SCIENCE CHINA
Life Sciences

• RESEARCH PAPER •

May 2013 Vol.56 No.5: 440–452
doi: 10.1007/s11427-013-4474-4

Mitochondrial genome sequences of *Artemia tibetiana* and *Artemia urmiana*: assessing molecular changes for high plateau adaptation

ZHANG HangXiao[1,2], LUO QiBin[3], SUN Jing[1,2], LIU Fei[1,2], WU Gang[1,2], YU Jun[1*] & WANG WeiWei[1,4*]

[1]*CAS Key Laboratory of Genome Sciences and Information, Beijing Institute of Genomics, Chinese Academy of Sciences, Beijing 101300, China;*
[2]*University of Chinese Academy of Sciences, Beijing 100094, China;*
[3]*Department of Genome Oriented Bioinformatics, Technische Universität München, Wissenschaftszentrum Weihenstephan, Freising 85350, Germany*
[4]*Division of Gastroenterology, Department of Medicine, University of Alberta, Edmonton T6G 2V2, Alberta, Canada*

Received November 29, 2012; accepted March 19, 2013

Brine shrimps, *Artemia* (Crustacea, Anostraca), inhabit hypersaline environments and have a broad geographical distribution from sea level to high plateaus. *Artemia* therefore possess significant genetic diversity, which gives them their outstanding adaptability. To understand this remarkable plasticity, we sequenced the mitochondrial genomes of two *Artemia tibetiana* isolates from the Tibetan Plateau in China and one *Artemia urmiana* isolate from Lake Urmia in Iran and compared them with the genome of a low-altitude *Artemia*, *A. franciscana*. We compared the ratio of the rate of nonsynonymous (*Ka*) and synonymous (*Ks*) substitutions (*Ka*/*Ks* ratio) in the mitochondrial protein-coding gene sequences and found that *atp8* had the highest *Ka*/*Ks* ratios in comparisons of *A. franciscana* with either *A. tibetiana* or *A. urmiana* and that *atp6* had the highest *Ka*/*Ks* ratio between *A. tibetiana* and *A. urmiana*. *Atp6* may have experienced strong selective pressure for high-altitude adaptation because although *A. tibetiana* and *A. urmiana* are closely related they live at different altitudes. We identified two extended termination-associated sequences and three conserved sequence blocks in the D-loop region of the mitochondrial genomes. We propose that sequence variations in the D-loop region and in the subunits of the respiratory chain complexes independently or collectively contribute to the adaptation of *Artemia* to different altitudes.

mitochondrial genome, *Artemia tibetiana*, sequence variation, high plateau species

Citation: Zhang H X, Luo Q B, Sun J, et al. Mitochondrial genome sequences of *Artemia tibetiana* and *Artemia urmiana*: assessing molecular changes for high plateau adaptation. Sci China Life Sci, 2013, 56: 440−452, doi: 10.1007/s11427-013-4474-4

Brine shrimps, *Artemia* (Crustacea, Anostraca), have a broad geographical distribution and inhabit hypersaline environments that vary considerably in their anionic composition, climate, and altitude. The major anions that contribute to anionic composition include chloride, sulfate, carbonate, or combinations of up to all three [1]. *Artemia* species can be found in altitudes from sea level to over 4000 m above sea level; for example, in Tibet (4000 m) and in the East African Plateau (over 5000 m). *Artemia* can survive in harsh environments other than in high salinity, including low atmospheric pressure, rarefied air, cold temperature, and long winters, as well as in climatological conditions that range from humid to arid [2]. *Artemia* species are capable of switching their reproduction strategies between sexual and parthenogenetic forms in a single life cycle [1,3,4], and populations can survive through cold winters by producing diapause cysts [5–7]. Because of their distinct features, *Ar-*

*Corresponding author (email: junyu@big.ac.cn; wangweiwei301@gmail.com)

temia draw significant scientific interest from several areas including ecology, aquaculture, physiology, ecotoxicology, and genetics [8–12].

Animal mitochondrial (mt) genomes are circular DNA molecules ~17 kb in length that encode the major enzymes for oxidative metabolism and ATP production. Mitochondrial genome sequences have been used to determine phylogenetic relationships among animals of different taxa because of the conservation of the protein-coding sequences and the variability of the non-coding portion [13–15]. The mt genome in animals typically contains 37 genes (13 protein-coding, two ribosomal, and 22 transfer RNA genes) and one major non-coding region, the displacement loop or D-loop. The D-loop is the control region with a number of regulatory elements, such as extended termination-associated sequences (ETASs) and conserved sequence blocks (CSBs) that are responsible for replication and transcription of the mtDNA [16–22]. Although, overall, animal mt genomes are highly conserved in size and protein-coding components, substantial variations, including substitutions and indels [22,23], do exist, providing clues for functional differentiation. In addition, frequent loss of tRNA genes has been observed in arthropod mtDNAs [24].

Previous phylogenetic studies demonstrated that *Artemia tibetiana* and *Artemia urmiana* were indistinguishable based on markers derived from mt gene sequences such as *16S rRNA* and *cox1*; therefore, it was proposed that these two species might share a close ancestry [4,25,26]. However, how and when they evolved into two species remains to be discovered. Geographic isolation may have led to or accelerated the speciation process of *A. tibetiana* and *A. urmiana* [27]. *A. tibetiana* is a highland species, distributed in the Tibetan Plateau, which has evolved to tolerate the hypoxia and low temperatures that are typical for high altitude environments. Over the past decade, it has been reported (in *Ochotona curzoniae* and *Pantholops hodgsoni*) that complexes I and IV of the mt electron transport system, such as the cytochrome c oxidase (COX) complexes, harbor amino acid (AA) variations which are considered to be related to high altitude adaptation [28,29]. The results from both these studies suggested that, in the hypoxia environment, sequence variations may be selected and eventually fixed in the mt genomes.

Previous studies into high altitude adaptation have focused on mammalian species because of their medical relevance; however, we propose that by investigating sequence variations among the protein-coding genes of animal mitochondria, especially across diverse animal species, some of the secrets of high-altitude adaptation may be revealed. In this study, we report three full-length mt genome sequences; two from *A. tibetiana* isolates (ARC 1609 and ARC 1610) and one from *A. urmiana* (ARC 1227). We also used the publicly available mt genome sequence of *A. franciscana* in a comparative analysis to identify sequence variations that may be the result of environmental adaptation. In addition, we acquired some sequences from an *A. franciscana* isolate (ARC 1590) collected from Huangnigou near the Bohai Bay in China to validate high-altitude specificity.

1 Materials and methods

1.1 Sample collection

All the strains used in this study were stored as cysts at the Laboratory of Aquaculture & Artemia Reference Center (ARC), Ghent University, Belgium. All the strains were assigned ARC code numbers, including four populations which were collected from salt lakes. Two lakes are in Tibet; one is in Nima (ARC 1609; accession number: JQ975177; ~4555 m above mean sea level (amsl); 31°55′0.00″N/ 87°52′59.88″E) and one is in Yangnapeng Co (ARC 1610; accession number: JQ975178; ~4800 m amsl; 32°18′59.91″N/ 89°45′59.93″E). The third lake is Lake Urmia (ARC 1227; accession number: JQ975176; ~1275 m amsl; 37°20′N/ 45°40′E) in West Azerbaijan in Iran. The fourth lake is in Huangnigou, Shangdong Province, northern China (ARC 1590; ~100 m amsl; 37°31′17.51″N/ 120°39′12.80″E) where the samples (representative of *A. franciscana* in China) were collected in 2002 (data not show) [26]. The reference genome sequence is from an isolate of *A. franciscana* (~10 m amsl) collected in San Francisco Bay Salterns, California, USA, and retrieved from Genbank (accession number: NC_001620) [30,31].

1.2 DNA isolation

To obtain genomic DNA, we ground 0.25 g of the cysts to fine powder in the presence of liquid nitrogen and then immediately added 1 mL of 2× CTAB buffer (2% CTAB, 0.1 mol L^{-1} Tris-HCl at pH 8.0, 20 mmol L^{-1} EDTA, and 1.4 mol L^{-1} NaCl). For the adult samples, we homogenized the entire animal in 200 μL 2× CTAB buffer. The mixture was incubated at 65°C for 2 h, then centrifuged at 13000 rpm for 10 min. The viscous aqueous supernatant was carefully collected into a fresh microcentrifuge. DNA was extracted from the supernatant with chloroform/iso-amylalcohol (24∶1), precipitated with ice-cold isopropanol, washed with 75% ethanol, and re-suspended in double-distilled water containing 20 μg mL^{-1} RNase.

1.3 PCR amplification

We designed 24 pairs of primers based on the *A. franciscana* mt genome sequence from Genbank (accession number: NC_001620) (Table S1). PCR amplifications were carried out in a final reaction volume of 25 μL containing 2.5 μL 10× PCR buffer, 2 μL dNTP, 10 pmol L^{-1} of each primer, and 0.25 unit EasyTaq DNA polymerase (QIAGEN, Germany). The thermo-cycling program was set as holding on 5 min at 94°C, 35 cycles of 30 s at 94°C, 30 s at 55°C, and

1 min at 72°C, followed by a final extension of 10 min at 72°C. The PCR products were analyzed by electrophoresis in 1.0% agarose gel.

