A Mitochondrial Genome Sequence of the Tibetan Antelope (*Pantholops hodgsonii*)

Shu-Qing Xu[1,3,4], Ying-Zhong Yang[2], Jun Zhou[1], Guo-En Jin[2], Yun-Tian Chen[2], Jun Wang[1], Huan-Ming Yang[1], Jian Wang[1], Jun Yu[1], Xiao-Guang Zheng[1,3]*, and Ri-Li Ge[1,2]*

[1] Beijing Genomics Institute, Chinese Academy of Sciences, Beijing 101300, China; [2] Research Center for High Altitude Medicine, Qinghai University, Xining 810016, China; [3] Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming 650223, China; [4] Graduate School, Chinese Academy of Sciences, Beijing 100049, China.

To investigate genetic mechanisms of high altitude adaptations of native mammals on the Tibetan Plateau, we compared mitochondrial sequences of the endangered *Pantholops hodgsonii* with its lowland distant relatives *Ovis aries* and *Capra hircus*, as well as other mammals. The complete mitochondrial genome of *P. hodgsonii* (16,498 bp) revealed a similar gene order as of other mammals. Because of tandem duplications, the control region of *P. hodgsonii* mitochondrial genome is shorter than those of *O. aries* and *C. hircus*, but longer than those of *Bos* species. Phylogenetic analysis based on alignments of the entire cytochrome b genes suggested that *P. hodgsonii* is more closely related to *O. aries* and *C. hircus*, rather than to species of the *Antilopinae* subfamily. The estimated divergence time between *P. hodgsonii* and *O. aries* is about 2.25 million years ago. Further analysis on natural selection indicated that the COXI (cytochrome c oxidase subunit I) gene was under positive selection in *P. hodgsonii* and *Bos grunniens*. Considering the same climates and environments shared by these two mammalian species, we proposed that the mitochondrial COXI gene is probably relevant for these native mammals to adapt the high altitude environment unique to the Tibetan Plateau.

Key words: tibetan antelope, mitochondrial genome, adaptation, COXI

Introduction

As the most prominent terrestrial highland on the earth, the Tibetan Plateau enacts great effects on global climate and biosphere. Its fauna and flora prospering on the plateau are constantly challenged by the harsh environment of hypoxia, low temperature, high solar radiation, and lack of biological production. Native animals of the Tibetan Plateau, surviving over thousands of years on the highland, have developed various physiological, behavioral, and morphological strategies to cope with these environmental challenges; some of the changes are certainly attributable to phenotypic plasticity and others are genetic. The genomic mean for discovering inheritable changes in a species is to sequence its genomes, including both nuclear and organellar genomes.

Pantholops hodgsonii (Tibetan antelope or chiru) is an excellent representative of the native mammalian species, which has perfectly adapted to the Tibetan Plateau. Living at elevations from 3,700 to 5,500 meters (*1*), the antelope can run for hours at a spectacular speed. Such high speed and stamina indicate that it is very effective in utilizing limited oxygen supply and other energy resources on the plateau. Therefore, we chose *P. hodgsonii* as an initial subject to look for any adaptive genetic changes, at both DNA and protein levels, and molecular mechanisms for high altitude adaptations of native mammals that thrive on the Tibetan Plateau.

During the past few decades, mitochondrial DNA (mtDNA) has been widely used in studies of evolutionary biology and population genetics. As the locality of energy metabolism and hosting unique genetic material, mitochondrion and its genome are particularly interesting for high altitude biology. Mammalian mitochondrial genome is a circular, double-stranded

*** Corresponding authors.**
E-mail: zhengxg@genomics.org.cn;
geriligao@hotmail.com

molecule with a length of about 16 Kb. In general, it contains 13 protein-coding genes, 2 rRNA genes, 22 tRNA genes, and a non-coding control region (*2–4*). The thirteen proteins encoded by the genome are all related to oxide metabolism.

In the current study, we report a complete mitochondrial genome sequence from a single *P. hodgsonii* individual and results from comparative analysis in search of genetic outcomes in living under high altitude environments.

Results

General features of the *P. hodgsonii* mitochondrial genome

The *P. hodgsonii* mitochondrial genome is 16,498 bp long, shorter than those of *Ovis aries* (NC_001941) and *Capra hircus* (NC_005044), which are 16,616 bp and 16,640 bp in length, respectively, yet longer than that of *Bos taurus* (NC_001567), *Bos indicus* (NC_005971), and *Bubalus bubalis* (NC_006295), which are 16,338 bp, 16,339 bp, and 16,359 bp in length, respectively. The size differences are resulted from different lengths of the control region among these species. Nucleotide composition analysis revealed that the *P. hodgsonii* mitochondrial genome is biased towards AT (A 33.59%, T 26.87%, G 13.11%, and C 26.41%); such an AT content is lower than those of *O. aries* and *C. hircus*. The *P. hodgsonii* mitochondrial genome encodes 13 proteins, 2 rRNAs, and 22 tRNAs (Figure 1). Eight tRNA genes and one protein gene are located on the light strand (Table 1). And the number of polymorphic sites at protein coding genes and RNA genes in the mitochondrial genomes of *P. hodgsonii*, *C. hircus* and *O. aries* is shown in Table 2.

