and Mongols) rather than farming groups (Hengbeis, Taojiazhais, Nuheliangs, South Hans and North Hans) (Fig.2). Meanwhile, according to the excavation report, animal skulls, bronze swords, as well as daggers but no farming tools were found at the Jinggouzi burial site. These burial objects implied that the ancient Jinggouzi people lived by hunting and stockbreeding (Wang et al. 2010). Moreover, a high ratio of $^{15}N$ isotopes in these bones of remains from the Western zone of the Jinggouzi cemetery was found, indicating that the animal food intake in their daily nutrition was rather high(Zhang et al. 2008). This finding also suggests that activities related to animal raising and hunting played a significant role in the subsistence of the Jinggouzi people at the time. Combined with the genetic data, we inferred that ancient Jinggouzi people were nomads.

Mitochondrial DNA and Y-chromosome DNA diversity were comparable to that of ancient and contemporary populations. Jinggouzi people exhibited the nearest genetic distance with Xianbeis among ancient groups (Figure 2), indicating that Jinggouzi people had the closest relationship with Xianbeis, who were the descendants of Donghus. Interestingly, both the mtDNA and Y-chromosome DNA frequency analyses indicate that Jinggouzis are most closely related to Oroqens (Fig. 3 and 4), who are originating from the north of Asia (Pu 1983). This result was consistent with that of MDS analysis. And

most scholars consider the Shiwei was the ancestor of Oroqens, who were also the Donghus' descendants. It can be inferred that Jinggouzis had some relationship with Donghus.

In addition, the location of the burial ground, Chifeng city of Inner Mongolia, was part of the Donghu territory, and the $^{14}C$ analysis dated the site to 2485 ± 45 B.C., coinciding with the time in which the Donghu flourished according to the Chinese historical records (Wang et al. 2010). Based on these findings, we inferred that the ancient Jinggouzi people associated with the Donghu culture, they might belong to Donghu population. It is of much value to trace the history of Donghu and know more about the development and disappearance of this ancient nomadic tribe. Moreover, it might help us to explore the structure of the ancient nomadic society.

*Acknowledgements*

We thank Professor Xingbo Zhao at the State Key Laboratories AgriBiotechnology in China Agricultural University for performing the mtDNA sequencing for confirmation. This study was supported by grants from the National Natural Science of China (Grant No. **J0930002**), National Social Science of China (Grant No. **11&ZD182**) and Special funds of the Central Colleges Basic Scientific Research Operating Expenses, Jilin University (Grant No. **450060445661**).

**Literature Cited**

Alzualde A, Izagirre N, Alonso S et al. 2006. Insights into the “isolation” of the Basques: mtDNA lineages from the historical site of Aldaieta (6th-7th centuries AD). *Am J Phys Anthropol*. 130:394-404.

Andrews RM, Kubacka I, Chinnery PF et al (1999) Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. Nat Genet 23:147.

Christine Keyser-Tracqui, Eric Crubézy, Bertrand Ludes. 2003. Nuclear and Mitochondrial DNA analysis of a 2000-year-old necropolis in the Egyin Gol valley of Mongolia. *Am J Hum Genet*. 73: 247-260.

Cooper A., Poinar H. N. 2000. Ancient DNA: do it right or not at all. *Science*. 289: 1139.

Derenko M., Malyarchuk B., Grzybowski T. et al. 2007. Phylogeographic analysis of mitochondrial DNA in northern Asian populations. *Am. J. Hum. Genet*. 81:1025–1041.

Derenko M., Malyarchuk B., Wozniak M. et al. 2006. The Diversity of Y-Chromosome Lineages in Indigenous Population of South Siberia. *Doklady Biological Sciences.*

Ewen Callaway. 2011. Ancient DNA reveals secrets of human history. *Nature*. 476:136-137.

Hammer M. F. 2002. The Y Chromosome Consortium. A nomenclature system for the tree of human Y-chromosomal binary haplogroups. *Genome. Res*. 12:339–348.

