

Article

Elevation as a selective force on mitochondrial respiratory chain complexes of the *Phrynocephalus* lizards in the Tibetan plateau

Yuanting JIN^{a,*}, Débora Y. C. BRANDT^{b,**}, Jiasheng LI^a, Yubin WO^a, Haojie TONG^a and Vladimir SHCHUR^c

^aCollege of Life Science, China Jiliang University, Hangzhou, 310018, China, ^bDepartment of Integrative Biology, University of California at Berkeley, Berkeley, CA, 94720-3140, USA and ^cInternational Laboratory of Statistical and Computational Genomics, National Research University Higher School of Economics, Moscow, Russia

*Address correspondence to Yuanting Jin. E-mail: jinyuanting@126.com.

**These authors contributed equally to this work

Handling editor: Maria Servedio

Received on 22 April 2020; accepted on 10 September 2020

Abstract

Animals living in extremely high elevations have to adapt to low temperatures and low oxygen availability (hypoxia), but the underlying genetic mechanisms associated with these adaptations are still unclear. The mitochondrial respiratory chain can provide >95% of the ATP in animal cells, and its efficiency is influenced by temperature and oxygen availability. Therefore, the respiratory chain complexes (RCCs) could be important molecular targets for positive selection associated with respiratory adaptation in high-altitude environments. Here, we investigated positive selection in 5 RCCs and their assembly factors by analyzing sequences of 106 genes obtained through RNA-seq of all 15 Chinese *Phrynocephalus* lizard species, which are distributed from lowlands to the Tibetan plateau (average elevation >4,500 m). Our results indicate that evidence of positive selection on RCC genes is not significantly different from assembly factors, and we found no difference in selective pressures among the 5 complexes. We specifically looked for positive selection in lineages where changes in habitat elevation happened. The group of lineages evolving from low to high altitude show stronger signals of positive selection than lineages evolving from high to low elevations. Lineages evolving from low to high elevation also have more shared codons under positive selection, though the changes are not equivalent at the amino acid level. This study advances our understanding of the genetic basis of animal respiratory metabolism evolution in extreme high environments and provides candidate genes for further confirmation with functional analyses.

Key words: elevation, positive selection, reptile, plateau, respiratory chain complexes

High elevation presents challenges to animal physiology due to lower temperatures and oxygen availability (Hayes and S. O'Connor 1999; Rezende and Dinizfilho 2012; Cai et al. 2013). More specifically, both of these limitations impose pressures on the respiratory metabolism of animals (Lachance and Tishkoff 2013; Lim et al. 2019).

The mitochondrial respiratory chain is an important step of respiratory metabolism, providing over 95% of the ATP of vertebrates (Acuña-Castroviejo et al. 2001). Therefore, the mitochondrial respiratory chain complexes (RCCs) are potential targets of natural selection in animals living in extremely high environments.

The mitochondrial respiratory chain consists of 5 protein complexes (I-V). These complexes are composed of 13 proteins encoded in the mitochondrial DNA (mtDNA) and 84 proteins encoded in nuclear DNA (nDNA; [Supplementary Table S1](#), also see in HUGO Gene Nomenclature Committee (HGNC) database: www.genenames.org). Furthermore, there is an increasing number of other nDNA encoded proteins that are known to participate in assembling these complexes (HGNC database: www.genenames.org). Recent explorations of the crystal structure of the RCCs revealed which subunits (or regions) play key functions, such as the proton or electron translocation channel ([Fiedorczuk et al. 2016](#); [Letts et al. 2016](#); [Lobo-Jarne and Ugalde 2018](#)). Most of the mtDNA-encoded (7/13) proteins cluster as a hydrophobic region within Complex I ([Supplementary Table S1](#)).

In the past 15–20 years, a large number of studies investigated natural selection in the 13 mtDNA encoded genes of the RCC in species adapted to high elevation ([Zink 2005](#); [Mishmar et al. 2006](#); [Jin et al. 2018](#)). These studies showed evidence of different selective pressures on mitochondrial genes among populations or closely-related species ([Schmidt et al. 2001](#); [Luo et al. 2003](#); [Hassanin et al. 2009](#); [Yu et al. 2011](#); [Li et al. 2013](#); [Camus et al. 2017](#); [Yuan et al. 2018](#)). However, the study of correlations between habitat elevation and genetic evidence of natural selection (e.g., the number of codons under positive selection) has been hampered by the fact that only a few codons (usually fewer than 10) of the mtDNA encoded RCC genes are positively selected. Moreover, there is little statistical power to compare the strength of positive selection among the 5 RCCs because a minority of their proteins are encoded in mtDNA (13/97), and they are not evenly distributed across complexes (e.g., 7 out of 13 mtDNA encoded proteins are in Complex I, and there are no mtDNA-encoded proteins in Complex II).