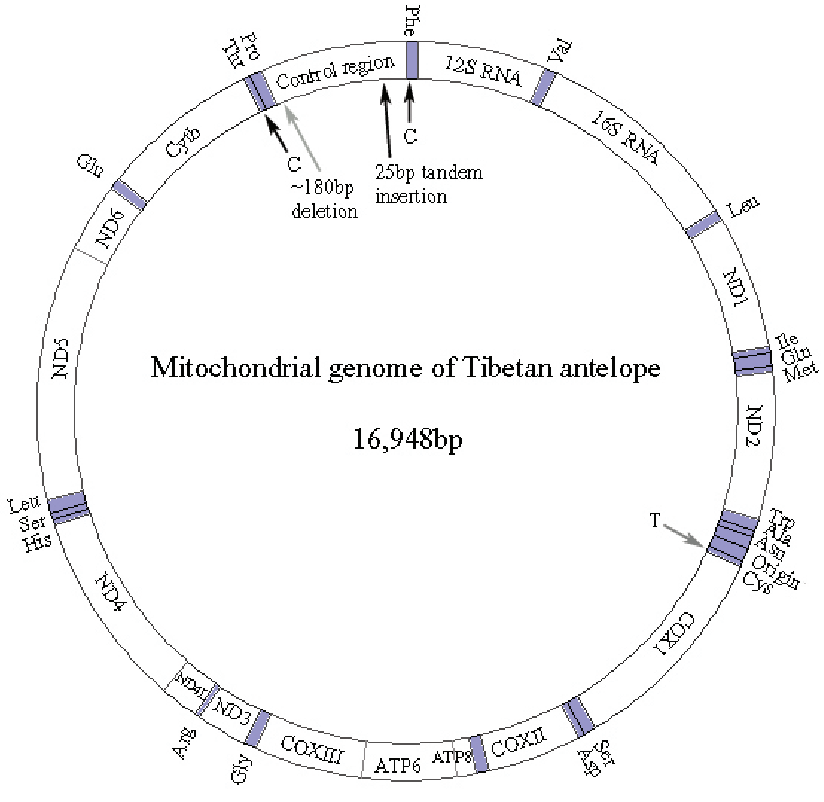

Fig. 1 Annotation of the *Pantholops hodgsonii* mitochondrial genome. The original figure was from Bae *et al* (*30*). ND1 to ND6 refer to NADH subunits; COXI, COXII, COXIII refer to cytochrome c oxidase subunits; ATP6 and ATP8 refer to ATPase6 and ATPase8; the tRNA genes are denoted by shade and are noted accordingly; Origin refers to Origin of L-strand replication. The ATPase8 and ATPase6 genes are overlapped for 40 bp. Some bases between genes are too few to be denoted in the figure.

Table 1 Components of *P. hodgsonii* Mitochondrial Genome

Gene	Direction	Nucleotide number	Start codon	Stop codon
tRNA Phe	F	1–69		
12S rRNA	F	70–1,026		
tRNA Val	F	1,027–1,093		
16S rRNA	F	1,094–2,659		
tRNA Leu	F	2,663–2,737		
NADH1	F	2,740–3,694	ATG	TAa
tRNA Ile	F	3,696–3,764		
tRNA Gln	R	3,762–3,833		
tRNA Met	F	3,836–3,904		
NADH2	F	3,905–4,946	ATA	Taa
tRNA Trp	F	4,947–5,013		
tRNA Ala	R	5,015–5,083		
tRNA Asn	R	5,085–5,157		
Origin of L-strand replication	R	5,158–5,189		
tRNA Cys	R	5,190–5,256		
tRNA Tyr	R	5,257–5,324		
COXI	F	5,326–6,870	ATG	TAA
tRNA Ser	R	6,868–6,936		
tRNA Asp	F	6,944–7,011		
COXII	F	7,013–7,696	ATG	TAA
tRNA Lys	F	7,700–7,767		
ATPase8	F	7,769–7,969	ATG	TAA
ATPase6	F	7,930–8,610	ATG	Taa
COXIII	F	8,610–9,393	ATG	Taa
tRNA Gly	F	9,394–9,462		
NADH3	F	9,463–9,809	ATA	TAa
tRNA Arg	F	9,810–9,878		
NADH4L	F	9,879–10,175	ATG	TAA
NADH4	F	10,169-11,546	ATG	Taa
tRNA His	F	11,547-11,616		
tRNA Ser	F	11,617-11,676		
tRNA Leu	F	11,678-11,747		
NADH5	F	11,748-13,568	ATA	TAA
NADH6	R	13,555-14,081	ATG	TAA
tRNA Glu	R	14,080-14,148		
Cytochrome b	F	14,153-15,292	ATG	AGA
tRNA Thr	F	15,296-15,366		
tRNA Pro	R	15,366-15,431		
Control region		15,432-16,498		

Protein coding genes

There are 13 protein-coding genes in the *P. hodgsonii* mitochondrial genome. Among these genes, eight use ATG as start codon and three (NADH2, NADH3, NADH5) use ATA as start codon. Some of these 13 protein genes are terminated with incomplete stop codons: NADH1, NADH3, and ATP6 are terminated with TA; COXIII (cytochrome c oxidase subunit III), NADH2, and NADH4 are terminated with T; the rest are terminated with TAA and AGA (Table 1). Presumably, these incomplete stop codons are accommodated post-transcriptionally in the mRNA maturation process, *i.e.* polyadenylation (*5*).

Table 2 Number of Polymorphic Sites at Protein Coding Genes and RNA Genes

Location	Gene	Number of mutations			
		All	*P. hodgsonii*	*C. hircus*	*O. aries*
1	tRNA-Phe	9	4	2	1
70	12s rRNA	101	39	29	31
1,027	tRNA-Val	10	5	3	2
1,094	16s rRNA	160	53	56	45
2,663	tRNA-Leu	4	4		
2,740	NU1M	149	53	48	42
3,762	tRNA-Gln (L)	1		1	
3,836	tRNA-Met	1		1	
3,905	NU2M	168	59	63	43
4,947	tRNA-Trp	3		3	
5,015	tRNA-Ala (L)	4	1	1	2
5,085	tRNA-Asn (L)	4		2	2
5,190	tRNA-Cys (L)	1			1
5,257	tRNA-Tyr (L)	4	2	1	1
5,326	COXI	241	100	65	71
6,868	tRNA-Ser (L)	4	1		3
6,944	tRNA-Asp	3	1	1	1
7,013	COXII	106	35	36	32
7,700	tRNA-Lys	12	4	3	4
7,769	ATP8	39	13	12	11
7,930	ATP6	131	44	45	39
8,610	COXIII	148	59	56	30
9,394	tRNA-Gly	4	1	2	
9,463	NU3M	57	25	13	17
9,810	tRNA-Arg	5	1	3	1
9,879	NU4L	45	10	10	23
10,169	NU4M	273	101	74	90
11,547	tRNA-His	8	3		4
11,617	tRNA-Ser	10	5	2	3
11,678	tRNA-Leu	2		2	
11,748	NU5M	365	126	110	121
13,555	NU6M (L)	69	19	25	24
14,080	tRNA-Glu (L)	6	3	2	1
14,153	CYB	185	60	62	57
15,296	tRNA-Thr	11	6	4	1
15,366	tRNA-Pro (L)	5		4	1

Number of polymorphic sites at protein coding genes and RNA genes in the mitochondrial genomes of *P. hodgsonii*, *C. hircus* and *O. aries*. The column "All" refers to the number of sites at which all three species are different; the column "*P. hodgsonii*" refers to number of the sites at which *P. hodgsonii* is different from the other two species.

We compared protein sequences between *P. hodgsonii* and other mammalian species. The *P. hodgsonii* CYTB, ND6, ND4L, and COXII bare higher homology to those of *C. hircus* than *O. aries* and other species; Its COXI, COXIII, NU1M, NU2M, NU3M, NU4M, NU5M, and NU6M are most similar to those of *O. aries* than *C. hircus* and other species. Interestingly, ATP8 of *P. hodgsonii* is 93.93% identical to that of *Bos grunniens* whereas it shares much identity with those of other species, including those of *O. aries* (89.39%) and *C. hircus* (83.33%). The nucleotide similarity of these *P. hodgsonii* proteins is in general higher when compared to *O. aries* and *C. hircus* than to other species. AT contents of the *P. hodgsonii* protein coding genes are higher than those of human and lower than those of mouse and rat (Data not shown).