Hua Zhong, Hong Shi, Xue-Bin Qi et al. 2010. Global distribution of Y-chromosome haplogroup C reveals the prehistoric migration routes of African exodus and early

settlement in East Asia. *Journal of Human Genetics*.55:428–435.

Karafet T. M., Mendez F. L., Meilerman M. B., et al. 2008. New binary polymorphisms reshape and increase resolution of the human Y chromosomal haplogroup tree. *Genome. Res*. 18: 830–838.

Kolman CJ, Sambuughin N, Bermingham E. 1996. Mitochondrial DNA analysis of Mongolian populations and implications for the origin of New World founders. *Genetics.* 142:1321-1334.

Kong QP, Yao YG, Liu M et al. 2003. Mitochondrial DNA sequence polymorphisms of five ethnic populations from northern China. *Hum Genetic*. 113:391-405.

Lin gan. 1989. *A History of Donghu*. Hohhot China: Inner Mongolia People's Publishing House.

Pääbo S., Poinar H., David S. et al. 2004. Genetic analyses from ancient DNA. *Annu. Rev. Genet.* 38:645–679.

Pu Qiu. 1983. *The Oroqens: China's nomadic hunters*. China Publications Centre (Guoji Shudian). Beijing China: Foreign Languages Press.

Shinoda K, Adachi N, Guillen S et al. 2006. Mitochondrial DNA analysis of ancient Peruvian highlanders. *Am J Phys Anthropol*, 131(1):98-107.

Tanaka M, Cabrera V M, González A M et al. 2004. Mitochondrial genome variation in eastern asia and the peopling of Japan. *Genome Res*. 14: 1832-1850.

Vernesi C, Caramelli D, Dupanloup I et al. 2004. The Etruscans: a population-genetic study. *Am J Hum Genet*. 74:694-704.

Wang Lixin, Ta La, Zhu Yonggang . 2010. *Lin Xi Jinggouzi: Archaeological Report*

*and Comprehensive Research on the late Bronze Age Cemetery.* Inner Mongolia Archaeological Institute, Jilin University Research Center for Chinese Frontier Archaeology eds. Beijing China: Science Press.

Xu Z, Zhang F, Xu B et al. 2008. Mitochondrial DNA evidence for a diversified origin of workers building First Emperor of China. *PLoS One.* 3(10):e3275.

Yao YG, Kong QP, Bandelt H-J et al. 2002a. Phylogeographic differentiation of mitochondrial DNA in Han Chinese. *Am J Hum Genet*. 70:635-651.

Yu CC, Xie L, Zhang XL et al. 2006. Genetic analysis on Tuoba Xianbei remains excavated from Qilang Mountain Cemetery in Qahar Right Wing Middle Banner of Inner Mongolia. *FEBS Lett*. 580(26):6242-6246.

Zhang Quanchao, Jacqueline T. ENG, Wang Lixin, et al. 2008. Stable isotope analysis of the human remains of Jinggouzi West Zone Cemetery in Linxi Country, Inner Mongolian. *Frontier archaeological research.* 7:322-327

Zhao Yong-Bin, Li Hong-Jie, Li Sheng-Nan et al. 2011. Ancient DNA Evidence Supports the Contribution of Di-Qiang People to the Han Chinese Gene Pool. *American Journal of Physical Anthropology*. 144(2):258–268.

