

## ORIGINAL ARTICLE

# Mitochondrial haplogroup M9a1a1c1b is associated with hypoxic adaptation in the Tibetans

Qian Li, Keqin Lin, Hao Sun, Shuyuan Liu, Kai Huang, Xiaoqin Huang, Jiayou Chu and Zhaoqing Yang

While hypoxic environment at high altitude remains a major challenge for travelers from low-altitude areas, Tibetans have adapted to the high-altitude environment. Mitochondria are the energy conversion and supplement centers in eukaryotic cells. In recent years, studies have found that the diversity of the mitochondrial genome may have a role in the adaptation to hypoxia in Tibetans. In this study, mitochondrial haplogroup classification and variant genotyping were performed in Tibetan and Han Chinese populations living at different altitudes. The frequencies of mitochondrial haplogroups B and M7 in the high-altitude population were significantly lower compared with those in the low-altitude population ( $P=0.003$  and  $0.029$ , respectively), whereas the frequencies of haplogroups G and M9a1a1c1b in the high-altitude group were significantly higher compared with those in the low-altitude group ( $P=0.01$  and  $0.002$ , respectively). The frequencies of T3394C and G7697A, which are the definition sites of haplogroup M9a1a1c1b, were significantly higher in the high-altitude group compared with that in the low-altitude group ( $P=0.012$  and  $0.02$ , respectively). Our results suggest that mitochondrial haplogroups B and M7 are associated with inadaptability to hypoxic environments, whereas haplogroups G and M9a1a1c1b may be associated with hypoxic adaptation. In particular, the T3394C and G7697A variants on haplogroup M9a1a1c1b may be the primary cause of adaptation to hypoxia. *Journal of Human Genetics* (2016) 61, 1021–1026; doi:10.1038/jhg.2016.95; published online 28 July 2016

## INTRODUCTION

High-altitude adaptation occurs when populations reside at high altitudes in hypoxic environment. With human migration on a global scale during the past tens of thousands of years, modern humans have faced severe environmental pressures, particularly at high altitude.<sup>1</sup> Exposure factors including low atmospheric pressure, low oxygen content, cold and solar radiation are common in high-altitude areas, with extreme hypobaric hypoxia as one of the major characteristics. At a height of 4000 m above sea level, the oxygen content in the air is only 60% of that at sea level.<sup>2</sup> Travelers from low-altitude areas often suffer from hypoxic symptoms and may develop mountain sickness, including high-altitude pulmonary hypertension and pulmonary edema,<sup>3,4</sup> whereas aboriginals in high-altitude areas have adapted to the hypobaric hypoxia environment. The Qinghai–Tibet Plateau is the highest and largest plateau in the world. It consists of an area of more than 2.5 million square kilometers with an average altitude of >4000 m.<sup>5</sup> Tibetans are mainly distributed among China's Tibet Autonomous Region and some other provinces, including Qinghai, Gansu, Sichuan and Yunnan. Compared with populations living at sea level, Tibetans have higher levels of carbon monoxide and pulmonary ventilation, and lower levels of oxygen saturation and hemoglobin.<sup>2,6–9</sup> These physiological differences suggest that Tibetans have adapted to their high-altitude hypoxic environment. Therefore, characterizing the genomic diversity of Tibetans living at different altitudes should reveal

the genetic basis and mechanisms of their high-altitude hypoxic adaptation.

Mitochondria are the main energy conversion and supplement centers in eukaryotic cells. Mitochondrial DNA (mtDNA) consists of 16 569 bp of circular DNA containing genes encoding 13 peptides, 2 rRNAs and 22 tRNAs, as well as containing an 1121-bp regulatory region (the D-Loop). The mitochondrial respiratory chain consists of five complexes. The mitochondrial genome encodes seven subunits of complex I, one subunit of complex III, three subunits of complex IV and two subunits of complex V.<sup>10,11</sup> Mitochondria are the location of oxidative phosphorylation and are significantly involved in cellular oxygen consumption. Therefore, the structure and function of mitochondria change and energy production is suppressed during hypoxia, possibly explaining the acute mountain sickness and high-altitude pulmonary edema among travelers from low-altitude areas. Research has shown that the mitochondrial content of the Tibetan population is lower compared with that in the Han population. However, the body's oxygen consumption is not affected because of the higher metabolic efficiency of mitochondria in the Tibetan population.<sup>2</sup> Research on the mitochondrial genome has shown that there is a significant difference in the frequency of the mitochondrial 3010G–3970C haplotype between high- and low-altitude populations. This haplotype is believed to be associated with improved adaptability to the low-oxygen environment among the Tibetan population.<sup>12</sup> Another study showed that the mitochondrial variant T3394C of the