1.4 Sequencing the PCR products

We sequenced the PCR amplified DNA using a capillary sequencer (ABI 3730 XL, ABI, Foster City, USA). The sequencing reaction was started with 3 μL (about 50 ng) of the PCR products purified on a Millipore 96-well clean-up plate. We obtained DNA sequences from both the 5′ and 3′ directions using the 24 PCR primer pairs. The sequencing amplification protocol was set up as one cycle of 15 s at 95°C, 35 cycles of 15 s at 95°C, followed by an annealing step of 15 s at 50°C, and a final extension of 1.5 min at 60°C.

1.5 Sequence analysis

The sequences were assembled using the Phred-Phrap-Consed package (phred: 0.020425.c; phrap: 1.080812; consed: 19.0) [32,33]. NVR (nucleotide variation rate) was calculated as the number of mutated nucleotides divided by the length of the gene and AAVR (amino acid variation rate) was calculated as the number of mutated residues divided by the AA length of the gene. The calculations were based on the results of a CLUSTALW 2.0 pairwise sequence alignment [34]. The ratios of the rate of nonsynonymous (*Ka*) and synonymous (*Ks*) substitutions (*Ka*/*Ks* ratios) were generated using the KaKs_Calculator (v1.2) based on the GY-HKY model [35]. Circular mt physical maps were drawn using DNAMAN 6.0.40 (http://www.lynnon.com/) and the tRNAs and protein domains were annotated using tRNAscan-SE 1.23 [36] and InterProScan (version 4.8) [37] respectively. DAMBE with the default parameter settings (version 5.3.00) was used to analyze codon usage bias and to calculate the relative synonymous codon usage (RSCU) values [38]. Three-dimensional structures of the encoded proteins were predicted by SWISS MODEL (default parameters) [39] and conserved elements (ETASs and CSBs) were identified using the EDIALIGN and PRETTYPLOT programs in the EMBOSS package [40]. The neighbor-joining trees were built with MEGA 3.1 (1000 bootstrap replicates) [41]. We calculated the divergence time of *A. tibetiana* and *A. urmiana* using the D-loop sequences with MEGA 3.1 based on the time point (5.5 million years ago, Mya) when *Artemia* began to diverge from a common ancestor distributed in the Mediterranean area into eight species [42].

2 Results

2.1 General features of the *Artemia* genomes

In this study, we sequenced and assembled three *Artemia* mt genomes; two from *A. tibetiana* isolates (collected from two separate lakes in Tibet) and one from *A. urmiana* using a piecemeal strategy: first, to PCR amplify the mt genome into 24 amplicons that overlap the entire mt genome, and then to sequence each amplicon to obtain a final complete assembly. As the *Artemia* mt genomes are typical animal mt genome [22] with 37 genes: 13 protein-coding genes (*cytb*, *nd5*, *nd4*, *nd4l*, *nd1*, *nd2*, *cox1*, *cox2*, *atp8*, *atp6*, *cox3*, *nd3*, *nd6*), 22 tRNA genes, and 2 rRNA (*12S rRNA*, *16S rRNA*) genes, and a major non-coding region called the D-loop. The two *A. tibetiana* mt genomes are 15826 bp (ARC 1609) and 15743 bp (ARC 1610) long and the *A. urmiana* mt genome is 15945 bp (ARC 1227). The *A. franciscana* reference genome (NC_001620) is 15822 bp [30,31] long. The length variations are mainly confined to the D-loop region (Figure 1, Table S2).

The GC content of the mt genomes from *A. tibetiana* and *A. urmiana* was relatively low around 37%, which is characteristic of animal mt genomes [43–46]. However, the GC content of mt genes varies regardless of whether they are protein-coding or RNA genes. In general, we found relatively higher GC content in both the *A. tibetiana* and *A. urmiana* genomes compared with in *A. franciscana* (37% vs. 36%), and the higher GC content was observed mainly in the protein-coding rather than the tRNA genes. Among the 13 protein-coding genes, the GC content of three genes in particular was higher by at least 10% in *A. tibetiana* and *A. urmiana* compared with in *A. franciscana* (*atp6*: ~16%; *atp8*: ~12%; *nd6*: ~12%); and the GC content of *nd4l* was higher by nearly 11% in *A. tibetiana* compared with in both *A. urmiana* and *A. franciscana.* Among the tRNA genes, the GC content of $tRNA^{Pro}$ was higher by almost 8% in *A. tibetiana* compared with in *A. urmiana* and *A. franciscana*. It was reported that biased base composition of mtDNAs accumulated during replication and is related to spontaneous deamination of the C and A bases on the heavy strand [47]. The three protein-coding genes (*atp6*, *atp8*, *nd6*) that show the most fluctuations in their GC content may be involved in environmental adaptation; however, further studies are required to investigate this proposal in more detail.

Among the 13 protein-coding genes in the mt genomes, in *A. tibetiana*, eight of them use ATG as the start codon (*nd2*, *cox1*, *cox2*, *atp6*, *cox3*, *nd5*, *cytb*, and *nd1*), three use ATT (*atp8*, *nd3*, and *nd4l*), one uses ATA (*nd4*), and one uses GTG (*nd6*). In *A. urmiana*, most of the genes use the same start codons as the corresponding genes in *A. tibetiana*; two genes (*atp8* and *nd4l*) that use ATC as the start codon are the only exceptions. There are more differences in start codon usage between the more phylogenetically distant species *A. franciscana* and *A. tibetiana*. In *A. franciscana* six genes (*nd2*, *cox2*, *nd5*, *nd4l*, *nd6*, and *cytb*) use different start codons from *A. tibetiana*; two of them (*nd2* and *nd5*) use ATT and the four others use GTG (*cox2*), ATC (*nd6*), ATA (*nd4l*), and ATA (*cytb*). Similarly, variations in stop codon usage were observed. In *A. urmiana*, *nd2* (TAA),

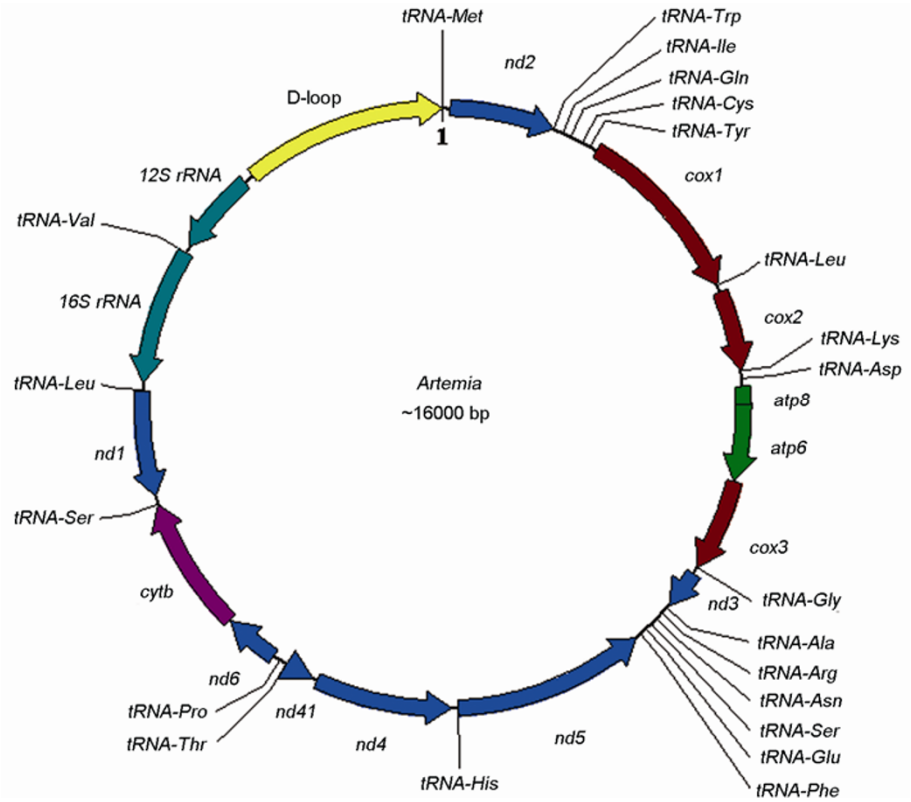

Figure 1 Circular physical map of the *A. tibetiana* mitochondrial genome. The *A. tibetiana* mt genome is ~16 kb long and has 37 genes. Thirteen of them are protein-coding genes that code for the subunits of the complexes of the mt electron transport system. Complex I encoding genes (*nd5*, *nd4*, *nd4l*, *nd1*, *nd2*, *nd3*, and *nd6*) are shown in blue; Complex III encoding gene (*cytb*) is in purple; Complex IV encoding genes (*cox1*, *cox2*, and *cox3*) are in brown; Complex V encoding genes (*atp8* and *atp6*) are in green. The other genes are RNA genes: 22 are tRNA genes and two are rRNA genes (*12S rRNA* and *16S rRNA*) shown in dark green. The major non-coding region (D-loop) is shown in yellow. The clockwise and counter-clockwise arrows indicate genes on the heavy and light strands, respectively.

cox3 (TAA) and *nd3* (TAG) use different stop codons than the corresponding *A. tibetiana* genes which use TAG, TAG, and TAA for *nd2*, *cox3*, and *nd3*, respectively. Four of the protein coding genes in *A. franciscana* have incomplete termination codons (*cox1*, *cox2*, *nd5*, and *atp6*), whereas, in *A. tibetiana* and *A. urmiana*, only *nd4* and *nd5* have incomplete termination signals (T).