**Table 1**. All primers used in this study

| HVSI fragments | Primers | | Length |
|---|---|---|---|
| HVSI-AB | L16017 5' TTCTCTGTTCTTTCATGGGGA 3'<br>H16251 5' GGAGTTGCAGTTGATGTGTGA 3' | | 235 bp |
| HVSI-CD | L16201 5' CAAGCAAGTACAGCAATCAAC 3'<br>H16409 5' AGGATGGTGGTCAAGGGA 3' | | 209 bp |
| Mitochondrial Haplogroups | Primers | Sites | Length |
| M | 10400T 5' TAATTATACAAAAAGGATTAGACTGTGCT 3'<br>10400C 5' TACAAAAAGGATTAGACAGAACC 3'<br>10400R 5' GAAGTGAGATGGTAAATGCTAG 3' | 10400 T/C | 149 bp (M)/142 bp |
| M10 | L10600 5'CTACTCTCATAACCCTCAAC3'<br>H10762 5'CATTGGAGTAGGTTTAGG3' | 10646 A/G | 163bp |
| D | 5178A 5' TGATCAACGCACCTGAAACAAGA 3'<br>5178C 5' GTCGCACCTGAAGCAAGC 3'<br>5178R 5' CCCATTTGAGCAAAAAGCC 3' | 5178A /C | 107 bp (D)/102 bp |
| D4 | 3010G 5'aTTGGATCAGGACtTCCCG3'<br>3010A 5'gctacaTGGATCAGGACAaCCCA3'<br>3010R 5'GATCACGTAGGACTTTAATCG3' | 3010G /A | 79bp(D4)/75 bp |
| G | L4735 5'TTACCTCCTCAAAGCAATACA3'<br>H4899 5'GGTATATGATTGAGATGGG3' | 4833G /A | 165 bp |
| C | L14318T 5'CCTTCATAAATTATTCAGCTTCCaACACTAT3'<br>L14318C5'aaaaagctaCATAAATTATTCAGCTTCCTACtCTAC3'<br>L14318R5'TTAGTGGGGTTAGCGATGGA3' | 14318 C/T | 110bp/115 bp |
| Y chromosome haplogroups | Primers | Sites | Length |
| C | M216S:5'TCACTTTTATATCCTCAACCA3'<br>M216-A:5'AATCTGAATTCTGACACTGC3' | M216 | 109 bp |
| C1 | M8S 5' GCAAGTTTAGTGCCTCAGTATC 3'<br>M8A 5' TAAAGACCCCAGGCAAGAC 3' | M8 | 115 bp |
| C2 | M38S 5'TGGCAATGGTATGTAGGC3'<br>M38A 5'GCTGGCACATCTGTCATAA3' | M38 | 145 bp |
| C3 | PK2S 5' CTGAGGCTGTGGCAGAGAG 3'<br>PK2A 5' AAGCCATTGGAGAATCAGGT 3' | PK2 | 120 bp |
| C5 | M356S 5' GCCCTGCTCAGTTGTTGT 3'<br>M356A 5' AAAATGATGGCCTGTGGTA 3' | M356 | 120 bp |
| Sex identification | Primers | | Length |
| AMG gene | XBamel-1 5'CCTGGGCTCTGTAAAGAATAG3'<br>XBamel-2 5'CAGAGCTTAAACTGGGAAGCTG3' | | 103 bp (F)<br>103 and 109 bp (M) |