M9 haplogroup resulted in improved mitochondrial complex I activity, suggesting that this site is specifically associated with adaptation to high-altitude hypoxia.<sup>13</sup>

In a previous study, we compared and analyzed the whole sequence of mitochondrial DNA in 40 Tibetan and 50 Han Chinese subjects and found significant frequency differences at 18 single-nucleotide polymorphisms (SNPs). Selective adaptation might exist for four of the protein-coding variants and three of the non-protein-coding variants.<sup>14</sup> As a follow-up, in this study, mitochondrial haplogroups were genotyped in three Tibetan populations living at different altitudes and a Han population living at low altitude. Nine SNPs in the non-D-loop region found in the previous study were also genotyped to determine whether the mitochondrial genome has a role in the adaptation to hypoxia in Tibetans.

## MATERIALS AND METHODS

### Subjects

A total of 144 unrelated Tibetans and 47 unrelated Han Chinese were collected from the Tibetan Plateau, the Yunnan–Guizhou Plateau and the North China Plain. The Tibetan populations came from three altitudes, including 50 individuals from Lhasa city in the Tibetan Autonomous Region at an altitude of 3650 m, 46 individuals from Guinan County from Qinghai Province at an altitude of 3100 m and 46 individuals from Nujiang County in Yunnan Province at an altitude of 1500 m. The Han population included 47 Han Chinese individuals from Zhouping County in Shandong Province from an altitude of 50 m. Peripheral blood samples were collected, and the B lymphocytes were transformed into immortalized lymphoblastoid cell lines by the Epstein–Barr virus. The acquired immortal lines were used for DNA extraction and further biological research. Genomic DNA of these immortalized lymphoblastoid cell lines was extracted using the AxyPre Blood Genomic DNA Miniprep Kit (Axygen, Suzhou, China) according to the recommendations of the manufacturer. This study was approved by the ethics committee of the Institute of Medical Biology, Peking Union Medical College (Kunming, China) and fulfilled the requirement of informed consent.

### mtDNA haplogroup classification

The mtDNA control region sequence of each individual was amplified using a primer pair (L15594: 5'-CGCCTACACAATTCTCCGATC-3' and H901: 5'-ACTTGGGTTAATCGTGTGACC-3') and under previously reported PCR conditions.<sup>15</sup> The PCR products were purified and sequenced using an ABI 3130 genetic analyzer (Applied Biosystems, Foster City, CA, USA) (Table 1). The sequences were compared with the revised Cambridge reference sequence<sup>11</sup> to identify variants in the mtDNA sequence in each subject. Based on PhyloTree (<http://phylotree.org>; mtDNA tree Build 16, 19 February 2014),<sup>16</sup> mtDNA haplogroups were classified using a previously described method,<sup>15</sup> and the results were estimated by the HaploGrep 2.0 software (<https://haplogrep.uibk.ac.at/>).

### mtDNA variant genotyping

We directly sequenced the PCR products to genotype the A1041G, T3394C, G4491A, T5628C, 5899-5900inC, T7142C, G7697A, T7738C and G13135A polymorphisms in all subjects. All primers for PCR amplification and sequencing were designed using the Primer 5 software (PREMIER Biosoft International, Palo Alto, CA, USA, Table 2). PCR amplification was performed

**Table 1** The sequencing primer of mtDNA control region

Primer	Sequence (5'–3')
29	GGTCTATCACCCCTATTAACCAC
L15996	CTCCACCATTAGCACCCAAAGC
H901	ACTTGGGTTAATCGTGTGACC
H16525	ACGTGTGGGCTATTAGGCT

Abbreviation: mtDNA, mitochondrial DNA.

under the following conditions: a predenaturation cycle at 94 °C for 5 min; 30 cycles at 94 °C for 30 s, 56 °C for 30 s, 72 °C for 30 s; and a final extension cycle at 72 °C for 10 min. The conservation index of these variants was analyzed by using MitoTool ([www.mitotool.org](http://www.mitotool.org)).<sup>17</sup>

### Statistical analyses

Statistical analyses were performed using the SPSS 17.0 software (Chicago, IL, USA). Pearson's  $\chi^2$  test was used to assess the significance of the differences in haplogroup and SNP frequencies between high- and low-altitude groups. Fisher's exact test was applied to those cases genotype < 5. *P*-values below 0.05 were considered to be statistically significant. All *P*-values were adjusted by using Bonferroni correction.

## RESULTS

### Distribution of mitochondrial haplogroups among the four populations

All of the mitochondrial haplogroups in the four studied populations living at different altitudes belonged to the M and N macrohaplogroups; a total of 24 haplogroups were observed among the different populations (Supplementary Table 1). Tibetans from the Tibet Autonomous Region were found to have 13 haplogroups, with D4, M9 and R9 as the major haplogroups. Tibetans from Qinghai Province were found to have 12 haplogroups, with D4, G and R9 as the major haplogroups. Tibetans from Yunnan Province were found to have 10 haplogroups, with D4 and M8 as the major haplogroups. Han Chinese from Shandong Province were found to have 14 haplogroups, with D4 and R9 as the major haplogroups (Table 3).