2.2 Sequence variations among the *Artemia* species and ecotypes

We observed only a limited number of nucleotide variations between the mt genomes of the two *A. tibetiana* isolates, ARC 1609 and ARC 1610 (Table S3), from different salt lakes in Tibet. The variations were mainly in four genes with relatively low NVRs (at least five times lower than at the interspecies level): *12S rRNA* (0.014), *nd4* (0.003), *nd1* (0.002), and *nd5* (0.001). Because the variations were few, we selected ARC 1609 as the representative of *A. tibetiana* in our comparative analysis of interspecies variations.

Among the 37 mt protein-coding genes of the three *Artemia* species, the NVR was higher when *A. urmiana* and *A. tibetiana* were compared with *A. franciscana*, and lower between *A. tibetiana* and *A. urmiana* (Figure 2, Table S3). We found many single nucleotide indels in the protein-coding genes of *A. tibetiana* and *A. urmiana* compared with *A. franciscana*. For example, in *A. tibetiana* and *A. urmiana*, *nd5* had four deletions and one insertion, *nd4* had a single nucleotide insertion and two deletions, *cox2* had a single nucleotide insertion and a single nucleotide deletion, and *cox3* had a single nucleotide deletion and an 8-bp insertion (in total there were 13-bp insertions in *cox3*) (Figure S1A–D). Interestingly, the indels in these genes did not lead to any frameshifts nor did they create non-sense transcripts because each indel either was accompanied by a compensating indel nearby or the indels were in multiples of three; however, the indels did cause variations in the AA sequences. For example, in *cox2*, several 2-bp indels led to five AA variations (Leu-142, Pro-143, Val-144, Tyr-146, and Asn-147) and 14 indels in *cox3* resulted in 16 AA variations (Leu-47, Ile-48, Thr-49, Ile-50/Val-50, Met-51, Ile-52, Val-53, Ser-54, Tyr-55, Asn-125, Pro-126, Phe-127, Gln-128, Val-129, Pro-130, and Pro-229). To confirm some of these variations experimentally, we re-sequenced these four genes (*nd5*, *nd4*, *cox2*, *cox3*) in a local *A. franciscana* sample collected in China, and found that some of the indels

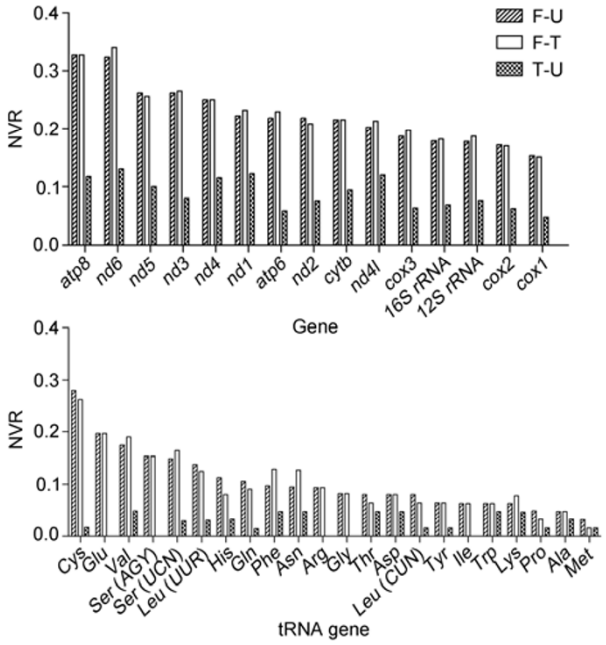

Figure 2 Nucleotide variation rates (NVRs) for the 37 mt genes among the reference *A. franciscana* (F), *A. tibetiana* (T), and *A. urmiana* (U).

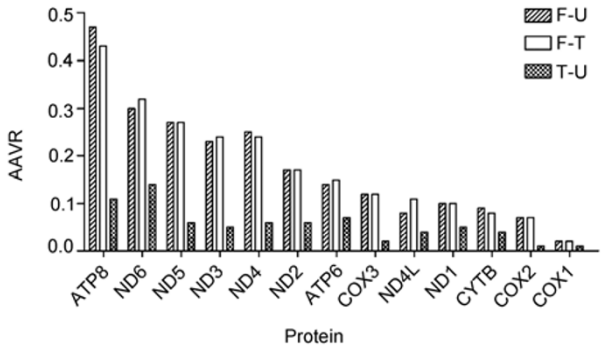

Figure 3 Amino acid variation rates (AAVRs) for the proteins encoded by the 13 mt protein-coding genes among the reference *A. franciscana* (F), *A. tibetiana* (T), and *A. urmiana* (U).

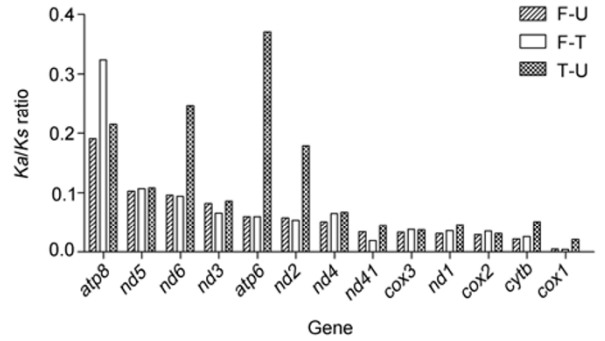

Figure 4 *Ka*/*Ks* ratios for the 13 mt protein-coding genes among the reference *A. franciscana* (F), *A. tibetiana* (T), and *A. urmiana* (U).

that we found in these genes did not exist in the mt genome of the local *A. franciscana* population. Assuming that sequencing errors are relatively random, it seems unlikely that so many consecutive indel pairs would be nearby by sequencing errors.

Nucleotide variations in protein-coding genes often lead to AA changes in the proteins that they encode. When *A. urmiana* and *A. tibetiana* were compared with the reference *A. franciscana*, *atp8* had the highest AAVRs, 0.47 and 0.43, respectively. Between *A. tibetiana* and *A. urmiana*, *nd6* had the highest AAVR (0.14), followed by *atp8* (0.11). *Cox1* had the lowest AAVRs: 0.02 between *A. franciscana* and *A. urmiana*, 0.02 between *A. franciscana* and *A. tibetiana*, and 0.01 between *A. tibetiana* and *A. urmiana* (Figure 3, Table S4). Overall the NADH dehydrogenases (Complex I) and ATP synthase (Complex V) showed higher diversity compared with other complexes, including Complex III and IV. The commonest variations involved the aliphatic AAs, Val, Ile, and Leu, which are all chemically similar (Figure S2).

The *Ka*/*Ks* ratio is of significance in phylogeny reconstruction and is especially useful for understanding evolutionary relations between protein-coding genes in closely related (recently diverged) species [48,49]. Between *A. franciscana* and *A. urmiana*/*A. tibetiana*, *atp8* had the highest *Ka*/*Ks* ratio (0.324 and 0.191, respectively) when compared with other genes; between *A. urmiana* and *A. tibetiana*, *atp6* had the highest Ka/Ks ratio (0.371) followed by *nd6* (0.247; Figure 4, Table S5). The *Ka*/*Ks* ratios of *nd2*, *atp6*, *cytb*, *cox1*, *nd4l*, *nd1*, and *nd6* were higher between *A. urmiana* and *A. tibetiana* compared with between *A. franciscana* and *A. urmiana*/*A. tibetiana*. This result suggested that these seven genes might play some role in high-altitude adaptation in this species. The variance analyses (NVR, AAVR, and GC content) also suggested that *atp8* and *atp6* may be under strong mutational pressure because they had higher values in all three measures than the other proteins.

2.3 tRNAs and codon usage

Most animal mt genomes encode 22 tRNAs [24]. We found many tRNA-related variations among the three *Artemia* species, including transitions, transversions, and indels. For example, in $tRNA^{Gln\ (CAA)}$, $tRNA^{Cys\ (TGC)}$, $tRNA^{His\ (CAC)}$, $tRNA^{Pro\ (CCA)}$ and $tRNA^{Leu\ (CTA)}$, all three types of variations were observed and some secondary structures, such as the TφC-loop and anticodon loop of $tRNA^{His\ (CAC)}$ and the amino acid arm of $tRNA^{Gln\ (CAA)}$, were affected (Figure S3).