**Table 2**. Sequence variation in the Jinggouzi population

| Type | Sample Code | HVSI (16038 -16391) 16000+ | SNPs | Haplogroup |
|---|---|---|---|---|
| T1 | 13,17[a],24[b],26 | 092-129-223-271-362 | 10400T 5178A 3010A | D4 |
| T2 | 3,5,6 | 158-224-245-292-362 | 10400T 5178A 3010A | D4 |
| T3 | 21 | 224-292-362 | 10400T 5178A 3010A | D4 |
| T4 | 18 | 223-224-292-362 | 10400T 5178A 3010A | D4 |
| T5 | 2 | 223-246-287-311-362 | 10400T 5178A 3010A | D4 |
| T6 | 8 | 192-223-311-362 | 10400T 5178A 3010A | D4 |
| T7 | 37[b] | 214-223-362 | 10400T 5178A 3010A | D4 |
| T8 | 23, 35 | 223-362 | 10400T 5178A 3010A | D4 |
| T9 | 34,36 | 223-362-368 | 10400T 5178A 3010A | D4 |
| T10 | 10 | 223-319-362 | 10400T 5178A 3010A | D4 |
| T11 | 11,25 | 173-223-274-291-362 | 10400T 5178A 3010A | D4 |
| T12 | 38[a],41 | 223-274-362 | 10400T 5178A 3010A | D4 |
| T13 | 33[b],39 | 223-245-362 | 10400T 5178A 3010A | D4 |
| T14 | 29 | 189-265-298-319-362 | 10400T 5178A | D5 |
| T15 | 14[a] | 209-223-274-362 | 10400T 5178A | D |
| T16 | 20,28,42 | 223-227-278-293-362 | 10400T 4833G | G2 |
| T17 | 30,40 | 093-223-227-234-278-309-362 | 10400T 4833G | G2 |
| T18 | 15,22,31 | 093-223-227-278-362 | 10400T 4833G | G2 |
| T19 | 4 | 093-223-278-362 | 10400T 4833G | G2 |
| T20 | 9 | 223-278-293-362 | 10400T 4833G | G2 |
| T21 | 27 | 093-223-298-327 | 10400T 14318C | C |
| T22 | 7,19[b] | 093-129-223-298-327 | 10400T 14318C | C |
| T23 | 12[a] | 129-150-223-298-327 | 10400T 14318C | C |
| T24 | 1 | 129-150-171-223-298-327 | 10400T 14318C | C |
| T25 | 16 | 129-223-298-327 | 10400T 14318C | C |
| T26 | 32 | 223-311 | 10400T 10646A | M10 |
| | 1[c] | 124-183-189-278-293-362 | | |
| | 2[c] | 183-189-261-311-365 | | |
| | 3[c] | 183-189-223-234-290-362 | | |
| | 4[c] | 192-223-290-319-362 | | |
| | 5[c] | 189-261-278-311-362 | | |
| | 6[c] | 223-290-319 | | |

[a] Samples were replicated independently in China Agricultural University.

[b] Cloned mtDNA HVSI segments of the samples.

[c] The mtDNA HVSI segments of researchers in this study.

## Figure legends

**Figure 1**. Geographic location of the Jinggouzi burial site and populations in this study. 1 Jinggouzis(This study), 2 Xianbeis (Yu et al. 2006), 3 Xiongnus (Christine et al. 2003), 4 Taojiazhais (Zhao et al. 2011), 5 Hengbeis (unpublished data), 6 Niuheliangs (unpublished data), 7 mausoleum-building workers (MBWs) (Xu et al. 2008). 8 Tuvinians (Derenko et al. 2007, Derenko et al. 2006), 9 Teleuts (Derenko et al. 2007, Derenko et al. 2006), 10 Altaians-Kizhis (Derenko et al. 2007, Derenko et al. 2006), 11 Evenks (Derenko et al. 2007, Derenko et al. 2006), 12 Mongolians (Kolman et al. 1996, Hua et al. 2010), 13 Oroqens (Kong et al. 2003, Hua et al. 2010), 14 Inner Mongolians (Kong et al. 2003, Hua et al. 2010), 15 Ewenkis (Kong et al. 2003, Hua et al. 2010), 16 Daurs (Kong et al. 2003, Hua et al. 2010), 17 Koreans (Tanaka et al. 2004), 18 North Hans (Yao et al. 2002a, Hua et al. 2010), 19 South Hans (Yao et al. 2002a, Hua et al. 2010).
1-7 were ancient populations and 8-19 were extant populations.

**Figure 2.** MDS plot based on $F_{ST}$ values of mtDNA HVSI.
The black squares represent ancient populations, and the hollow squares represent extant populations.