**Table 2** The PCR primers and DNA sequencing primers of mtDNA SNPs

Site	Primer name	PCR primer sequence (5'–3')	Sequencing primer
A1041G	L905	ACACGATTAACCCAAGTCAAT	H1163
	H1163	TGAAGCACCGCCAGGTCCTTTGAG	
T3394C	L2691	TGAAGAGGCGGGCATAACACA	H3708
	H3708	GAGATTGTTGGGCTACTGCT	
G4491A	L4144	TTCCGCTACGACCAACTCATA	H5334
	H5334	ATTAGCGCTAGGTAGAAGTAGAGGTT	
T5628C	L5113	CATTCTACTACTCAACT	H5915
	H5915	GTAGAGAATAGTCAACGG	
5899-5900inC	L5113	CATTCTACTACTCAACT	H6251
	H6251	CTGCTCCGGCCTCCACTATA	
T7142C	L6962	GACTGGCATTGTATTAGCA	L6962
	H7779	ATGGCGGGCAGGATAGTTC	
G7697A	L6962	GACTGGCATTGTATTAGCA	H7779
	H7779	ATGGCGGGCAGGATAGTTC	
T7738C	L6962	GACTGGCATTGTATTAGCA	H7779
	H7779	ATGGCGGGCAGGATAGTTC	
G13135A	L13007	CAGCCCAATTAGGTCTC	L13007
	H13173	AGAGTGGTGATAGCGCTAAG	

Abbreviations: mtDNA, mitochondrial DNA; SNP, single-nucleotide polymorphism.

**Table 3 Distribution of haplogroup frequencies in four populations**

Haplogroup	XiZang	QingHai	YunNan	ShangDong
	Tibetan (n = 50)	Tibetan (n = 46)	Tibetan (n = 48)	Han (n = 47)
A	6 (12.0%)	5 (10.9%)	5 (10.4%)	2 (4.3%)
B	1 (2.0%)		4 (8.3%)	7 (14.9%)
D4	10 (20.0%)	8 (17.4%)	19 (39.6%)	10 (21.3%)
D5		1 (2.2%)	1 (2.1%)	
G	4 (8.0%)	8 (17.4%)		2 (4.3%)
H	1 (2.0%)	5 (10.9%)	2 (4.2%)	2 (4.3%)
M5		2 (4.3%)		
M7				5 (10.6%)
M8	5 (10.0%)	3 (6.5%)	10 (20.8%)	4 (8.5%)
M9a	11 (22.0%)	3 (6.5%)	2 (4.2%)	
M10				1 (2.1%)
M12			1 (2.1%)	
M13	1 (2.0%)	1 (2.2%)		
M35	1 (2.0%)			
M43				1 (2.1%)
M62	1 (2.0%)	2 (4.3%)		
M74				1 (2.1%)
N9				2 (4.3%)
N10				1 (2.1%)
P6			1 (2.1%)	
R9	8 (16.0%)	7 (15.2%)	3 (6.3%)	8 (17.0%)
T				1 (2.1%)
U8		1 (2.2%)		
X2m	1 (2.0%)			

**Table 4 Distribution of mtDNA SNPs in four populations**

Allele	XiZang	QingHai	YunNan	ShanDong
	Tibetan (n = 50)	Tibetan (n = 46)	Tibetan (n = 48)	Han (n = 47)
A1041G	11 (22%)	5 (10.9%)	2 (4.2%)	
T3394C	13 (26%)	3 (6.5%)	3 (6.3%)	2 (4.2%)
G4491A	11 (22%)	3 (6.5%)	2 (4.2%)	
T5628C		1 (2.2%)	1 (2.1%)	2 (4.2%)
5899-5900inC	8 (16%)	4 (8.7%)		
T7142C	4 (8%)	2 (4.4%)		1 (2.1%)
G7697A	7 (14%)	3 (6.5%)		
T7738C				1 (2.1%)
G13135A	1 (2%)	1 (2.2%)		2 (4.2%)

Abbreviations: mtDNA, mitochondrial DNA; SNP, single-nucleotide polymorphism.

**mtDNA SNP distribution among the four populations**

The mtDNA SNP distribution among the four populations varied (Table 4). Specifically, the T5628C and T7738C variants were missing in Tibetans living in the Tibet Autonomous Region. Additionally, we did not detect the 7738C variant among Qinghai Tibetans. Only four variants (A1041G, T3394C, G4491A and T5628C) were detected among Yunnan Tibetans. Five variants (T3394C, T5628C, T7142C, T7738C and G13135A) were detected in Shandong Han Chinese.