We analyzed mutations of four protein coding genes in details, including CYTB, COXI, COXII, and COXIII, among *O. aries*, *C. hircus*, *B. grunniens*, *P. hodgsonii*, and a few other species of the Bovidae family, attempting to predict possible functional implications for mutations discovered from comparative analyses. In the case of COXI, there are five mutations affecting amino acids in *P. hodgsonii*, four in *B. grunniens*, three in *C. hircus*, and none in *O. aries*, as compared with *B. taurus* (Table 3). The amino acid changes in *P. hodgsonii* and *B. grunniens* occur mostly at sites between 400 and 500 (Table 3, Figure 2), whereas in *C. hircus* the mutations were found within 150 to 512. The region from 400 to 500 constitutes the transmembrane components X, XI, and XII.

Table 3 Amino Acid Mutations of the COXI Gene

Species	Mutation	E/B	Neighbor rigidity	Rigidity	Volume change	Charge change	Polarity change
C. hircus	V155I	64	−0.2351483	−0.34353	26.7	0	0
	A308T	5	−0.2854679	−0.4675	27.5	0	0
	N512S	80	2.12195984	2.659928	−114.1	0	−2
P. hodgsonii	F8Y	136	0.2337137	−0.0661	3.7	0	0
	D407T	76	0.63863865	0.801412	5	1	−1
	M449V	129	0.05870691	1.199486	−22.9	0	0
	L467V	34	0.18495316	0.683123	−26.7	0	0
	T509V	52	−0.2661298	−0.22177	23.9	0	−1
B. grunniens	Y440S	108	−0.3272085	0.142943	−193.6	0	0
	S441P	5	0.35454652	−0.54554	112.7	0	1
	I453V	51	0.1775126	−0.22684	−26.7	0	0
	F476L	96	0.08862233	0.404623	−23.2	0	0

Amino acid mutations of the COXI gene for *C. hircus*, *P. hodgsonii*, and *B. grunniens*, as compared with *B. taurus*. The influences of these mutations are predicted by the methods referred from Mirkovic *et al* (*31*).

```
B.taurus       -MFINRWLFS TNHKDIGTLY LLFGAWAGMV GTALSLLIRA ELGQPGTLLG DDQIYNVVVT  60
B.bubalis      -......... .......... .......... .......... .......... ..........  60
B.grunniens    -......... .......... .......... .......... .......... ..........  60
P.hodgsonii    -.......Y. .......... .......... .......... .......... ..........  60
S.scrofa       -..V....Y. .......... .......... .......... .......... .......I..  60
O.aries        -......... .......... .......... .......... .......... .......I..  60
M.muntjak      -......... .......... .......... .......... .......... .......I..  60
C.familiaris   -......... .......... .......... .......... .......... .......I..  60
E.caballus     -......... .......... .......... .......... .......... .......I..  60
M.musculus     -......... .......... .......... ....I..... .....A.... .......I..  60
C.hircus       -......... .......... .......... .......... .......... .......I..  60
H.sapiens      -..AD..... .......... ........VL .......... .....N.... N.H....I..  60
E.robustus     -..M...... .......... .......... ..G....... .....S.I.. .......L..  60
U.arctos       -..MS..... .......... .......... .......... .....A.... .......I..  60
U.maritimus    -..M...... .......... .......... .......... .....A.... ..........  60
X.laevis       -.A.T..... .......... .V......L. .......... ..S....... .......I..  60
S.crocodilurus -.T.T..F.. .......... .I.......A .......... ..S....... ..........  60
G.gallus       MT........ .......... .I..T....A .......... .......... .......I..  60
A.albifrons    MT........ .......... .I........ .......... .......... .......I..  60
B.rerio        -.T.T..F.. .......... .V........ .......... ..S..A.... .......I..  60
T.rubripes     -.A.T..F.. .......... .V........ .......... ..S..A.... .......I..  60
```

```
B.taurus       AHAFVMIFFM VMPIMIGGFG NWLVPLMIGA PDMAFPRMNN MSFWLLPPSF LLLLASSMVE  120
B.bubalis      .......... .......... .......... .......... .......... ..........  120
B.grunniens    .......... .......... .......... .......... .......... ..........  120
P.hodgsonii    .......... .......... .......... .......... .......... ..........  120
S.scrofa       .......... .......... .......... .......... .......... ..........  120
O.aries        .......... .......... .......... .......... .......... ..........  120
M.muntjak      .......... .......... .......... .......... .......... ..........  120
C.familiaris   .......... .......... .......... .......... .......... ..........  120
E.caballus     .......... .......... .......... .......... .......... ........I.  120
M.musculus     .......... ...M...... .......... .......... .......... ..........  120
C.hircus       .......... .......... .......... .......... .......... ..........  120
H.sapiens      .......... .......... .......... .......... .........L ......A...  120
E.robustus     .......... .......... .......... .......... .......... ...M......  120
U.arctos       .......... .......... .......... .......... .......... ..........  120
U.maritimus    .......... .......... .......... .......... .......... ..........  120
X.laevis       ....I..... .......... .......... .......... .......... .......G..  120
S.crocodilurus .......... ...V...... .......... .......... .........L .......GI.  120
G.gallus       .......... .......... .......... .......... .......... .......T..  120
A.albifrons    .......... .......... .......... .......... .......... .......T..  120
B.rerio        .......... ....L..... .......... .......... .......... .......G..  120
T.rubripes     .......... .......... ...I...... .......... .......... .......G..  120

B.taurus       AGAGTGWTVY PPLAGNLAHA GASVDLTIFS LHLAGVSSIL GAINFITTII NMKPPAMSQY  180
B.bubalis      .......... .......... .......... .......... .......... ..........  180
B.grunniens    .......... .......... .......... .......... .......... ..........  180
P.hodgsonii    .......... .......... .......... .......... .......... ..........  180
S.scrofa       .......... .......... .......... .......... .......... ..........  180
O.aries        .......... .......... .......... .......... .......... ..........  180
M.muntjak      .......... .......... .......... .......... .......... ..........  180
C.familiaris   .......... .......... .......... .......... .......... ..........  180
E.caballus     .......... .......... .......... .......... .......... ......L...  180
M.musculus     .......... .......... .......... .......... .......... .......T..  180
C.hircus       .......... .......... .......... .....I.... .......... ..........  180
H.sapiens      .......... ......YS.P .......... .......... .......... .......T..  180
E.robustus     .......... .......... .......... .......... .......... .......T..  180
U.arctos       .......... .......... .......... .....I.... .........V ..........  180
U.maritimus    .......... .......... .......... .....I.... .......... ..........  180
X.laevis       .......... .......... .......... .....I.... ........T. ..........  180
S.crocodilurus .......... .......... .......... .......... ........C. ....A..T..  180
G.gallus       .......... .......... ......A..– HY........ .......... ......L...  180
A.albifrons    .......... .......... ......A... .....I.... ........A. ......L...  180
B.rerio        .......... .......... .......... .......... ........T. .....TI...  180
T.rubripes     .......... .......... .......... .......... .......... ......I...  180

B.taurus       QTPLFVWSVM ITAVLLLLSL PVLAAGITML LTDRNLNTTF FDPAGGGDPI LYQHLFWFFG  240
B.bubalis      .......... .......... .......... .......... .......... ..........  240
B.grunniens    .......... .......... .......... .......... .......... ..........  240
P.hodgsonii    .........L .......... .......... .......... .......... ..........  240
S.scrofa       .........L .......... .......... .......... .......... ..........  240
O.aries        .........L .......... .......... .......... .......... ..........  240
M.muntjak      .........L .......... .......... .......... .......... ..........  240
C.familiaris   .........L .......... .......... .......... .......... ..........  240
E.caballus     .........L ........A. .......... .......... .......... ..........  240
M.musculus     .........L .......... .......... .......... .......... ..........  240
C.hircus       .........L .......... .......... .......... .......... ..........  240
H.sapiens      .........L .......... .......... .......... .......... ..........  240
E.robustus     .........L .......... .......... .......... .......... ..........  240
U.arctos       .........L .......... .......... .......... .......... ..........  240
U.maritimus    .........L .......... .......... .......... .......... ..........  240
X.laevis       .........L .......... .......... .......... .........V ..........  240
S.crocodilurus .........L .......... .......... ........S. .........V ..........  240
G.gallus       .........L ...I...... .......... .......... .......... ..........  240
A.albifrons    .........L ...I...... .......... .......... .......... ..........  240
B.rerio        .......A.L V......... .......... .......... .......... ..........  240
T.rubripes     .......A.L .......... .......... .......... .......... ..........  240
```