We calculated the RSCU value to estimate codon usage bias (Table S6). In *A. franciscana*, the RSCU value for the CCU codon was found to be greater than 1, indicating that *A. franciscana* preferentially used CCU for Pro; on the other hand, *A. urmiana* and *A. tibetiana* preferentially used both CCC and CCU for Pro (RSCU>1). A similar situation was found in GGU codon usage for Gly and ACA for Thr. *A. franciscana* tended to use codons with relatively lower GC

content than *A. urmiana and A. tibetiana* and this may be related to replication-associated mutational pressure and natural selection [50].

2.4 Protein-coding genes

The *Artemia* mtDNA encodes 13 polypeptides and they are usually highly conserved to ensure the precision of ATP production. However, substantial variations, including SNPs and indels, do exist (Figure S1A–D). In both *A. tibetiana* and *A. urmiana*, the five AA variations in COX2 are in a β-sheet on the inner-membrane space side of the mitochondria that contains adenylate kinase [51] and the 16 AA variations in COX3 all caused changes in the 3D structure of the protein. In *A. tibetiana*, nine of the AA variations in COX3 were in helix II (Box A, Figure 5), six were in helix IV (Box B, Figure 5), and one AA insertion was between helixes VI and VII (Box C, Figure 5). The variations in Box B extended the length of helix IV and the insertion in Box C changed the conformation of the loop region between helixes VI and VII (Figure 5B). The AA variation patterns in Boxes A–C are similar in *A. tibetiana*, *A. urmiana*, and the *A. franciscana* collected in China, but different in the reference *A. franciscana* sequence. These variations in the COX3 protein may be candidate for species-specific mutations in Asian *Artemia* species. Further population-based studies are needed to confirm this observation.

ATP8 is an intrinsic membrane protein that is 53 AA long [52]. It was predicted to have three conserved regions: an N-terminal signal-peptide (1–25), a transmembrane stem (10–30), and a C-terminal positively charged region that was proposed to face the matrix space (30–53) (Figure 6A).

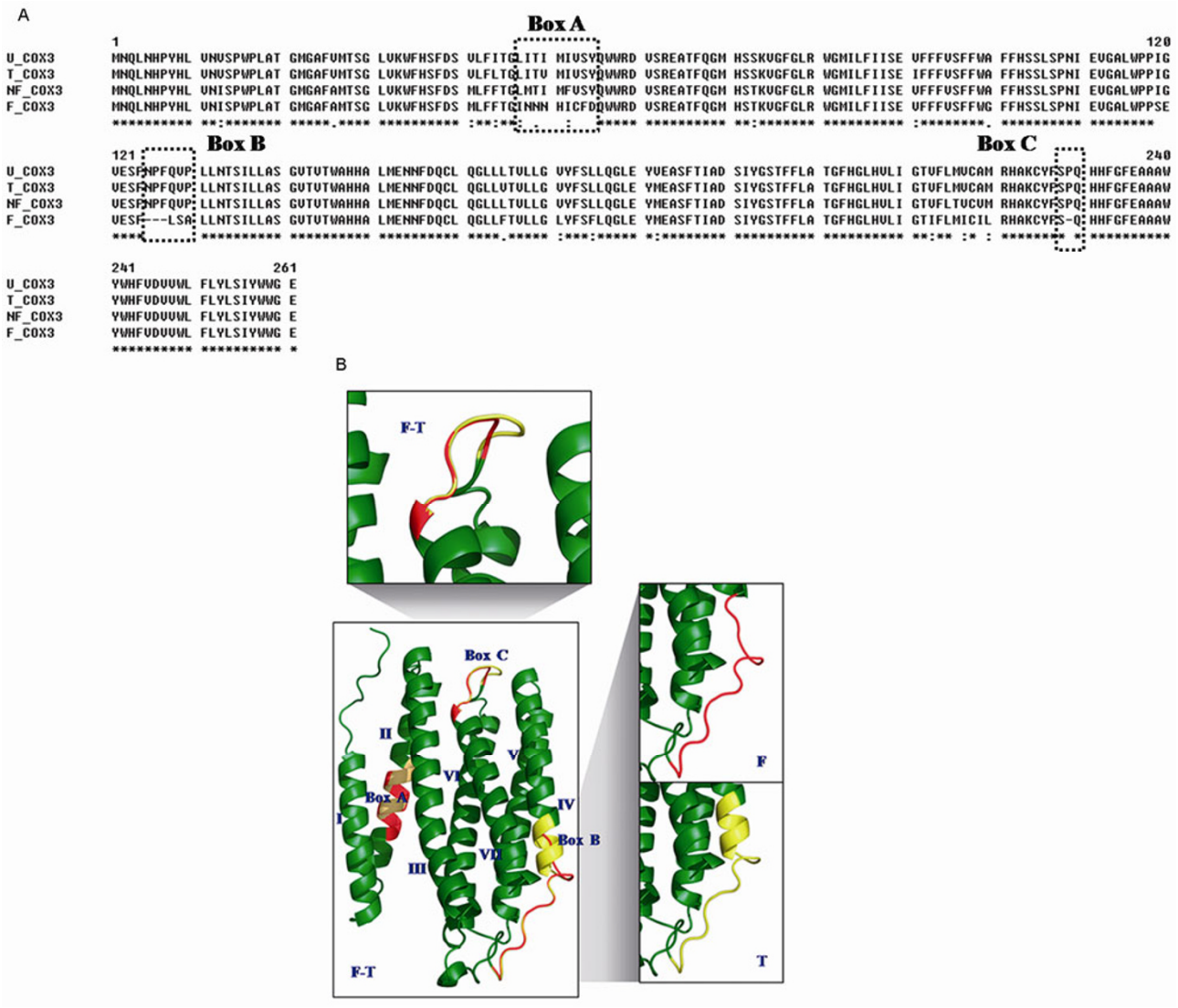

Figure 5 Sequence variations in COX3 and their distribution in its 3D structure. A, Partial alignment of the COX3 sequences from the reference *A. franciscana* (F), *A. tibetiana* (T), *A. urmiana* (U), and the local *A. franciscana* (NF). The three highly variable regions in which single nucleotide indels were found are shown (Boxes A, B, and C). B, Predicted 3D structure of COX3 in the reference *A. franciscana* (F, red) and *A. tibetiana* (T, yellow). The variable regions are enlarged in the surrounding panels. The helix numbers are marked in blue.

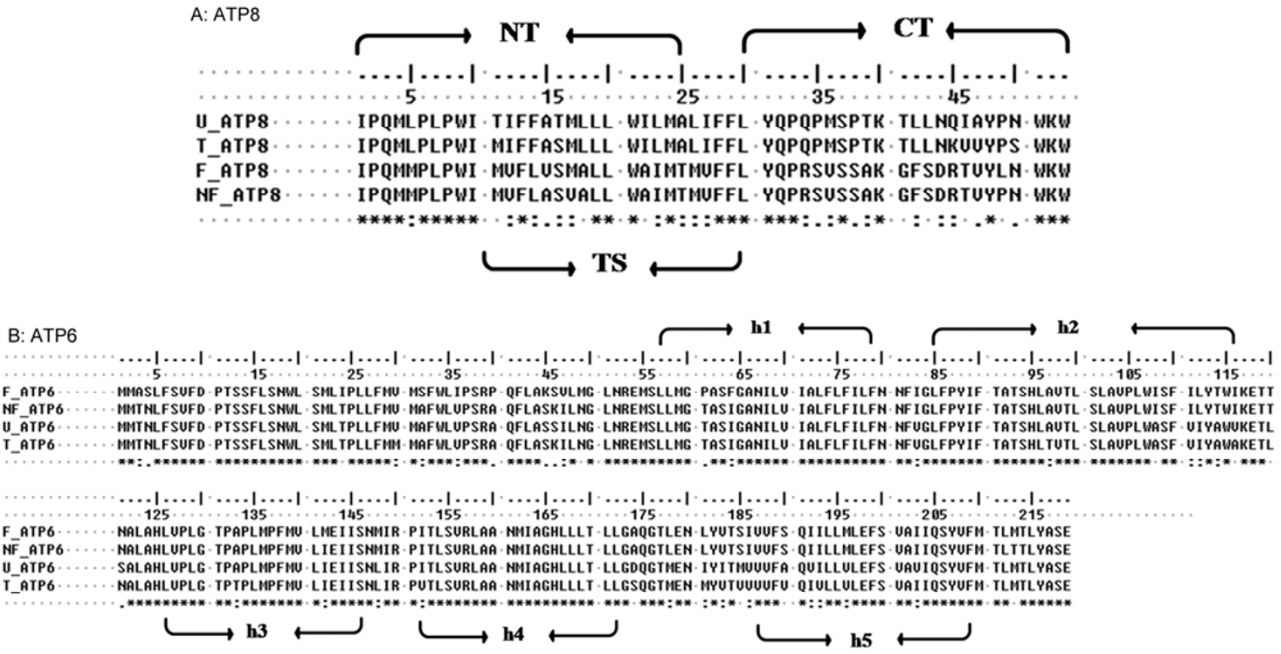

Figure 6 Amino acid sequence alignments and the putative secondary structure of ATP8 and ATP6 from the reference *A. franciscana* (F), *A. tibetiana* (T), *A. urmiana* (U), and the local *A. franciscana* (NF). * indicates positions where the amino acid residues (AAs) are identical in all four sequences. : indicates positions where the AAs are highly conserved (very similar). . indicates positions where the AAs are somewhat similar. Blank indicates positions were the AAs are either dissimilar or there are gaps. NT, N-terminal signal-peptide; TS, transmembrane stem; CT, C-terminal positively charged region.