**Association between haplogroups and high-altitude adaptation**

We divided the four populations into a high-altitude group (populations living at an altitude over 3000 m) and a low-altitude group (those living at an altitude below 3000 m). Tibetans living in the Tibet Autonomous Region and Qinghai Province were included in the high-altitude group, and Yunnan Tibetans and Shandong Hans were

**Table 5 Comparison of the haplogroup frequencies between high- and low-altitude group**

Haplogroup	High-altitude	Low-altitude	P-value
	group (n = 96)	group (n = 95)	
A	11	7	0.333
B	1	11	0.003 <sup>a</sup>
D4	18	29	0.059
D5	1	1	1.00
G	12	2	0.010 <sup>a</sup>
H	6	4	0.747
M5	2	0	0.497
M7	0	5	0.030 <sup>a</sup>
M8	8	14	0.166
M9a	14	2	0.003 <sup>a</sup>
M10	0	1	0.497
M12	0	1	0.497
M13	2	0	0.497
M35	1	0	1.00
M43	0	1	0.497
M62	3	0	0.246
M74	0	1	0.497
N9	0	2	0.246
N10	0	1	0.497
P6	0	1	0.497
R9	15	11	0.415
T	0	1	0.497
U8	1	0	1.00
X2m	1	0	1.00

<sup>a</sup>P-value <0.05.

classified into the low-altitude group. The results show a haplogroup difference between high- and low-altitude groups. The frequencies of haplogroups B and M7 in the high-altitude group were significantly lower compared with those in the low-altitude group ( $P=0.003$  and  $0.03$ , respectively). In contrast, the frequencies of haplogroups G and M9a in the high-altitude group were significantly higher compared with those in the low-altitude group ( $P=0.003$  and  $0.01$ , respectively; Table 5). However the the  $P$ -values were  $>0.05$  by using Bonferroni correction.

**Association between mtDNA SNPs and high-altitude adaptation**

The frequencies of A1041G, T3394C, G4491A, 5899-5900inC and G7697A in the high-altitude group were significantly higher compared with those in the low-altitude group ( $P<0.05$ ), of which 5899-5900inC and G7697A were only found in the high-altitude populations (Table 6), but the  $P$ -values were  $>0.05$  by using Bonferroni correction. Conservation analysis of four variants (A1041G, T3394C, G4491A and G7697A) that may cause amino-acid changes showed that A1041G and G4491A were not conserved during evolution, whereas T3394C and T7697C were highly conserved (Table 7).

**Analysis of mitochondrial haplogroups combined with SNPs**

As T3394C and G7697A are definition sites of the M9a haplogroup, this haplogroup was further divided into M9a1b and M9a1a1c1b. In this study, we found these two haplogroups to be present in the high-altitude group, whereas M9a1a1c1b was not present in the low-altitude group because G7697A was missing from this group. The frequency of haplogroup M9a1a1c1b in the high-altitude group was significantly higher compared with that in the low-altitude group ( $P=0.002$ ; Table 8), and the  $P$ -values were  $<0.05$  after Bonferroni correction.

A combined analysis of the T3394C and G7697A variants together with the haplogroup showed that G7697A was only found in the high-altitude group on haplogroup M9a1a1c1b. Additionally, T3394C was mainly present in the high-altitude group on haplogroup M9a1a1c1b. In addition, T3394C was also found in populations with the haplogroups M9a1b, D4 and M8 in the high-altitude group, whereas it was also present in haplogroups M9a1b, D4 and M7 in the low-altitude group (Table 9).

## DISCUSSION

In this study, we compared mitochondrial haplogroups and nine SNPs previously reported to be potentially associated with adaptation to hypoxia in four Chinese populations living in places with different altitudes, divided into high- and low-altitude groups divided by an altitude of 3000 m. We found that the frequencies of the A1041G, T3394C, G4491A, 5899-5900inC and G7697A variants in the high-altitude group were significantly higher compared with those in the low-altitude population ( $P < 0.05$ ). Furthermore, the frequencies of haplogroups B and M7 in the high-altitude population were significantly lower compared with those in the low-altitude population ( $P < 0.05$ ). In contrast, the frequencies of haplogroups G and M9a in the high-altitude population were significantly higher compared with those in the low-altitude population ( $P < 0.05$ ). In addition, the frequency of haplogroup M9a1a1c1b, a subhaplogroup of M9a, in the high-altitude population was significantly higher compared with that in the low-altitude population ( $P = 0.002$ ), whereas there was no significant difference in the distribution of subhaplogroup M9a1b between the low- and high-altitude groups.