**Figure 3**. The frequency of mtDNA haplogroups in different populations.

**Figure 4**. The frequency of haplogroup C of Y-SNP in different populations.


Russia
10
11
9
8
3
15
16
13
14
Kazakhstan
Mongolia
12
1
2
6
18a
Uzbekistan
Kyrgyzstan
China
Turkmenistan
Japan
18c
17
4
18e
5
18b
South Korea
7
18d
Afghanistan
Iran
Pakistan
19e
East China Sea
Nepal
19d
19a
19c
India
19b
Myanmar (Burma)
Oman
Philippine Sea
Thailand
South China Sea
Bay of Bengal
Andaman Sea
Arabian Sea
Vietnam
Philippines
Gulf of Thailand
Laccadive Sea
Malaysia
Sea of Okhotsk

**Figure 1**

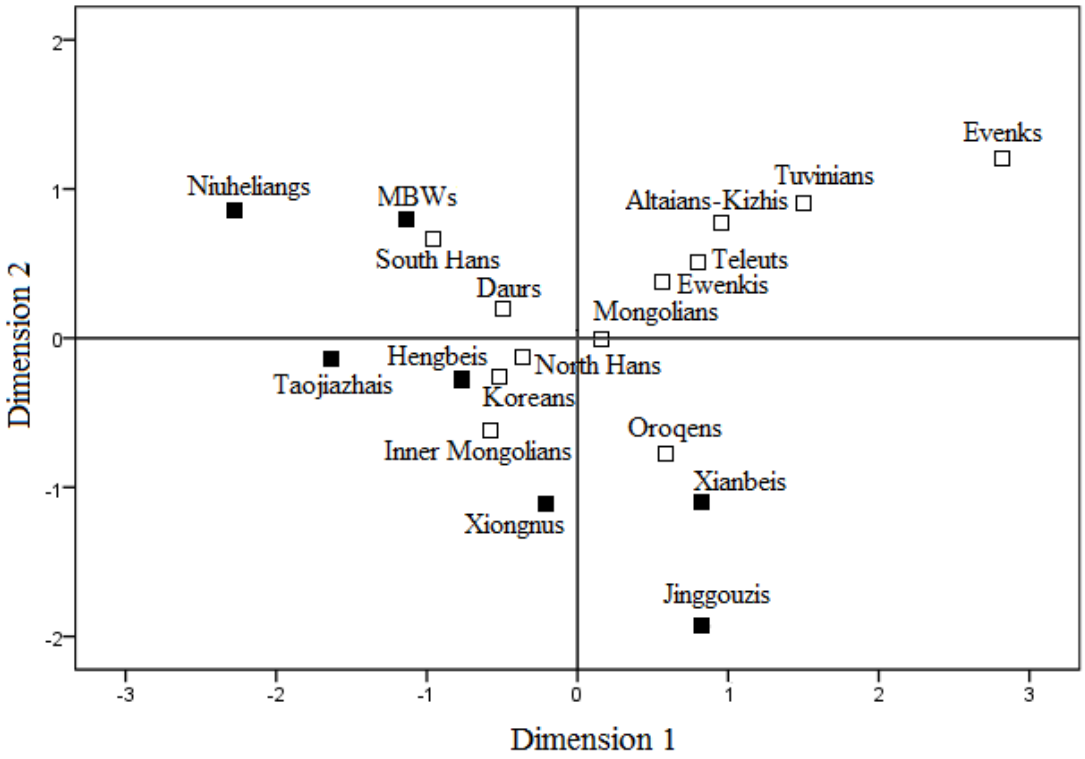
Evenks
Tuvinians
Niuheliangs
MBWs
Altaians-Kizhis
South Hans
Teleuts
Ewenkis
Daurs
Mongolians
Hengbeis
North Hans
Taojiazhais
Koreans
Oroqens
Inner Mongolians
Xianbeis
Xiongnus
Jinggouzis
Dimension 2
Dimension 1
2
1
0
-1
-2
-3
-2
-1
0
1
2
3