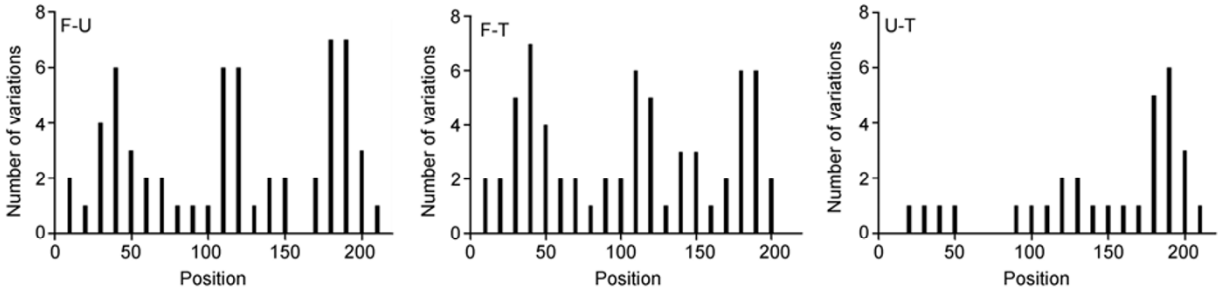

Figure 7 The AA variation counts for ATP6 (in a 20-AA window and 1-AA step in a total length of 219 AAs) for comparisons between the reference *A. franciscana* (F), *A. tibetiana* (T), and *A. urmiana* (U). All the AA variations are the result of substitutions in the gene sequences because no indels were found in the *atp6* mt genes in the *Artemia* species in this study. Indeed, until now, no indels have been reported in *atp6* mt genes in other animals.

ATP8 had the highest AAVR (Figure 3, Table S4) and *Ka*/*Ks* ratio among the 13 encoded proteins from *A. franciscana* compared with those from *A. urmiana*/*A. tibetiana* (Figure 4, Table S5). In *A. tibetiana* and *A. urmiana,* we identified 17 AA changes in ATP8; eight were in the transmembrane stem and nine were in the C-terminal (Table S7A). These changes may be related to the early divergence of the major *Artemia* species. Interestingly, between the two recently diverged species, *A. tibetiana* and *A. urmiana,* we identified six AA changes; in *A. tibetiana*, four of them (AAs at positions 45, 46, 47, and 50) were different from the AA changes identified in *A. urmiana* and *A. franciscana*. These species-specific variations in the ATP8 from *A. tibetiana* are likely candidates as high-altitude adaptation related.

ATP6 has five predicted transmembrane helixes in its 219 AA long sequence. In *A. urmiana* and *A. tibetiana*, ATP6 had the highest *Ka*/*Ks* ratio among the 13 highly conserved protein-coding genes (Figure 4, Table S5). The five transmembrane helixes were denoted as h1 (57–79), h2 (85–116), h3 (126–146), h4 (152–172), and h5 (187–209; Table S7B). The h1 and h2 domains had no substitutions and h3 had only one. There were also seven substitutions in the non-functional domains; four of them were between the two well-conserved h4 and h5 domains that lie towards the C-termini. The h4 and h5 helixes are amphiphilic α-helixes in which the charged/polar residues lie along one face of the helix. It has been proposed that these helixes may be involved in proton translocation [53,54]. We identified five substitutions, including two AA changes between Val and

Ala and three between Ile and Val in these two TMH domains (Figure 6B). We also found that there were more AA variations in the N-termini between *A. franciscana* and *A. urmiana/A. tibetiana*, while, between *A. urmiana* and *A. tibetiana*, the AA variations were concentrated in the C-termini (Figure 7).

2.5 The D-loop region

To determine the similarities among the *Artemia* species, we performed clustering analyses using the neighbor-joining (NJ) method and the D-loop sequences with the *Daphnia* sequence as the outgroup [44] (Figure 8A). The sequences grouped into three major clades, and the tree topology was consistent with phylogenetic trees that were built previously based on DNA barcoding and restriction fragment length polymorphism (RFLP) [4,26]. Based on the D-loop sequences, we calculated the divergence time of *A. tibetiana* and *A. urmiana* as ~3.47 Mya (Figure 8A), and based on the whole mt genome sequences, as ~2.06 Mya (Figure S4).

We detected two ETASs (ETAS1 and ETAS2) and three CSBs (CSB1, CSB2, and CSB3) with long conserved sequences, and a central region between the ETASs and CSBs (Figure 8B) in the D-loops of the *Artemia* species. We found that there were more variations at the interspecies level (among *A. tibetiana*, *A. urmiana*, and *A. franciscana*) than at the intraspecies level (between the two *A. tibetiana* isolates). We also observed obvious variations between *A. tibetiana* and *A. urmiana*; for example, in CSB1 in *A. urmiana*, there was a TTTT in place of a CCG in *A. tibetiana*, and, between ETAS1 and ETAS2, a long insertion (90 bp) was found only in *A. tibetiana* and not in *A. urmiana* (Figure 8C). Similarly, an 80-bp spacer between CSB2 and CSB3 was found in the D-loop of the two *A. franciscana* isolates, but not in either *A. tibetiana* or *A. urmiana* (Figure 8D).

3 Discussion

In this study, we sequenced three *Artemia* mt genomes, and conducted comparative analysis and 3D structure modeling to assess the potential implications of the sequence variations in the protein-coding and non-coding (the D-loop) regions of the genome. Consistent differences between the variations in *A. tibetiana*, *A. urmiana*, and the local *A. franciscana* (China isolate) compared with the reference *A. franciscana* (USA isolate) were observed in the protein-coding genes and these differences were interpreted as Asian species-specific. Variations in the protein-coding genes between *A. tibetiana* and *A. urmiana* were regarded as high-altitude adaptations.

3.1 Protein-coding genes

In general, when comparing the reference *A. franciscana* with *A. tibetiana* and *A. urmiana*, we observed that the NVR and AAVR were high. *Atp8* had the highest variation rate among the 13 highly conserved coding-genes. Some of the variations, such as those in *cox2* and *cox3*, appeared to be Asian-specific, while others between *A. tibetiana* and *A. urmiana*, such as those in *atp6*, appeared to be related to the species' high-altitude adaptation.

Of the 13 mt proteins, COX1 and COX2, the core subunits of Complex IV, are directly involved in electron transfer and proton translocation processes, whereas COX3, a member of the catalytic core of Complex IV, is thought to be a regulator [55,56]. Because lipid binding plays important roles in the assembly, structure, and activity of these proteins [51], mutations in the highly conserved residues that are involved in the interaction between the COX proteins and lipid molecules may lead to functionally defective complexes [51,57]. We found indels and variations in *cox2* and some of them led to AA variations within the functional domains of the COX2 protein. Five of the variations were in the β-sheet regions that are known to be involved in the interaction between the lipid bilayer and COX. More interestingly, we found COX3-specific AA changes in *A. tibetiana* which were predicted to lead to conformational changes in the 3D structure of the protein, which might affect the activity or assembling efficiency of Complex IV. Moreover, recent research indicated that COX induced the expression of nuclear hypoxia genes, possibly via a pathway that involves protein nitration [58]. Luo et al. [28] suggested that COX may play an important role in hypoxia adaptation, based on an analysis of the mt genome of *Ochotona curzoniae*. It has also been reported that mt COX1 may be related to plateau adaptation in Tibetan antelope [29]. Oxygen is the ultimate electron acceptor, in a process that is catalyzed by COX; therefore, modifications in COX might be expected to increase its ability to cope with reduced oxygen supply. Here, we propose that some of the AA variations in COX2 and COX3 may be related to hypoxia adaptation.