It has been reported that the frequencies of the B, M7 and G haplogroups are 0–9%, 0–2.3% and 1.8–16.4%, respectively, in Tibetan populations living in areas at an altitude higher than 3000 m.<sup>13,18</sup> The frequencies of these haplogroups have been found to be 12–30.4%, 3.3–18.6% and 0–7.8%, respectively, in 13 low-altitude

Han populations.<sup>19</sup> Thus, previous studies suggest that the frequencies of haplogroups B and M7 in the high-altitude Tibetan group are lower compared with those in low-altitude Han populations, whereas the frequency of haplogroup G is higher in the high-altitude Tibetan group. In our study, we also found that the frequencies of haplogroups B and M7 in the high-altitude Tibetan group were significantly lower compared with those in the low-altitude group. Previous studies suggest that haplogroups B and M7 may be the risk factors of acute mountain sickness in Han populations,<sup>20</sup> and haplogroup B might be the risk factor for high-altitude pulmonary edema among Han populations.<sup>21</sup> As these haplogroups are considered to be risk factors of acute mountain sickness and high-altitude pulmonary edema in Han populations, we hypothesize that these two haplogroups may contribute to inadaptability to hypoxic environment. Additionally, because the frequency of haplogroup G in the high-altitude Tibetan group was significantly higher compared with that in the low-altitude population in our study, this haplogroup might be involved in the adaptation to hypoxia. Thus, we speculate that some variants on the B, M7 and G haplogroups may cause functional alterations that affect adaptability to high-altitude hypoxic environments.

Haplogroup M9 is divided into two subhaplogroups, M9a and E. Haplogroup E is mainly found in Southeast Asia and on some islands, including Taiwan.<sup>22,23</sup> M9a is widely distributed in populations of eastern Asia, central Asia and northern Asia at various frequencies, with the highest frequency in the Tibetan population.<sup>24,25</sup>

**Table 6 Comparison of the frequency of mtDNA SNPs between high- and low-altitude group**

Allele	High-altitude group (n = 96)	Low-altitude group (n = 95)	P-value
A1041G	16	2	0.001 <sup>a</sup>
T3394C	16	5	0.012 <sup>a</sup>
G4491A	14	2	0.003 <sup>a</sup>
T5628C	1	3	0.621
5899-5900inC	12	0	<0.001 <sup>a</sup>
T7142C	6	1	0.118
G7697A	10	0	0.020 <sup>a</sup>
T7738C	0	1	1
G13135A	2	2	1

Abbreviations: mtDNA, mitochondrial DNA; SNP, single-nucleotide polymorphism.  
<sup>a</sup> $P$ -value < 0.05.

**Table 7 Conservation analysis of four SNPs with significant difference between high- and low-altitude group**

Allele	Conservation index (%)	Reported (population context)	Reported (disease context)	Haplogroup-specific variant
A1041G	34.61	Yes	No	M9a1
T3394C	94.23	Yes	Yes	C4a4, M9a'b, D4g1a, J1c1
G4491A	17.30	Yes	No	M9, M73b, J1c14
T7697C	98.07	Yes	Yes	C1d1, M9a1a1c1b

Abbreviation: SNP, single-nucleotide polymorphism.

**Table 8 Comparison of subhaplogroup frequencies of M9a between high- and low-altitude group**

Haplogroup	High-altitude group (n = 96)	Low-altitude group (n = 95)	P-value
M9a			
M9a1a1c1b	10	0	0.002 <sup>a</sup>
M9a1b	4	2	0.683

<sup>a</sup> $P$ -value < 0.05.

**Table 9 The distribution of 3394C and 7697A mutations in haplogroups in high- and low-altitude populations**

Haplogroup	High-altitude group		Low-altitude group	
	T3394C (n = 16)	G7697A (n = 10)	T3394C (n = 5)	G7697A (n = 0)
M9a1a1c1b	10	10		
M9a1b	4		2	
D4	1		2	
M7			1	
M8	1			

The frequency of haplogroup M9a differs among different Tibetan populations, ranging from 2.3 to 25.4%.<sup>18,26</sup> The average frequency of haplogroup M9a in the Tibetan population is 16.4%, and there is a significant difference in the frequency of haplogroup M9a between high-altitude Tibetan populations and low-altitude Han populations ( $P < 0.0001$ ).<sup>13</sup> In our study, the frequency of haplogroup M9a in the three Tibetan groups ranged from 4.2 to 22%, whereas it was not found in the Han population. Moreover, there was a significant difference in the frequency haplogroup M9a between the high-altitude Tibetan group and the low-altitude populations. However, when haplogroup M9a was further divided into subhaplogroups, M9a1b was found in both the high- and low-altitude populations, without a significant difference in distribution between the populations. However, subhaplogroup M9a1a1c1b was only found in the high-altitude population with a frequency of 10.42%, and this was significantly higher compared with that in the low-altitude populations ( $P = 0.002$ ). Therefore, we speculate that the M9a1a1c1b subhaplogroup might be associated with adaptation to hypoxia. Variants in this haplogroup may lead to changes in mitochondrial function, enabling adaptation to the hypoxic environment. Interestingly, four out of five variants (A1041G, T3394C, G4491A and G7697A) found to be significantly different between the high- and low-altitude populations in our study were definition sites of haplogroup M9a1a1c1b.