B.taurus	HPEVYILILP	GFGMISHIVT	YYSGKKEPFG	YMGMVWAMMS	IGFLGFIVWA	HHMFTVGMDV	300
B.bubalis	....	....	....	....	....	....	300
B.grunniens	....	....	....	....	....	....	300
P.hodgsonii	....	....	....	....	....	....	300
S.scrofa	....	....	....	....	....	....	300
O.aries	....	....	....	....	....	....	300
M.muntjak	....	....	....	....	....	....	300
C.familiaris	....	....	....	....	....	....	300
E.caballus	....	....	....	....	....	....	300
M.musculus	....	...I...V..	....	....	....	.......L..	300
C.hircus	....	....	....	....	....	....	300
H.sapiens	....	....	....	....	....	....	300
E.robustus	....	....	....	........V.	....	....	300
U.arctos	....	....	....	....	....	....	300
U.maritimus	....	....	....	....	....	....	300
X.laevis	....	....	....	....	..L.......	......DLN.	300
S.crocodilurus	....	.......V..	..A.......	....	....	....	300
G.gallus	....	.......V.A	..A.......	........L.	....	......R...	300
A.albifrons	....	...I...V..	....	........L.	....	....	300
B.rerio	....	...I...V.A	..A.......	.........A	..L.......	....	300
T.rubripes	....	.........A	....	.........A	..L.......	....	300

B.taurus	DTRAYFTSAT	MIIAIPTGVK	VFSWLATLHG	GNIKWSPAMM	WALGFIFLFT	VGGLTGIVLA	360
B.bubalis	....	....	....	....	....	....	360
B.grunniens	....	....	....	....	....	....	360
P.hodgsonii	....	....	....	....	....	....	360
S.scrofa	....	....	....	L	....	....	360
O.aries	....	....	....	....	....	....	360
M.muntjak	....	....	....	....	....	....	360
C.familiaris	....	....	....	L	....	....	360
E.caballus	....	....	....	L	....	....	360
M.musculus	....	....	....	L	....	.........S	360
C.hircus	........T.	....	....	....	....	....	360
H.sapiens	....	....	....	S.M...A.VL	....	....	360
E.robustus	....	....	....	........L.	....	....	360
U.arctos	....	....	....	....	....	....	360
U.maritimus	....	....	....	....	....	....	360
X.laevis	....	....	.......M..	.T...DAP.L	....	....	360
S.crocodilurus	....	....	....	.I...DA.LL	....	....	360
G.gallus	....	........I.	....	.T...D.P.L	....	I.........	360
A.albifrons	....	........I.	....	.T...D.P.L	....	I.........	360
B.rerio	....	....	....	.A...ETP.L	....	....	360
T.rubripes	....	....	....	.S...ETP.L	....	....	360