The F_1F_0-type ATP synthase (ATPase) complex consists of a membrane-embedded domain (F_0) as the proton channel (allowing conversion of the proton gradient) and a transmembrane segment (F_1, generating ATP by the phosphorylation of ADP) [59–61]. ATP6 and ATP8 are the two mtDNA-encoded subunits that are incorporate into Complex V where they play essential roles in the final assembly of ATPase [52,62]. Studies in yeast demonstrated that the N-terminal signal-peptide domain of ATP8 is responsible for its function [63], whereas the C-terminal domain is required for the assembly of F_0 [64] and the transmembrane stem accommodates charged AA residues [65]. The *Artemia* ATP8 proteins shares a common four-residue sequence motif (IPQM) at their N-termini, which is similar to the ATP8-specific motif (MPQM) identified previously in other

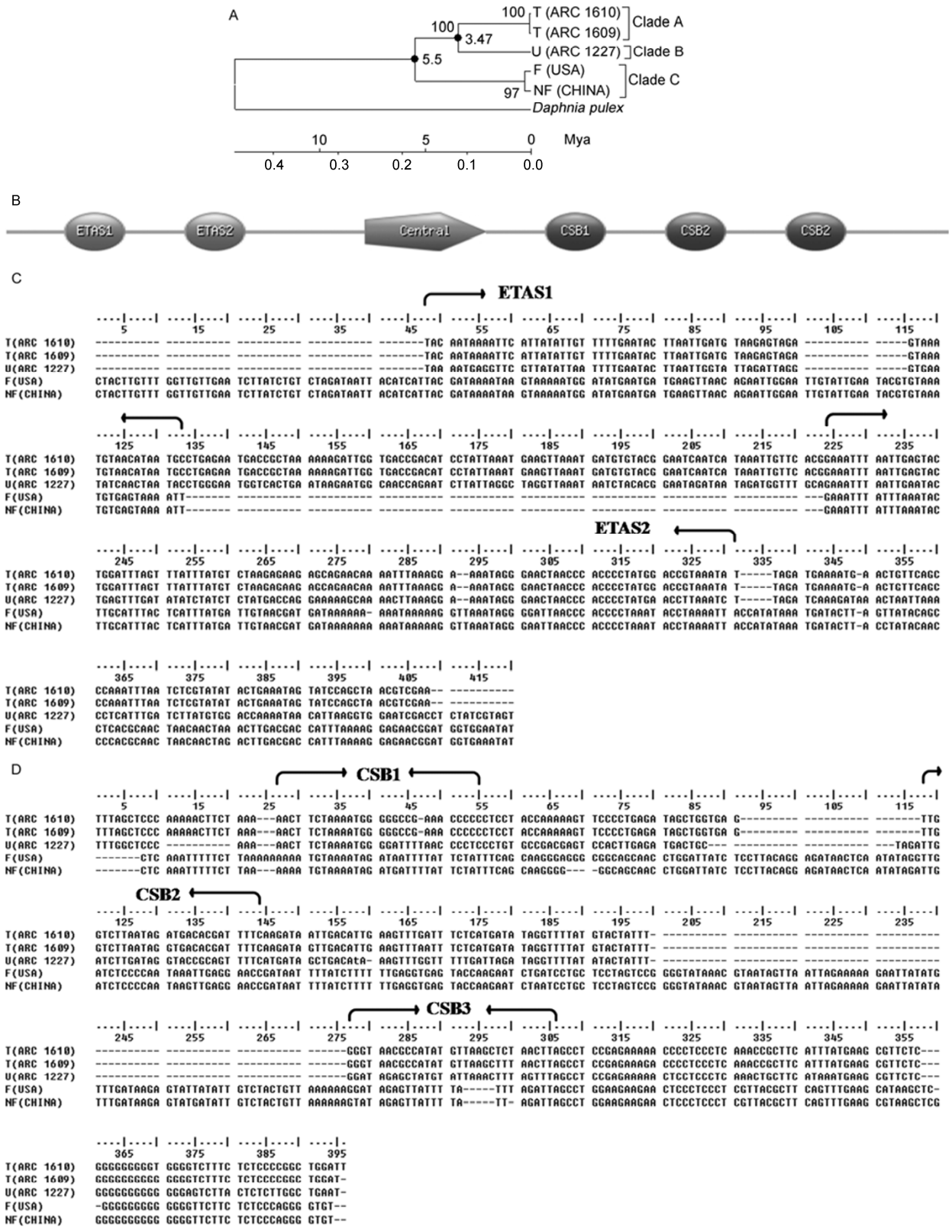

Figure 8 Analyses of the non-coding D-loop region of the *Artemia* mt genomes. A, Phylogeny of the *Artemia* D-loop region. B, Schematic diagram of the putative conserved elements in the *Artemia* D-loop region. The two ETASs (ETAS1 and ETAS2) and three CSBs (CSB1, CSB2, and CSB3) with the long conserved sequence (Central) are shown. C, Sequence alignment of the ETAS region from the reference *A. franciscana* (F), *A. tibetiana* (T), *A. urmiana* (U), and the local *A. franciscana* (NF). D, Sequence alignment of the CSB region from the reference *A. franciscana* (F), *A. tibetiana* (T), *A. urmiana* (U), and the local *A. franciscana* (NF).

animals [66]. In the present study, we detected AA changes in both the transmembrane stem and C-terminal domains of ATP8, and importantly, based on the physiochemical properties of the variable AA, we determined that these changes may affect the hydrophobicity, charge, and conformation of the entire complex. Therefore, we propose that variations in the transmembrane stem and C-terminal domains of ATP8 may influence the assembly and functioning of Complex V, causing differences in its ability to produce ATP efficiently. How these variations might affect its ATPase activity needs further investigation. Because *atp8* showed the highest adaptive variations in both *A. tibetiana* and *A. urmiana* compared with in *A. franciscana*, we propose that it may have evolved faster and was exposed to stronger natural selection than the other mt genes.

ATP6 was predicted to have five hydrophobic transmembrane α-helices, an N-terminus that protruded into the inner membrane space, and a C-terminus that projected into the matrix. Together with other membrane integral proteolipid subunits, ATP6 forms the proton channel of mt ATPase [53,61,67]. *Atp6* had a higher *Ka*/*Ks* ratio between the two geographic species, *A. tibetiana* and *A. urmiana*, even though they are closely related, compared with the *Ka*/*Ks* ratios among phylogenetically distant species. The AA variations suggested that *atp6* was under a strong selection as a result of environmental pressure and geographic isolation; 15 AA changes, concentrated at the C-terminal region, were detected between *A. urmiana* and *A. tibetiana*. Although the h4 and h5 helixes were highly-conserved and are known to be involved in proton translocation, they also harbored five substitutions including two AA alterations between Val and Ala and three AA alterations between Ile and Val. These variations in h4 and h5 may affect proton translocation that could change the balance of ATP and energy production [68,69]. *Atp6* has been reported to be highly variable in the mtDNA of humans who live in the arctic zones [69–72] and *atp6* has been found to be under enhanced adaptive selection in Tibetans living at the higher altitudes [68]. Therefore, we propose that the *atp6* mutations in the *A. tibetiana* mt genome are candidates for *Artemia* adaptation to extreme cold and hypoxia.

3.2 Non-coding regions

The *Artemia* D-loop is similar to the D-loops in other animal mt genomes. It is rich in repetitive sequences and contains most of the sequence variations that have been found in mt genomes. However, the D-loop contains major regulatory elements for replication and transcription (ETASs and CSBs) and both these regions are conserved among mammals. ETASs, located at the 5′ end of the control region, contain *cis* elements for the regulation of DNA synthesis of the heavy strand [73]. CSBs are located at the 3′ end of the control region and contain major regulatory elements: two promoters (HSP and LSP) and the replication origin of the heavy strand (O_H). It was reported that this domain has a stable cloverleaf-like secondary structure [74]. CSB1 is the only essential element that is common to all mammalian mt genomes, whereas CSB2 and CSB3 are functionally different and species-specific [75,76]. It has been proposed that related species may develop different regulatory elements for gene expression and DNA replication [75,77].

Mutations in the non-coding region are also known to play functional roles, such as altering mt energy production. D-loop mutants, for example, may not generate defective mtDNA gene products but they may alter the ratio of the mt and nuclear gene products [71]. We detected many variations and large indels in the D-loop regions among the three *Artemia* species. In general, these differences were found mainly between *A. franciscana* and *A. tibetiana*/*A. urmiana*. We also identified variations within the ETAS1 and CSB1 domains between *A. tibetiana* and *A. urmiana* which are phylogenetically close. Because of the changes that we observed in the regulatory elements between the two geographical isolates, we propose that sequence variations in the non-coding region may contribute to high-altitude adaptation. Pinar et al reviewed the known variants in the control region of human mtDNA and pointed out their associations with longevity and climatic adaptation [71].

Based on our phylogenetic analysis, we determined the divergence time between *A. tibetiana* and *A. urmiana* as 3.47 Mya (based on the D-loop sequences). This estimation is consistent with the formation of the Tibetan Plateau (~3 Mya) [78–80], suggesting that a geographic speciation process between *A. tibetiana* and *A. urmiana* may have coincided with high-altitude adaptation. Moreover, given its limited ability to migrate, *Artemia* may serve as an excellent model to validate the hypothesis that high-altitude adaptation and speciation are related to geographic isolation as the Himalayas were created and rose up.

This work was supported by the National Natural Science Foundation of China (30221004) to Wang WeiWei and the National Basic Research Program from Ministry of Science and Technology of China (2011CB944100) to Yu Jun.