The A1041G variant of 12sRNA has an important role in the maintenance of RNA structure. When the A allele is changed to a G, the consequent RNA conformational change may cause a decrease in free energy, thus making the RNA structure more stable. The ND2 4491A variant may cause the eighth amino acid to change from valine to isoleucine, causing a conformational change from a helix to a sheet, leading to a decrease in the free energy of the protein, and increasing the stability of the protein.<sup>14</sup> In our study, there were significant differences in the frequencies of A1041G, G4491A and 5899-5900inC between the high- and low-altitude populations. Both A1041G and G4491A are definition sites of haplogroup M9a or other haplogroups, and there is insufficient evidence indicating the association between these two sites and human disorders. Conservation analysis showed that these two variants are not evolutionarily conserved and that they are polymorphic. Additionally, 5899-5900inC is a non-coding variant and a definition site of some mitochondrial haplogroups. Therefore, we deduced that the A1041G, G4491A, and 5899-5900inC variants may not affect adaptation to hypoxia in high-altitude populations, although these three sites had different frequencies among populations living at different altitudes.

The G7697A variant of cytochrome *c* oxidase subunit II (COXII) leads to a change of valine to isoleucine at amino-acid position 38, leading to a slight change in the protein conformation and a slight increase in protein stability.<sup>14</sup> The frequency of G7697A in the Tibetan population is significantly higher compared with that in the Beijing Han population.<sup>14</sup> In our study, no G7697A alleles were found in the low-altitude populations, and the frequency of G7697A in the high-altitude populations was significantly higher ( $P = 0.02$ ).

The T3394C variant in ND1 results in a change from tyrosine at residue 30 to histidine, thus changing the hydrophobicity of the protein. As seen in structural predictions of the protein, this variant may cause a change in protein conformation from a compact sphere to an incompact strip, thus decreasing the stability of the protein. The T3394C variant is also associated with Leber's hereditary optic neuropathy and diabetes. In low-altitude populations, T3394C variants may increase the risk of Leber's hereditary optic neuropathy.<sup>27-30</sup> A previous study reported that the frequency of T3394C in the

high-altitude population was significantly higher Leber's compared with that in the low-altitude Han population.<sup>13,14</sup> Our study also found this to be the case ( $P = 0.012$ ). The conservation analysis also showed that G7697A and T3394C are highly conserved variants.

Combining the analysis of T3394C and G7697A with the M9a1a1c1b haplogroup in our study showed that all samples with this haplogroup contain the T3394C and G7697A variants. The presence of T3394C on haplogroup M9a may enhance the activity of mitochondrial complex I, which may cause an expression level change of hypoxia-inducible factor 1 and 2. The increase in complex I activity may allow the body to adapt to the hypoxic environment.<sup>13</sup> The G7697A variant is located on the COXII encoding cytochrome *C* oxidase subunit 2 (COXII). Cytochrome *C* oxidase, also known as mitochondrial complex IV, is the rate-limiting enzyme of the oxidative phosphorylation process in the mitochondrial respiratory chain and is directly involved in mitochondrial oxygen consumption and energy generation. The COXII subunit carries binding sites for cytochrome, which can be combined with cytochrome *C*, consequently leading to an oxidoreduction reaction at the mitochondrial cytoplasmic surface. Finally, electrons are transferred to O<sub>2</sub>, and H<sup>+</sup> is pumped out of the mitochondrial inner membrane for ATP synthesis. It is possible that a conformational change of the protein caused by G7697A may affect the activity of cytochrome *C* oxidase, leading to changes in the body's oxygen consumption and ATP production rate that eventually affect oxygen consumption in hypoxic environment. Haplogroup M9a1a1c1b is likely involved in adaptation to hypoxia, mainly due to the T3394C and G7697A variants. These two variants can improve the activity of the mitochondrial respiratory chain complexes I and IV, thereby allowing the body to adapt to hypoxic environment.