B.taurus	NSSLDIVLHD	TYYVVAHFHY	VLSMGAVFAI	MGGFVHWFPL	FSGYTLN**D**TW	AKIHFAIMFV	420
B.bubalis	....	....	....	....	....	....	420
B.grunniens	....	....	....	....	....	....	420
P.hodgsonii	....	....	....	....	.......**T**..	....	420
S.scrofa	....	....	....	....	.......**Q**A.	.....V....	420
O.aries	....	....	....	....	....	....	420
M.muntjak	....	....	....	....	.......**N**..	.....V....	420
C.familiaris	....	....	....	....A.....	....	.....T....	420
E.caballus	....	....	....	....	.......**Q**..	.....T....	420
M.musculus	....	....	....	.A........	...F..D...	..A.......	420
C.hircus	....	....	....	....	....	....	420
H.sapiens	....	....	....	....I.....	......D**Q**.Y	.....T...I	420
E.robustus	....	....	....	....	.......**T**..	T....M....	420
U.arctos	....	....	....	....A.....	.......**N**..	.....M...I	420
U.maritimus	....	....	....	....	.......**N**..	.....M...I	420
X.laevis	......M...	....	....	....I.....	.T...H**E**..	.....GV..A	420
S.crocodilurus	....	....	....	.A........	...F..H**P**A.	T....GT..I	420
G.gallus	......A...	....	....	LA..T.....	.T.F..H**P**S.	T.A..GV..T	420
A.albifrons	......A...	....	....	LA..T.....	LT.F..H**Q**..	..A..GV..T	420
B.rerio	....	....	....	.A........	.T.....**S**V.	T....GV..I	420
T.rubripes	....	....	....	..A.......	......H**S**..	T....GV..I	420

```
B.taurus       GVNMTFFPQH FLGLSGMPRR YSDYPDAYTM WNTISSMGSF ISLTAVMLMV FIIWEAFASK 480
B.bubalis      .......... .......... .......... .......... .......... .......... 480
B.grunniens    .......... .......... SP........ ...V...... .......... ......L... 480
P.hodgsonii    .......... .......... .........V .......... .......V.I .......... 480
S.scrofa       .......... .......... .........A .......... .........I .......... 480
O.aries        .......... .......... .......... .......... .........I .......... 480
M.muntjak      .......... .......... .......... .......... .........I .......... 480
C.familiaris   .......... .......... .........T ...V...... .........I .M........ 480
E.caballus     .......... .......... .........T .......... .........I .M........ 480
M.musculus     .......... .......... .........T ...V...... ......LI.I .M........ 480
C.hircus       .......... .......... .......... .......... .........I .......... 480
H.sapiens      ...L...... .......... .........T ..IL..V... .........I .M........ 480
E.robustus     ...L...... .......... .........T .......... .........I .......T.. 480
U.arctos       .......... .......... .........T ...V...... .........I .M........ 480
U.maritimus    .......... .......... .........T ...V...... .........I .M........ 480
X.laevis       ...L...... ....A..... .........L ...V..I..L ...V..IM.M ........A. 480
S.crocodilurus .......... ...A...... .........I ..A...I..L ...VG.VM.M .......ST. 480
G.gallus       ...L...... ...A...... .........L ...L..I..L ..M...IMLM ..V....SA. 480
A.albifrons    ...L...... ...A...... .........L ......I..L ..MV..IMLM .......SA. 480
B.rerio        ...L...... ...A...... ........AL ...V..I..L ...V..IMFL ..L....TA. 480
T.rubripes     ...L...... ...A...... ........AL ..SV..I..M V..V..IMFL ..L....TA. 480

B.taurus       REVLTVDLTT TNLEWLNGCP PPYHTFEEP- TYVNLK----  520
B.bubalis      ...S...... .......... .........- ......----  520
B.grunniens    .......... .......... .........- ......----  520
P.hodgsonii    .......... .......... .........- V.....----  520
S.scrofa       ...SA.E..S ......H... .........- ..I...----  520
O.aries        .......... .......... .........- ......----  520
M.muntjak      ......E... .......... .........- ......----  520
C.familiaris   ...AM.E... ..I...H... .........- ...IQ.----  520
E.caballus     ...S..E..S ......H... .........- A.....----  520
M.musculus     ...MS.SYAS ......H... .........- ...KV.----  520
C.hircus       .......... .......... .........- ...S..----  520
H.sapiens      .K..M.EEPS M.....Y... .........- V.MKS-----  520
E.robustus     ....A....S .......... .........- AF..P.WS--  520
U.arctos       ...AV.E..S ..I...H... .........- A..T..----  520
U.maritimus    ...AV.E..S ..I...H... .........- A..T..----  520
X.laevis       ...T.YE..S .M....Q... T....LKTSL VQI.HQMIKS  520
S.crocodilurus ....SLEA.N ......H... ..F..Y...- .F.QTSRE--  520
G.gallus       .K..QPE..A ..I..IH... .........- AF.QVQE---  520
A.albifrons    .K..QPE... ..V..IH... .........- AF.QVQE---  520
B.rerio        ....S.E..A ..V...H... .........- AF.QIQSN--  520
T.rubripes     ...QS.E..M ..V...H... .........- AF.QTQT---  520
```

Fig. 2 Alignment of amino acid sequences of the COXI gene using CLUSTAL W1.83. The species included were *Bos taurus* (P00396), *Bubalus bubalis* (YP_087084), *Bos grunniens* (YP_133786), *Pantholops hodgsonii*, *Sus scrofa* (NP_008636), *Ovis aries*(NP_008408), *Muntiacus muntjak* (AY225986), *Canis familiaris* (NP_008473), *Equus caballus* (NP_007162), *Mus musculus* (NP_904330), *Capra hircus* (NP_877405), *Homo sapiens* (NP_536845), *Eschrichtius robustus* (NP_944635), *Ursus arctos* (Q8SJI5), *Ursus maritimus* (Q8SJH3), *Xenopus laevis* (P00398), *Shinisaurus crocodilurus* (NC_005959), *Gallus gallus* (P18943), *Anser albifrons* (NP_777304), *Brachydanio rerio* (Q9MIY8), and *Takifugu rubripes* (NP_694917). Mutations of *P. hodgsonii* were highlighted as bold letters.

RNA genes

There are 22 tRNA genes identified in the *P. hodgsonii* mitochondrial genome, typical for mammalian mitochondrial genomes (*6–9*). Lengths of these tRNAs range from 65 to 74 bp. Some indels occur in the dihydrouridine and TψC arms (Figure 1). The 12S rRNA and 16S rRNA genes are 957 bp and 1,566 bp in length, respectively.

Control region

Mitochondrial control region of *P. hodgsonii* is 1,067 bp in length, shorter than that of *O. aries* (1,180 bp), *C. hircus* (1,212 bp), domestic dog (1,270 bp), and domestic horse (1,192 bp), but longer than that of *B. taurus* (910 bp), *B. indicus* (913 bp), *B. bubalis* (910 bp), and *B. grunniens* (894 bp). By using the Tandem Repeats Finder (*10*), we found a 75-bp tandem repeats that vary among different species: four in *O. aries* (Ref. *6*; another *O. aries* haplotype in the same study has three repeats), two in *C. hircus*, and two in *P. hodgsonii*. The 75-bp repeats appear at the same location for these three species, close to the 5′-end of the last prolinyl tRNA (Pro) gene (Figure 1). Sequence consensuses of these 75-bp repeats are very similar with identities of 86.7% between *O. aries* and *P. hodgsonii*, 74.0% between *C. hircus* and *P. hodgsonii*, and 76.6% between *C. hircus* and *O. aries*. In addition, *P. hodgsonii* has two additional 25-bp tandem repeats, which have not been found in *Bos* species. This repeat is inserted at the 3′-end of the *P. hodgsonii* control region, close to phenyalaninyl tRNA (Phe) gene (Figure 1).

The 75-bp repeat has a lower GC content (25%) than the average of the control region (39%), whereas the 25-bp repeat has slightly higher GC content (41%) than the average. Although we only found two 75-bp-long tandems in control region of the *C. hircus* mitochondrial genome, there is no loss of length in the *C. hircus* control region, compared to the same region of *O. aries*.