**Figure 2**

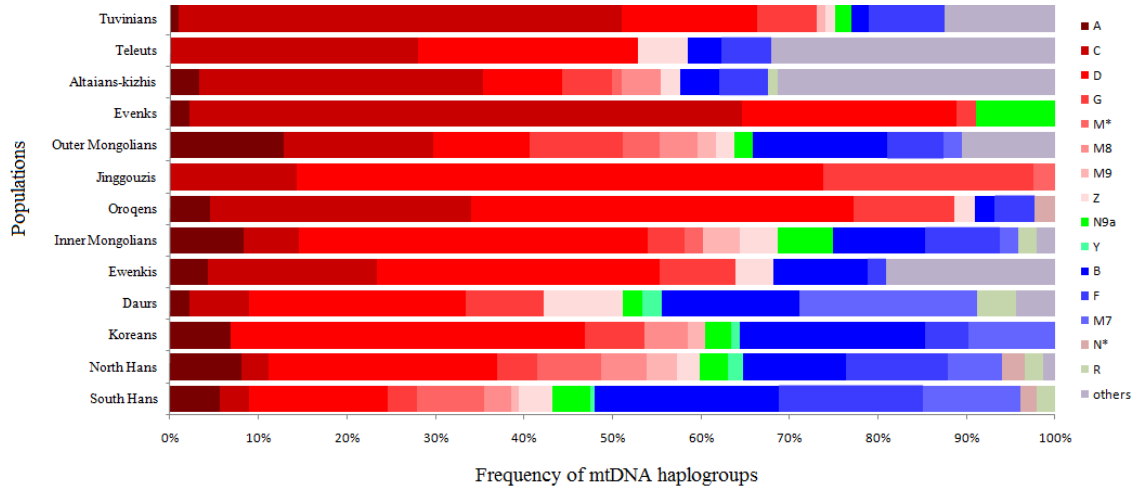
Tuvinians
Teleuts
Altaians-kizhis
Evenks
Outer Mongolians
Jinggouzis
Oroqens
Inner Mongolians
Ewenkis
Daurs
Koreans
North Hans
South Hans
Populations
0%
10%
20%
30%
40%
50%
60%
70%
80%
90%
100%
Frequency of mtDNA haplogroups
A
C
D
G
M*
M8
M9
Z
N9a
Y
B
F
M7
N*
R
others

**Figure 3**

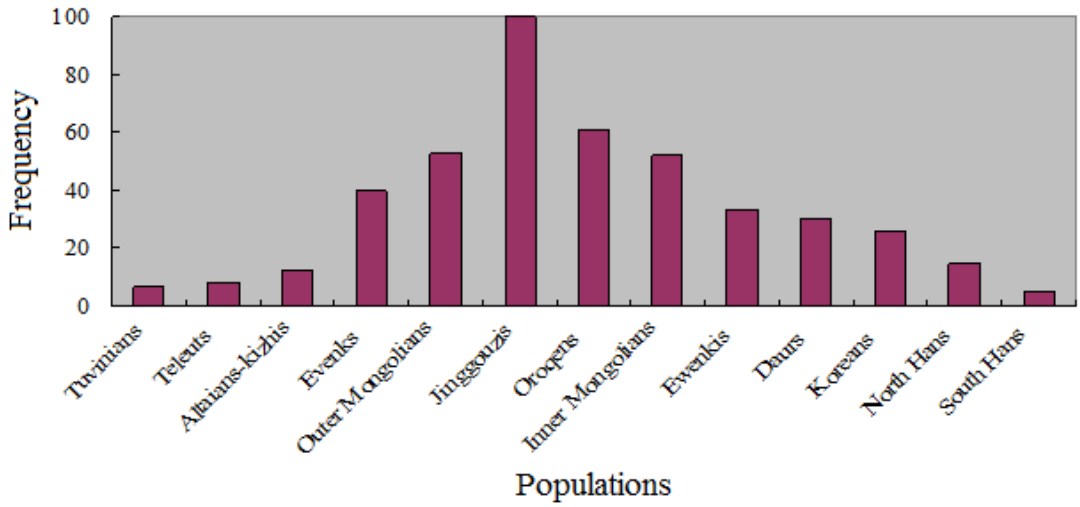
100
80
60
40
20
0
Frequency
Tuvinians
Teleuts
Altaians-Kizhis
Evenks
Outer Mongolians
Jinggouzis
Oroqens
Inner Mongolians
Ewenkis
Daurs
Koreans
North Hans
South Hans
Populations

**Figure 4**