MtDNA is subject to features of maternal inheritance, such as the lack of recombination and a high evolutionary rate, which make it feasible for studies on human evolution, population migration and environmental adaptation. Additionally, mitochondrial diversity is under selection pressure of the environment. Therefore, certain haplogroups of mtDNA may be enriched in particular environments.<sup>20</sup> In our study, we found two mitochondrial haplogroups (M9a1a1c1b and G) that were enriched in populations living in hypoxic environments. We also explored the possible hypoxic adaptation mechanisms of haplogroup M9a1a1c1b. Meanwhile, it should be noted that we divided different ethnic groups into a low-altitude group and a high-altitude group to minimize bias caused by differences between ethnic groups. We have carried out the Bonferroni correction for multiple test, and the corrected *P*-values were  $> 0.05$ . Although our statistical power was limited by the sample size in this study, we attempted to minimize bias by analyzing and comparing our SNP frequencies with previously published data. And our results still show a significant trend of natural selection on mitochondrial haplogroups.

In summary, the mitochondrial haplogroups B and M7 may be negatively correlated with adaptation to hypoxia and may cause inadaptability to a hypoxic environment. In contrast, haplogroups G and M9a1a1c1b are enriched in populations living in hypoxic environments. Haplogroup M9a1a1c1b carries the T3394C and G7697A variants, which may increase the activity of mitochondrial respiratory chain complexes, thereby leading to hypoxic adaptation. Therefore, we hypothesize that the mitochondrial genome of the Tibetans living at altitudes above 3000 m may be under selection from the high-altitude hypoxic environment.

#### CONFLICT OF INTEREST

The authors declare no conflict of interest.

## ACKNOWLEDGEMENTS

We are grateful to the participants of this study. The study was supported by the National Natural Science Foundation of China (No. 30971578) and the Chinese National High Technology Research and Development Program (No. 2012AA021802).