Phylogenetic analysis and time of divergence

Complete cytochrome b of *P. hodgsonii*, *O. aries*, *C. hircus*, and other species of the Bovidae family were used to construct phylogenetic trees by using PAUP*(4b10) (*11*), with *Giraffe camelopardalis* as the outgroup. Our result indicates that *P. hodgsonii* is more related to *O. aries*, *C. hircus*, and *Oreotragus oreotragus* (klipspringer) (Figure 3), consistent with previously studies (*12, 13*). Based on the phylogenetic tree and a substitution rate of 0.056 per site per million years between two taxa (*14*), we estimated that *P. hodgsonii* and *O. aries* divided about 2.25 million years ago (Table 4).

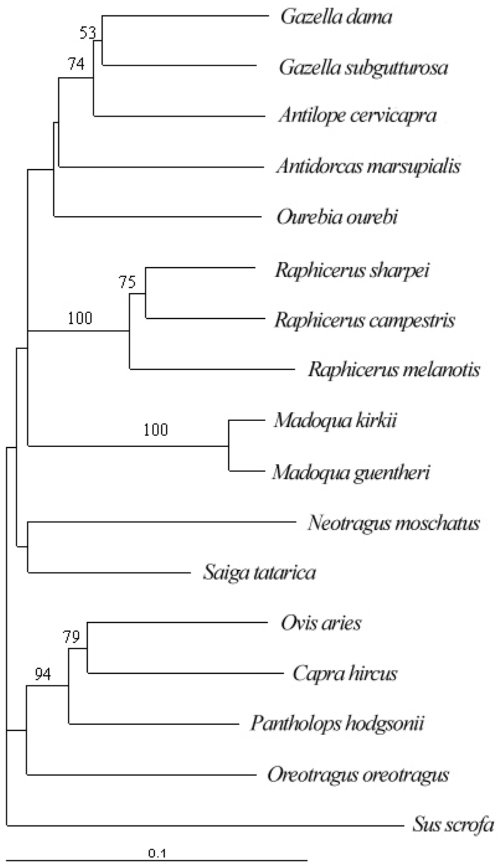

Fig. 3 Neighbor-joining tree based on the complete cytochrome b genes of some Bovidae species. The tree was constructed by using PAUP*(4b10), with parameters estimated by MODELTEST3.6 and with *Giraffe camelopardalis* as the outgroup. Number of bootstrap replication is 1000. Bootstrap values over 50% are shown on the branches.

Table 4 Molecular Divergence and Estimated Divergence Time Between Native Mammals on the Tibetan Plateau and Their Lowland Relatives

Species	Molecular divergence	Divergence time (million years before presence)
P. hodgsonii–C. hircus	0.124	2.22
P. hodgsonii–O. aries	0.126	2.25
O. aries–C. hircus	0.124	2.21
B. grunniens–B. taurus	0.089	1.59

Based on complete sequences of mitochondrial cytochrome b and a substitution rate of 0.056 per site per million years between two mammalian taxa.

Positive selection analysis

We used PAML (*15*) to analyze natural selection on mitochondrial genes of *P. hodgsonii*. The numbers of N (nonsynonymous substitutions) and S (synonymous substitutions) of each branch were calculated for eight genes that are more than 1,000 bp in length by *codmel* in PAML package. We found that N/S ratios of the COXI gene are significantly higher for the *P. hodgsonii* branch and the *B. grunniens* branch, compared with other branches (Fisher test $P<0.001$; Figure 4). The difference between N/S ratios of the *B. grunniens* branch and the *P. hodgsonii* branch is not significant (Fisher test $P>0.1$). These results suggested that the COXI gene has probably undergone positive selection in *P. hodgsonii* and *B. grunniens*. Despite the fact that COXI of *P. hodgsonii* has only three unique amino acid changes, it has other variation shared with one or more other species, which may collectively alter the function of the protein itself or interactions with other components in the mitochondrial respiration system. Experimentation and detailed structural analysis are of essence to pinpoint the relatedness of structure and function underscored by genetic changes. The sequence alignment of a highly variable region of COXI from selected species is shown in Figure 2.

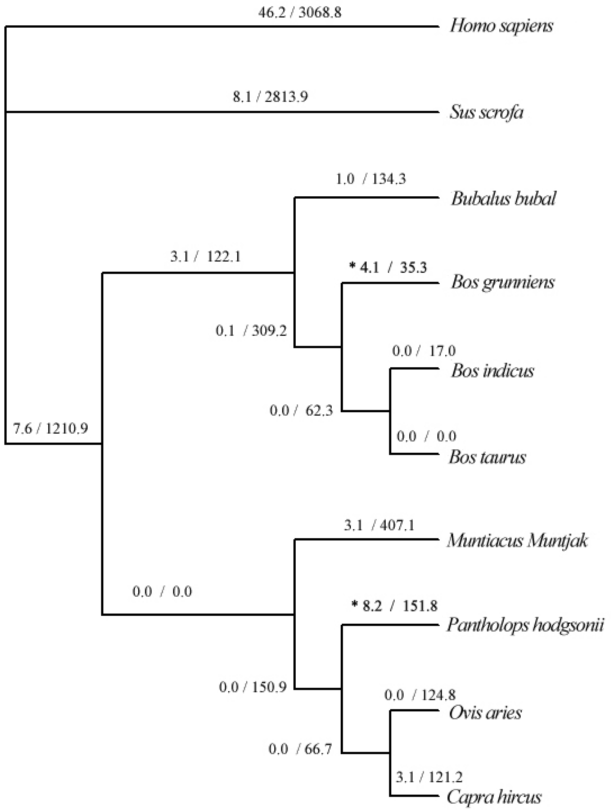

Fig. 4 The COXI gene was used to analyze N/S ratio of each branch within Artiodactyla mammals. Phylogeny was inferred with PAUP*(4b10). Numbers on the branches are mutation of nonsynonymous/synonymous for protein coding region of the COXI gene, as obtained with codeml in PAML package. * Fisher test P-value $<$ 0.001. Take other branches (except *B. grunniens* and *P. hodgsonii* branches) as background for testing. The difference between the *B. grunniens* and *P. hodgsonii* branches is not significant (Fisher test P-value $>$ 0.1).

Discussion

In the current study, we sequenced a complete mitochondrial genome from one *P. hodgsonii* individual, as the first attempt in a long-term research effort to understand the genetic basis of hypoxic adaptations of native fauna and flora on the Tibetan Plateau. Similar to other mammalian mitochondrial genomes, the *P. hodgsonii* mitochondrial genome contains 13 protein genes, 2 rRNA genes, 22 tRNA genes, and one control region. Its control region has two 75-bp tandem repeats near the last tRNA gene, whereas *O. aries*, *C. hircus*, and *Bos* species have three/four, two, and none, respectively. Although we only identified two 75-bp repeats in the control region of *C. hircus*, it is still possible that there might have been four units since *C. hircus* mitochondrial genome is very similar to the *O. aries* sequence in length and two of these four repeat units may occur early in time, resulted in poor homology due to mutations over a long decaying period. There are two additional 25-bp tandem repeats near the end of *C. hircus*'s control region, suggesting that short tandem repeats occur frequently among the mitochondrial genomes of *O. aries*, *C. hircus*, and *P. hodgsonii*.