1 Triantaphyllidis G V, Abatzopoulos T J, Sorgeloos P. Review of the biogeography of the genus *Artemia* (Crustacea, Anostraca). J Biogeogr, 1998, 25: 213–226
2 Van Stappen G, Sui L Y, Xin N H, et al. Characterisation of high-altitude *Artemia* populations from the Qinghai-Tibet Plateau, PR China. Hydrobiologia, 2003, 500: 179–192
3 Campos-Ramos R, Maeda-Martinez A M, Obregon-Barboza H, et al. Mixture of parthenogenetic and zygogenetic brine shrimp *Artemia* (Branchiopoda : Anostraca) in commercial cyst lots from Great Salt Lake, UT, USA. J Exp Mar Biol Ecol, 2003, 296: 243–251
4 Baxevanis A D, Kappas I, Abatzopoulos T J. Molecular phylogenetics and asexuality in the brine shrimp *Artemia*. Mol Phylogenet Evol, 2006, 40: 724–738
5 Busa W B, Crowe J H. Intracellular pH regulates transitions between dormancy and development of brine shrimp (*Artemia salina*) embryos. Science, 1983, 221: 366–368
6 Clegg J S, Golub A L. Protein synthesis in *Artemia salina* embryos.

II. Resumption of RNA and protein synthesis upon cessation of dormancy in the encysted gastrula. Dev Biol, 1969, 19: 178–200
7 Hand S C, Gnaiger E. Anaerobic dormancy quantified in *Artemia* embryos: a calorimetric test of the control mechanism. Science, 1988, 239: 1425–1427
8 Abatzopoulos Th J, Beardmore J A, Clegg J S, et al., eds. *Artemia*: Basic and Applied Biology. Berlin & Heidelberg: Springer-Verlag, 2002
9 Wu G, Zhang H X, Sun J, et al. Diverse LEA (late embryogenesis abundant) and LEA-like genes and their responses to hypersaline stress in post-diapause embryonic development of *Artemia franciscana*. Comp Biochem Phys B, 2011, 160: 32–39
10 MacRae T H, Hu Y, Bojikova-Fournier S, et al. The structural stability and chaperone activity of artemin, a ferritin homologue from diapause-destined *Artemia* embryos, depend on different cysteine residues. Cell Stress Chaperon, 2011, 16: 133–141
11 Van Stappen G, Manaffar R, Zare S, et al. Sediment cores from Lake Urmia (Iran) suggest the inhabitation by parthenogenetic *Artemia* around 5000 years ago. Hydrobiologia, 2011, 671: 65–74
12 Sant'Ana A E G, dos Santos A F, Cavada B S, et al. Toxicity of some glucose/mannose-binding lectins to *Biomphalaria glabrata* and *Artemia salina*. Bioresource Technol, 2010, 101: 794–798
13 Segawa R D, Aotsuka T. The mitochondrial genome of the Japanese freshwater crab, *Geothelphusa dehaani* (Crustacea: Brachyura): evidence for its evolution via gene duplication. Gene, 2005, 355: 28–39
14 Rogaev E I, Moliaka Y K, Malyarchuk B A, et al. Complete mitochondrial genome and phylogeny of pleistocene mammoth *Mammuthus primigenius*. PLoS Biol, 2006, 4: e73
15 Boore J L. Complete mitochondrial genome sequence of *Urechis caupo*, a representative of the phylum Echiura. BMC Genomics, 2004, 5: 67
16 Fernandez-Silva P, Enriquez J A, Montoya J. Replication and transcription of mammalian mitochondrial DNA. Exp Physiol, 2003, 88: 41–56
17 Yasukawa T. Overview of mammalian mitochondrial DNA replication and transcription. Tanpakushitsu Kakusan Koso, 2005, 50: 1727–1731
18 Kuhn K, Streit B, Schwenk K. Conservation of structural elements in the mitochondrial control region of *Daphnia*. Gene, 2008, 420: 107–112
19 Carrodeguas J A, Vallejo C G. Mitochondrial transcription initiation in the Crustacean *Artemia franciscana*. Eur J Biochem, 1997, 250: 514–523
20 Shadel G S, Clayton D A. Mitochondrial transcription initiation. Variation and conservation. J Biol Chem, 1993, 268: 16083–16086
21 Bonawitz N D, Clayton D A, Shadel G S. Initiation and beyond: multiple functions of the human mitochondrial transcription machinery. Mol Cell, 2006, 24: 813–825
22 Gissi C, Iannelli F, Pesole G. Evolution of the mitochondrial genome of Metazoa as exemplified by comparison of congeneric species. Heredity, 2008, 101: 301–320
23 Hassanin A, Leger N, Deutsch J. Evidence for multiple reversals of asymmetric mutational constraints during the evolution of the mitochondrial genome of Metazoa, and consequences for phylogenetic inferences. Syst Biol, 2005, 54: 277–298
24 Jühling F, Pütz J, Bernt M, et al. Improved systematic tRNA gene annotation allows new insights into the evolution of mitochondrial tRNA structures and into the mechanisms of mitochondrial genome rearrangements. Nucleic Acids Res, 2012, 40: 2833–2845
25 Baxevanis A D, Triantaphyllidis G V, Kappas I, et al. Evolutionary assessment of *Artemia tibetiana* (Crustacea, Anostraca) based on morphometry and 16s rRNA RFLP analysis. J Zool Syst Evol Res, 2005, 43: 189–198
26 Wang W, Luo Q, Guo H, et al. Phylogenetic analysis of brine shrimp (*Artemia*) in China using DNA barcoding. Genom Proteom Bioinform, 2008, 6: 155–162
27 He S, Cao W, Chen Y. The uplift of Qinghai-Xizang (Tibet) Plateau and the vicariance speciation of glyptosternoid fishes (Siluriformes: Sisoridae). Sci China Ser C-Life Sci, 2001, 44: 644–651
28 Luo Y, Gao W, Gao Y, et al. Mitochondrial genome analysis of *Ochotona curzoniae* and implication of cytochrome *c* oxidase in hypoxic adaptation. Mitochondrion, 2008, 8: 352–357
29 Xu S Q, Yang Y Z, Zhou J, et al. A mitochondrial genome sequence of the Tibetan Antelope (*Pantholops hodgsonii*). Genomics Proteomics Bioinformatics, 2005, 3: 5–17
30 Perez M L, Valverde J R, Batuecas B, et al. Speciation in the *Artemia* genus: mitochondrial DNA analysis of bisexual and parthenogenetic brine shrimps. J Mol Evol, 1994, 38: 156–168
31 Ramon Valverde J, Batuecas B, Moratilla C, et al. The complete mitochondrial DNA sequence of the Crustacean *Artemia franciscana*. J Mol Evol, 1994, 39: 400–408
32 Ewing B, Green P. Base-calling of automated sequencer traces using phred. II. Error probabilities. Genome Res, 1998, 8: 186–194
33 Gordon D, Abajian C, Green P. Consed: a graphical tool for sequence finishing. Genome Res, 1998, 8: 195–202
34 Larkin M A, Blackshields G, Brown N P, et al. Clustal W and Clustal X version 2.0. Bioinformatics, 2007, 23: 2947–2948
35 Zhang Z, Li J, Zhao X Q, et al. Kaks_calculator: calculating Ka and Ks through model selection and model averaging. Genom Proteom Bioinform, 2006, 4: 259–263
36 Lowe T M, Eddy S R. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res, 1997, 25: 955–964
37 Quevillon E, Silventoinen V, Pillai S, et al. InterProScan: protein domains identifier. Nucleic Acids Res, 2005, 33: W116–120
38 Xia X, Xie Z. DAMBE: software package for data analysis in molecular biology and evolution. J Hered, 2001, 92: 371–373
39 Arnold K, Bordoli L, Kopp J, et al. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. Bioinformatics, 2006, 22: 195–201
40 Olson S A. EMBOSS opens up sequence analysis. European molecular biology open software suite. Brief Bioinform, 2002, 3: 87–91
41 Kumar S, Tamura K, Nei M. MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. Brief Bioinform, 2004, 5: 150–163
42 Abreu-Grobois F A, Beardmore J A. Genetic differentiation and speciation in the brine shrimp *Artemia*. Prog Clin Biol Res, 1982, 96: 345–376
43 Yamauchi M M, Miya M U, Nishida M. Complete mitochondrial DNA sequence of the Japanese spiny lobster, *Panulirus japonicus* (Crustacea: Decapoda). Gene, 2002, 295: 89–96
44 Crease T J. The complete sequence of the mitochondrial genome of *Daphnia pulex* (Cladocera: Crustacea). Gene, 1999, 233: 89–99
45 Clary D O, Wolstenholme D R. The *Drosophila* mitochondrial genome. Oxf Surv Eukaryot Genes, 1984, 1: 1–35
46 Webster B L, Rudolfova J, Horak P, et al. The complete mitochondrial genome of the bird schistosome *Trichobilharzia regenti* (Platyhelminthes: Digenea), causative agent of cercarial dermatitis. J Parasitol, 2007, 93: 553–561
47 Reyes A, Gissi C, Pesole G, et al. Asymmetrical directional mutation pressure in the mitochondrial genome of mammals. Mol Biol Evol, 1998, 15: 957–966
48 Fay J C, Wu C I. Sequence divergence, functional constraint, and selection in protein evolution. Annu Rev Genomics Hum Genet, 2003, 4: 213–235
49 Kimura M. The neutral theory of molecular evolution—a review of recent evidence. Japan J Genet, 1991, 66: 367–386
50 Lin Q, Cui P, Ding F, et al. Replication-associated mutational pressure (RMP) governs strand-biased compositional asymmetry (SCA) and gene organization in animal mitochondrial genomes. Curr Genom, 2012, 13: 28–36
51 Qin L, Sharpe M A, Garavito R M, et al. Conserved lipid-binding sites in membrane proteins: a focus on cytochrome *c* oxidase. Curr Opin Struct Biol, 2007, 17: 444–450
52 Fernandez-Vizarra E, Tiranti V, Zeviani M. Assembly of the oxidative phosphorylation system in humans: what we have learned by studying its defects. Biochim Biophys Acta, 2009, 1793: 200–211