- 1 Bigam, A., Bauchet, M., Pinto, D., Mao, X., Akey, J. M., Mei, R. *et al.* Identifying signatures of natural selection in Tibetan and Andean populations using dense genome scan data. *PLoS Genet.* **6**, e1001116 (2010).
- 2 Beall, C. M. Two routes to functional adaptation: Tibetan and Andean high-altitude natives. *Proc. Natl Acad. Sci. USA* **104**(suppl. 1): 8655–8660 (2007).
- 3 Leon-Velarde, F., Maggiorini, M., Reeves, J. T., Aldashev, A., Asmus, I., Bernardi, L. *et al.* Consensus statement on chronic and subacute high altitude diseases. *High Alt. Med. Biol.* **6**, 147–157 (2005).
- 4 Wu, T. Life on the high Tibetan plateau. *High Alt. Med. Biol.* **5**, 1–2 (2004).
- 5 Wu, T. The Qinghai-Tibetan plateau: how high do Tibetans live? *High Alt. Med. Biol.* **2**, 489–499 (2001).
- 6 Beall, C. M. Andean, Tibetan, and Ethiopian patterns of adaptation to high-altitude hypoxia. *Integr. Comp. Biol.* **46**, 18–24 (2006).
- 7 Beall, C. M., Brittenham, G. M., Strohl, K. P., Blangero, J., Williams-Blangero, S., Goldstein, M. C. *et al.* Hemoglobin concentration of high-altitude Tibetans and Bolivian Aymara. *Am. J. Phys. Anthropol.* **106**, 385–400 (1998).
- 8 Moore, L. G. Human genetic adaptation to high altitude. *High Alt. Med. Biol.* **2**, 598–604 (2001).
- 9 Wu, T., Wang, X., Wei, C., Cheng, H., Wang, X., Li, Y. *et al.* Hemoglobin levels in Qinghai-Tibet: different effects of gender for Tibetans vs. Han. *J. Appl. Physiol.* **98**, 598–604 (2005).
- 10 Schon, E. A., DiMauro, S. & Hirano, M. Human mitochondrial DNA: roles of inherited and somatic variants. *Nat. Rev. Genet.* **13**, 878–890 (2012).
- 11 Andrews, R. M., Kubacka, I., Chinnery, P. F., Lightowlers, R. N., Turnbull, D. M., Howell, N. *et al.* Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat. Genet.* **23**, 147 (1999).
- 12 Luo, Y., Gao, W., Liu, F. & Gao, Y. Mitochondrial nt3010G-nt3970C haplotype is implicated in high-altitude adaptation of Tibetans. *Mitochondrial DNA* **22**, 181–190 (2011).
- 13 Ji, F., Sharpley, M. S., Derbeneva, O., Alves, L. S., Qian, P., Wang, Y. *et al.* Mitochondrial DNA variant associated with Leber hereditary optic neuropathy and high-altitude Tibetans. *Proc. Natl Acad. Sci. USA* **109**, 7391–7396 (2012).
- 14 Gu, M., Dong, X., Shi, L., Shi, L., Lin, K., Huang, X. *et al.* Differences in mtDNA whole sequence between Tibetan and Han populations suggesting adaptive selection to high altitude. *Gene* **496**, 37–44 (2012).
- 15 Zhang, A. M., Jia, X., Bi, R., Salas, A., Li, S., Xiao, X. *et al.* Mitochondrial DNA haplogroup background affects LHON, but not suspected LHON, in Chinese patients. *PLoS ONE* **6**, e27750 (2011).
- 16 van Oven, M. & Kayser, M. Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. *Hum. Mutat.* **30**, E386–E394 (2009).
- 17 Fan, L. & Yao, Y. G. MitoTool: a web server for the analysis and retrieval of human mitochondrial DNA sequence variations. *Mitochondrion* **11**, 351–356 (2010).
- 18 Qin, Z., Yang, Y., Kang, L., Yan, S., Cho, K., Cai, X. *et al.* A mitochondrial revelation of early human migrations to the Tibetan Plateau before and after the last glacial maximum. *Am. J. Phys. Anthropol.* **143**, 555–569 (2010).
- 19 Yao, Y. G., Kong, Q. P., Bandelt, H. J., Kivisild, T. & Zhang, Y. P. Phylogeographic differentiation of mitochondrial DNA in Han Chinese. *Am. J. Hum. Genet.* **70**, 635–651 (2002).
- 20 Li, F. X., Ji, F. Y., Zheng, S. Z., Yao, W., Xiao, Z. L. & Qian, G. S. MtDNA haplogroups M7 and B in southwestern Han Chinese at risk for acute mountain sickness. *Mitochondrion* **11**, 553–558 (2011).
- 21 Liu, Y. J., Gao, W. X., Li, S. Z., Huang, X. W., Chen, Y., Liu, F. Y. *et al.* Mitochondrial haplogroup D4 confers resistance and haplogroup B is a genetic risk factor for high-altitude pulmonary edema among Han Chinese. *Genet. Mol. Res.* **11**, 3658–3667 (2012).
- 22 Soares, P., Trejaut, J. A., Loo, J. H., Hill, C., Mormina, M., Lee, C. L. *et al.* Climate change and postglacial human dispersals in southeast Asia. *Mol. Biol. Evol.* **25**, 1209–1218 (2008).
- 23 Tabbada, K. A., Trejaut, J., Loo, J. H., Chen, Y. M., Lin, M., Mirazón-Lahr, M. *et al.* Philippine mitochondrial DNA diversity: a populated viaduct between Taiwan and Indonesia? *Mol. Biol. Evol.* **27**, 21–31 (2009).
- 24 Torroni, A., Miller, J. A., Moore, L. G., Zamudio, S., Zhuang, J., Droma, T. *et al.* Mitochondrial DNA analysis in Tibet: implications for the origin of the Tibetan population and its adaptation to high altitude. *Am. J. Phys. Anthropol.* **93**, 189–199 (1994).
- 25 Yao, Y. G., Kong, Q. P., Wang, C. Y., Zhu, C. L. & Zhang, Y. P. Different matrilineal contributions to genetic structure of ethnic groups in the silk road region in China. *Mol. Biol. Evol.* **21**, 2265–2280 (2004).
- 26 Zhao, M., Kong, Q. P., Wang, H. W., Peng, M. S., Xie, X. D., Wang, W. Z. *et al.* Mitochondrial genome evidence reveals successful Late Paleolithic settlement on the Tibetan Plateau. *Proc. Natl Acad. Sci. USA* **106**, 21230–21235 (2009).
- 27 Brown, M. D., Torroni, A., Reckord, C. L. & Wallace, D. C. Phylogenetic analysis of Leber's hereditary optic neuropathy mitochondrial DNA's indicates multiple independent occurrences of the common variants. *Hum. Mutat.* **6**, 311–325 (1995).
- 28 Liang, M., Guan, M., Zhao, F., Zhou, X., Yuan, M., Tong, Y. *et al.* Leber's hereditary optic neuropathy is associated with mitochondrial ND1 T3394C variant. *Biochem Biophys. Res. Commun.* **383**, 286–292 (2009).
- 29 Liu, S. M., Zhou, X., Zheng, F., Li, X., Liu, F., Zhang, H. M. *et al.* Novel variants found in mitochondrial diabetes in Chinese Han population. *Diabetes Res. Clin. Pract.* **76**, 425–435 (2007).
- 30 Puomila, A., Hamalainen, P., Kivioja, S., Savontaus, M. L., Koivumaki, S., Huoponen, K. *et al.* Epidemiology and penetrance of Leber hereditary optic neuropathy in Finland. *Eur. J. Hum. Genet.* **15**, 1079–1089 (2007).

Supplementary Information accompanies the paper on Journal of Human Genetics website (<http://www.nature.com/jhg>)