Phylogenetic analysis on cytochrome b genes revealed that *P. hodgsonii* is more closely related with *O. aries*, *C. hircus*, and *O. oreotragus*, rather than other antelope species (the Antilopinae subfamily). Based on a substitution rate of 0.056 per site per million years between two mammalian taxa, we estimated that *P. hodgsonii* and *O. aries* divided about 2.25 million years ago. It was reported previously that climates around the Tibetan Plateau have undergone rapid changes at about 3.6–2.6 million years ago (*16*, *17*). Our results are in close agreement with the proposed climatic changes around the Tibetan Plateau.

To investigate whether the *P. hodgsonii* mitochondrial genes have been selected under the hypoxic environment of the plateau, we compared functional genes (CYTB, COXI, COXII, and COXIII) encoded in the mitochondrial genomes among *P. hodgsonii* and other distant relatives, including *O. aries*, *C. hircus*, and *B. grunniens*. In the COXI gene, we found that most of the mutations affecting amino acid sequences in *P. hodgsonii* and *B. grunniens* occur at sites between 400 and 500, whereas mutations in *C. hircus* occur at sites between 120 and 512. The region of 400 to 500 contains the transmembrane components X, XI, and XII, suggesting that nonsynonymous mutations unique to the two high altitude-adapted species may have functional implications. Although this process of functional prediction may not be very reliable, our results suggested to the likely targets for the next experimentation to verify the novel conjecture.

Evidence from N/S ratios shows that the COXI gene has more functional mutations in *P. hodgsonii* and *B. grunniens* compared with other mammals (Figure 4), providing further evidence that the COXI gene might have undergone positive selection among the native mammalian species of the plateau. The cytochrome c oxidase is the last step of the electron transport chain. It is consisted of 13 subunits, of which three subunits are encoded by the mitochondrial genome (COXI, II, and III), and ten subunits are encoded by the nuclear genome. Functional core of the enzyme complex is composed of subunits 1, 2, and 3 (*18*). Numerous studies have shown that some of the subunits of the COX gene have higher nonsynonymous substitution rate in primate than in other animals, such as COXI (*19*), COXII (*20*), and the nuclear coded COXIV (*21*) and COXVII (*22*). Several pieces of evidence suggest that structure and activity of cytochrome c oxidase may have adaptive changes during physiological hypoxia in mouse and rat cells (*23*). Expression of the mitochondrial genome encoded subunit COXI can decrease due to hypoxia while the enzyme efficiency remained. It suggested that expression of the COXI gene was regulated by the oxygen content. In addition, significant higher expression of COXI mRNA was observed in mammalian tissues, such as kidney and heart, during hibernation, and the change was not found in euthermic animals (*24*, *25*).

Since *P. hodgsonii* and *B. grunniens* are both well-adapted to the same environment, the Tibetan Plateau, natural selection may have resulted in similar genetic signatures in their genomes, including the nuclear and mitochondrial genomes. Similar seemingly function-associated mutations and dN/dS ratios at COXI of *P. hodgsonii* and *B. grunniens* provided useful clues for further studies and functional confirmations on the role of mtDNA-encoded COX subunits on adaptation of native mammals to the unique Tibetan Plateau.

Materials and Methods

DNA extraction and sequencing

Blood samples of *P. hodgsonii* individuals were collected from the Kekexili Natural Reservation in Qinghai Province, China, in December 2004. Samples were stored at 4℃ for a few days before whole genomic DNA was extracted from the whole blood with standard salt-extraction method.

PCR primers and sequencing primers were designed based on a sequence of *O. aries* mitochondrial genome (NC_001941; ref. *6*). PCR reactions were conducted on a PTC-200 thermal cycler with the following conditions: an initial denaturation step of 95℃ (3 min) followed by 34 cycles of 95℃ (30 s), 58℃/56℃ (30 s), and 72℃ (90 s) followed by 72℃ for 10 min. PCR products were purified by Montage PCR Cleanup Kit (Millipore, Billerica, USA) and sequenced with ABI 3730xl DNA Analyzer (Applied Biosystems, Foster City, USA).

Sequence analysis

Base calling was performed with Phred (*26*) at the lowest Phred quality values control of 20. The sequences were assembled with Phrap (http://www.phrap.org) in a default setting. Sequence contigs were further finished, compared, and annotated in referencing to that of the *O. aries* sequence (NC_001941). Genes of tRNA were defined with tRNAscan-SE 1.2 (http://www.genetics.wustl.edu/eddy/tRNAscan-SE/). Comparative analysis was performed by using BLAST (*27*) and CLUSTAL W1.83 (*28*). Tandem repeats were defined with the program Tandem Repeats Finder (*10*).

Phylogenetic analysis

Complete sequences of cytochrome b genes from *P. hodgsonii*, *O. aries*, *C. hircus*, and other selected species of the Bovidae family were used to construct phylogenetic trees, with *G. camelopardalis* as an outgroup. The sequences used were from *P. hodgsonii*, *O. oreotragus* (AF022052), *Antilope cervicapra* (AF022058), *Ourebia ourebi* (AF320574), *Neotragus moschatu* (AF022069), *Madoqua kirki* (AF022070), *Madoqua guentheri* (AF022071), *Saiga tatarica* (AF064487), *Raphicerus sharpei* (AF022050), *Gazella dama* (AF025954), *Gazella subgutturosa* (AF036282), *Raphicerus campestris* (AF022068), *Raphicerus melanotis* (AF022053), *Antidorcas marsupialis* (AF022054), *O. aries* (NC_001941), *C. hircus* (NC_005044), *Bubalus bubal* (NC_006295), *B. grunniens* (NC_006380), *B. indicus* (NC_005971), *B. taurus* (NC_006853), and *G. camelopardalis* (AB001612).

Sequences were aligned with CLUSTAL W1.83 in default options. Evolutionary models and parameters of the aligned sequences were estimated by MODELTEST3.6 (*29*). Phylogenetic trees were constructed with the neighbor joining arithmetic method using PAUP*(4b10). Bootstrap analysis of 1,000 replicates was performed to estimate robustness of the tree.