53 Nagley P. Eukaryote membrane genetics: the F_o sector of mitochondrial ATP synthase. Trends Genet, 1988, 4: 46–51
54 Cox G B, Fimmel A L, Gibson F, et al. The mechanism of ATP synthase: a reassessment of the functions of the b and a subunits. Biochim Biophys Acta, 1986, 849: 62–69
55 da Fonseca R R, Johnson W E, O'Brien S J, et al. The adaptive evolution of the mammalian mitochondrial genome. BMC Genomics, 2008, 9: 119
56 Hosler J P, Ferguson-Miller S, Mills D A. Energy transduction: proton transfer through the respiratory complexes. Annu Rev Biochem, 2006, 75: 165–187
57 Svensson-Ek M, Abramson J, Larsson G, et al. The X-ray crystal structures of wild-type and EQ(I-286) mutant cytochrome *c* oxidases from *Rhodobacter sphaeroides*. J Mol Biol, 2002, 321: 329–339
58 Benamar A, Rolletschek H, Borisjuk L, et al. Nitrite-nitric oxide control of mitochondrial respiration at the frontier of anoxia. Biochim Biophys Acta-Bioenerg, 2008, 1777: 1268–1275
59 Bowler M W, Montgomery M G, Leslie A G, et al. Ground state structure of F_1-ATPase from bovine heart mitochondria at 1.9 Å resolution. J Biol Chem, 2007, 282: 14238–14242
60 Yoshida M, Muneyuki E, Hisabori T. ATP synthase—a marvellous rotary engine of the cell. Nat Rev Mol Cell Biol, 2001, 2: 669–677
61 Stock D, Leslie A G, Walker J E. Molecular architecture of the rotary motor in ATP synthase. Science, 1999, 286: 1700–1705
62 Wittig I, Meyer B, Heide H, et al. Assembly and oligomerization of human ATP synthase lacking mitochondrial subunits a and A6L. Biochim Biophys Acta, 2010, 1797: 1004–1011
63 Stephens A N, Khan M A, Roucou X, et al. The molecular neighborhood of subunit 8 of yeast mitochondrial F_1F_0-ATP synthase probed by cysteine scanning mutagenesis and chemical modification. J Biol Chem, 2003, 278: 17867–17875
64 Papakonstantinou T, Galanis M, Nagley P, et al. Each of three positively-charged amino acids in the C-terminal region of yeast mitochondrial ATP synthase subunit 8 is required for assembly. Biochim Biophys Acta, 1993, 1144: 22–32
65 Papakonstantinou T, Law R H, Nagley P, et al. Non-functional variants of yeast mitochondrial ATP synthase subunit 8 that assemble into the complex. Biochem Mol Biol Int, 1996, 39: 253–260
66 Hong S, Pedersen P L. Mitochondrial ATP synthase: a bioinformatic approach reveals new insights about the roles of supernumerary subunits g and A6L. J Bioenerg Biomembr, 2004, 36: 515–523
67 Devenish R J, Prescott M, Boyle G M, et al. The oligomycin axis of mitochondrial ATP synthase: OSCP and the proton channel. J Bioenerg Biomembr, 2000, 32: 507–515
68 Gu M L, Wang Y J, Shi L, et al. Comparison on mitochondrial *atp6*, *atp8* and *cytb* genes between Chinese Tibetans in three different zones: detecting the signature of natural selection on mitochondrial genome. Yi Chuan, 2009, 31: 147–152
69 Elson J L, Turnbull D M, Howell N. Comparative genomics and the evolution of human mitochondrial DNA: assessing the effects of selection. Am J Hum Genet, 2004, 74: 229–238
70 Mishmar D, Ruiz-Pesini E, Golik P, et al. Natural selection shaped regional mtDNA variation in humans. Proc Natl Acad Sci USA, 2003, 100: 171–176
71 Coskun P E, Ruiz-Pesini E, Wallace D C. Control region mtDNA variants: longevity, climatic adaptation, and a forensic conundrum. Proc Natl Acad Sci USA, 2003, 100: 2174–2176
72 Bhopal R S, Rafnsson S B. Could mitochondrial efficiency explain the susceptibility to adiposity, metabolic syndrome, diabetes and cardiovascular diseases in South Asian populations? Int J Epidemiol, 2009, 38: 1072–1081
73 Madsen C S, Ghivizzani S C, Hauswirth W W. Protein binding to a single termination-associated sequence in the mitochondrial DNA D-loop region. Mol Cell Biol, 1993, 13: 2162–2171
74 Brown G G, Gadaleta G, Pepe G, et al. Structural conservation and variation in the D-loop-containing region of vertebrate mitochondrial DNA. J Mol Biol, 1986, 192: 503–511
75 Saito S, Tamura K, Aotsuka T. Replication origin of mitochondrial DNA in insects. Genetics, 2005, 171: 1695–1705
76 Zhang D X, Szymura J M, Hewitt G M. Evolution and structural conservation of the control region of insect mitochondrial DNA. J Mol Evol, 1995, 40: 382–391
77 Schultheis A S, Weigt L A, Hendricks A C. Arrangement and structural conservation of the mitochondrial control region of two species of Plecoptera: utility of tandem repeat-containing regions in studies of population genetics and evolutionary history. Insect Mol Biol, 2002, 11: 605–610
78 Tapponnier P, Zhiqin X, Roger F, et al. Oblique stepwise rise and growth of the Tibet Plateau. Science, 2001, 294: 1671–1677
79 Royden L H, Burchfiel B C, van der Hilst R D. The geological evolution of the Tibetan Plateau. Science, 2008, 321: 1054–1058
80 Zhong Dalai D L. Rising process of the Qinghai-Xizang (Tibet) plateau and its mechanism. Sci China Ser D-Earth Sci, 1996, 39: 369–379

Supporting Information

Table S1 The 24 primers pairs used in this study

Table S2 Characteristics of the three mitochondrial genomes of the reference *A. franciscana* (F), *A. urmiana* (U), and *A. tibetiana* (T)

Table S3 Nucleotide variation rates (NVRs) of the 37 mitochondrial genes among the reference *A. franciscana* (F), *A. tibetiana* (T), and *A. urmiana* (U)

Table S4 Amino acid variation rates (AAVRs) in the protein-coding genes among the reference *A. franciscana* (F), *A. tibetiana* (T), and *A. urmiana* (U)

Table S5 *Ka*/*Ks* ratios of the 13 protein-coding genes among the reference *A. franciscana* (F), *A. tibetiana* (T), and *A. urmiana* (U)

Table S6 Codon usage in the reference *A. franciscana* (F), *A. tibetiana* (T), and *A. urmiana* (U)

Table S7 Amino acid variations in ATP8 and ATP6 among the reference *A. franciscana* (F), *A. tibetiana* (T), and *A. urmiana* (U)

Figure S1 Alignment of the *nd5* (A), *nd4* (B), *cox2* (C), and *cox3* (D) genes. Indels are highlighted in red.

Figure S2 Heatmap of amino acid variations among the reference *A. franciscana* (F), *A. tibetiana* (T), and *A. urmiana* (U). The heatmap shows a full view of the amino acid variations among *A. franciscana*, *A. tibetiana*, and *A. urmiana*. The values of the color-bar indicate the variation counts.

Figure S3 Putative secondary structures of $tDNA^{Cys}$, $tDNA^{Gln}$, $tDNA^{His}$, $tDNA^{Leu}$, and $tDNA^{Pro}$. The putative secondary structures of $tDNA^{Cys}$, $tDNA^{Gln}$, $tDNA^{His}$, $tDNA^{Leu}$, and $tDNA^{Pro}$ are predicted by tRNAscan-SE 1.23 [36]. The red boxes indicate different regions of tDNA.

Figure S4 Phylogeny of the *Artemia* species based on the whole mitochondrial genome sequences.

The supporting information is available online at life.scichina.com and www.springerlink.com. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.