Analysis of positive selection

Eight protein coding genes with the length of over 1,000 bp were used to calculate N/S in each branch by using PAML, and phylogenetic trees were constructed with PAUP*(4b10). The F3X4 model for codon frequency and free ratio model of ωratio assumption was performed with the genetic codon of mammalian mitochondrial, and the other parameters were used as default. The DNA sequences were aligned as protein sequences. The following sequences were also included in the analysis from *Homo sapiens* (X93334), *Sus scrofa* (NC_000845), *Bubalus bubal* (NC_006295), *B. grunniens* (NC_006380), *B. indicus* (NC_005971), *B. taurus* (NC_006853), *Muntiacus muntjak* (AY225986), *O. aries* (NC_001941), *C. hircus* (NC_005044), and *P. hodgsonii.*

Acknowledgements

We are grateful to Dr. Ming-Sheng Yu of the Qinghai Academy of Animal and Veterinary Science, Qinghai University, staff of the Kekexili Natural Reservation, and students of Qinghai University Medical School for sample collection. Xiao-Mei Tong, Yan Guo, Jun Li, and staff of the Beijing Genomics Institute helped through molecular experiment and data analysis. Dr. Wen Wang provided invaluable comments for data analysis and the manuscript. This work was supported by Chinese Academy of Sciences (grant to Jun Yu, No. KSCX2-SW-331), National Natural Science Foundation of China (grant to Ri-Li Ge, No. 30393133), and Natural Research Foundation of Qinghai (grant to Ri-Li Ge, No. 2003-N-120).

References

1. Nowak, R.M. (ed.) 1999. *Walkers Mammals of the World*, Sixth Edition. The Johns Hopkins University Press, Baltimore, USA.
2. Boore, J.L. 1999. Animal mitochondrial genomes. *Nucleic Acids Res.* 27: 1767-1780.
3. Blier, P.U., *et al.* 2001. Natural selection and the evolution of mtDNA-encoded peptides: evidence for intergenomic co-adaptation. *Trends Genet.* 17: 400-406.
4. Bannikova, A.A. 2004. Molecular markers and modern phylogenetics of mammals. *Zh. Obshch. Biol.* 65: 278-305.
5. Ojala, D., *et al.* 1981. tRNA punctuation model of RNA processing in human mitochondria. *Nature* 290: 470-474.
6. Hiendleder, S., *et al.* 1998. The complete mitochondrial DNA sequence of the domestic sheep (*Ovis aries*) and comparison with the other major ovine haplotype. *J. Mol. Evol.* 47: 441-448.
7. Anderson, S., *et al.* 1981. Sequence and organization of the human mitochondrial genome. *Nature* 290: 457-465.
8. Lin, C.S., *et al.* 1999. Complete nucleotide sequence of pig (*Sus scrofa*) mitochondrial genome and dating evolutionary divergence within Artiodactyla. *Gene.* 236: 107-114.
9. Parma, P., *et al.* 2003. The complete nucleotide sequence of goat (*Capra hircus*) mitochondrial genome. Goat mitochondrial genome. *DNA Seq.* 14: 199-203.
10. Benson, G. 1999. Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res.* 27: 573-580.
11. Swofford, D.L. 2002. PAUP*: phylogenetic analysis using parsimony (and other methods). Version 4 beta 10. CD-ROM. Sinauer Associates, Inc., Publishers, Sunderland, USA.
12. Hassanin, A., *et al.* 1998. Molecular systematics of the subfamily Caprinae (Artiodactyla, Bovidae) as determined from cytochrome b sequences. *J. Mammal. Evol.* 5: 217-236.
13. Matthee, C.A. and Robinson, T.J. 1999. Cytochrome b phylogeny of the family Bovidae: resolution within the Alcelaphini, Antilopini, Neotragini, and Tragelaphini. *Mol. Phylogenet. Evol.* 12: 31-46.

14. Zheng, X., *et al.* 2003. Historical demography and genetic structure of sister species: deermice (*Peromyscus*) in the North American temperate rain forest. *Mol. Ecol.* 12: 711-724.
15. Yang, Z. 1997. PAML: a program package for phylogenetic analysis by maximum likelihood. *Comput. Appl. Biosci.* 13: 555-556.
16. An, Z., *et al.* 2001. Evolution of Asian monsoons and phased uplift of the Himalaya-Tibetan plateau since Late Miocene times. *Nature* 411: 62-66.
17. Li, J.J. *et al.* 2001. Late Cenozoic intensive uplift of Qinghai-Xizang Plateau and its impacts on environments in surrounding area. *Quarter. Sci.* 21: 381-391.
18. Cooper, C.E., *et al.* 1991. Cytochrome c oxidase: structure, function, and membrane topology of the polypeptide subunits. *Biochem. Cell. Biol.* 69: 586-607.
19. Wu, W., *et al.* 2000. Molecular evolution of cytochrome c oxidase subunit I in primates: is there coevolution between mitochondrial and nuclear genomes? *Mol. Phylogenet. Evol.* 17: 294-304.
20. Adkins, R.M. and Honeycutt, R.L. 1994. Evolution of the primate cytochrome c oxidase subunit II gene. *J. Mol. Evol.* 38: 215-231.
21. Wu, W., *et al.* 1997. Molecular evolution of cytochrome c oxidase subunit IV: evidence for positive selection in simian primates. *J. Mol. Evol.* 44: 477-491.
22. Schmidt, T.R., *et al.* 1999. Molecular evolution of the COX7A gene family in primates. *Mol. Biol. Evol.* 16: 619-626.
23. Vijayasarathy, C., *et al.* 2003. Adaptive changes in the expression of nuclear and mitochondrial encoded subunits of cytochrome c oxidase and the catalytic activity during hypoxia. *Eur. J. Biochem.* 270: 871-879.
24. Hittel, D.S. and Storey, K.B. 2002. Differential expression of mitochondria-encoded genes in a hibernating mammal. *J. Exp. Biol.* 205: 1625-1631.
25. Carey, H.V., *et al.* 2003. Mammalian hibernation: cellular and molecular responses to depressed metabolism and low temperature. *Physiol. Rev.* 83: 1153-1181.
26. Ewing, B. and Green, P. 1998. Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res.* 8: 186-194.
27. Altschul, S.F., *et al.* 1990. Basic local alignment search tool. *J. Mol. Biol.* 215: 403-410.
28. Thompson, J.D., *et al.* 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22: 4673-4680.
29. Posada, D. and Crandall, K.A. 1998. MODELTEST: testing the model of DNA substitution. *Bioinformatics* 14: 817-818.
30. Bae, J.S., *et al.* 2004. The mitochondrial genome of the firefly, *Pyrocoelia rufa*: complete DNA sequence, genome organization, and phylogenetic analysis with other insects. *Mol. Phylogenet. Evol.* 32: 978-985.
31. Mirkovic, N., *et al.* 2004. Structure-based assessment of missense mutations in human BRCA1: implications for breast and ovarian cancer predisposition. *Cancer Res.* 64: 3790-3797.