Phrynocephalus is an important group of lizards distributed in a wide range in western China. It is divided into alpine viviparous species into cold and hypoxic arid or semiarid areas of the Qinghai-Tibetan Plateau (QTP) and oviparous species typically in lowland sand dunes, deserts, or areas of the Gobi desert outside of the QTP ([Zhao 1999](#)). The group likely originated in lowlands, a clade diverged ~10 million years ago and occupied highlands, and both the lowlands and the highlands clades further diversified into different species ([Jin and Brown 2013](#); [Solovyeva et al. 2018](#)). The maximum elevation difference between species in the viviparous and oviparous groups is near 5,000 m, and the elevation range remains large (~2,307–5,040 m) among viviparous species within the QTP and ~60–1,300 m for oviparous species outside of QTP ([Jin et al. 2008, 2013](#); [Jin and Brown 2013](#)). This group of closely-related species evolving in different elevation gradients, geographically isolated from each other by mountain barriers with little overlapping area ([Figure 1](#)), provides a suitable case to test positive selection during evolutionary history of these species.

We have recently identified 4 codons in mtDNA-encoded respiratory chain genes under long-term positive selection (i.e., under selection in the whole tree topology), and 6 codons under short-term positive selection analyzed by branch-site models (i.e., under branch-specific selection) in *Phrynocephalus* lizards ([Jin et al. 2018](#)). We found no overlap among the 6 codons under branch-specific positive selection (i.e., different codons were under selection in each branch). Interestingly, 2 of the codons under branch-specific selection correspond to amino-acid residues responsible for proton translocation. Although previous studies have found interesting potential targets of selection within respiratory chain genes, they have not investigated how selective pressures on the respiratory chain vary

among different protein complexes and among *Phrynocephalus* lineages living at different elevations. To increase the power of statistical tests comparing selection among groups of genes and among environments (altitude, in the case of this study), it is important to include nDNA-coded genes, since most of the proteins in the RCCs are encoded in the nucleus.

To date, many studies have shown that mitochondrial genes could be positively selected along environmental gradients ([Fonseca et al. 2008](#); [Foote et al. 2011](#); [Zhou et al. 2014](#); [Ngatia et al. 2019](#)). Recently, several studies using the 13 RCC genes of mitochondrial genomes indicated that altitudinal or longitudinal associated selection occurred in different animal groups ([Luo et al. 2003](#); [Mishmar et al. 2006](#); [Fonseca et al. 2008](#); [Jin et al. 2018](#)). Some of these studies indicated that Complex I was under stronger selective pressures than other complexes ([Mishmar et al. 2006](#); [Yu et al. 2011](#); [Jin et al. 2018](#)). However, this hypothesis could never be formally tested only using mitochondrial genes due to the following reasons: 1) too few candidate genes (only 13) are encoded in the mitogenome and they are not distributed evenly among the 5 RCCs, for example, there is no mtDNA encoded protein within Complex II, and most of the mtDNA encoded proteins (7 among 13) were exclusively clustered as a hydrophobic region of Complex I; 2) too few selected codons were discovered, statistically impeding a comparison of selective pressure among branches.

To overcome these limitations, we obtained the protein coding sequence (CDS) of nuclear genes representing the 5 mitochondrial RCCs of all Chinese *Phrynocephalus* lizards, as well as complex assembly factors, which are functionally closely related with the RCCs themselves. With these data, we address the following questions: 1) Are any of the RCCs under stronger selective pressure than others? 2) Do species evolving from low to high elevations show stronger signals of natural selection acting on their respiratory chain genes than species that have not undergone such dramatic elevation changes in their evolutionary history? 3) Do species evolving from low to high elevation show patterns of parallel evolution at either the codon or gene level? This study will help elucidate the genetic basis of adaptation of the mitochondrial RCCs of animals in response to elevation changes.

Materials and Methods

Specimens

A total of 15 *Phrynocephalus* specimens collected in 2016–2017 from elevations ranging from 602 m to 4,541 m were used in this study ([Supplementary Table S2](#)). All 6 viviparous species distributed in the QTP and 9 Chinese oviparous species/lineages living outside of the QTP were included. Equal amounts of total RNA extracted from each of 5 tissues (brain, liver, muscle, heart, and lung) of each specimen were combined for RNA-seq libraries. These species provided a comprehensive coverage of deep evolutionary lineages in this group. All experiments were performed in accordance with guidelines from the China Council on Animal Care and were approved by the Ethics Committee for Animal Experiments at China Jiliang University.

RNA sequencing

A total of 3 μ g RNA per sample was used as input material for the RNA sample preparations. Sequencing libraries were generated using NEBNext[®] UltraTM RNA Library Prep Kit for Illumina[®] (Illumina, San Diego, CA, USA) following manufacturer's

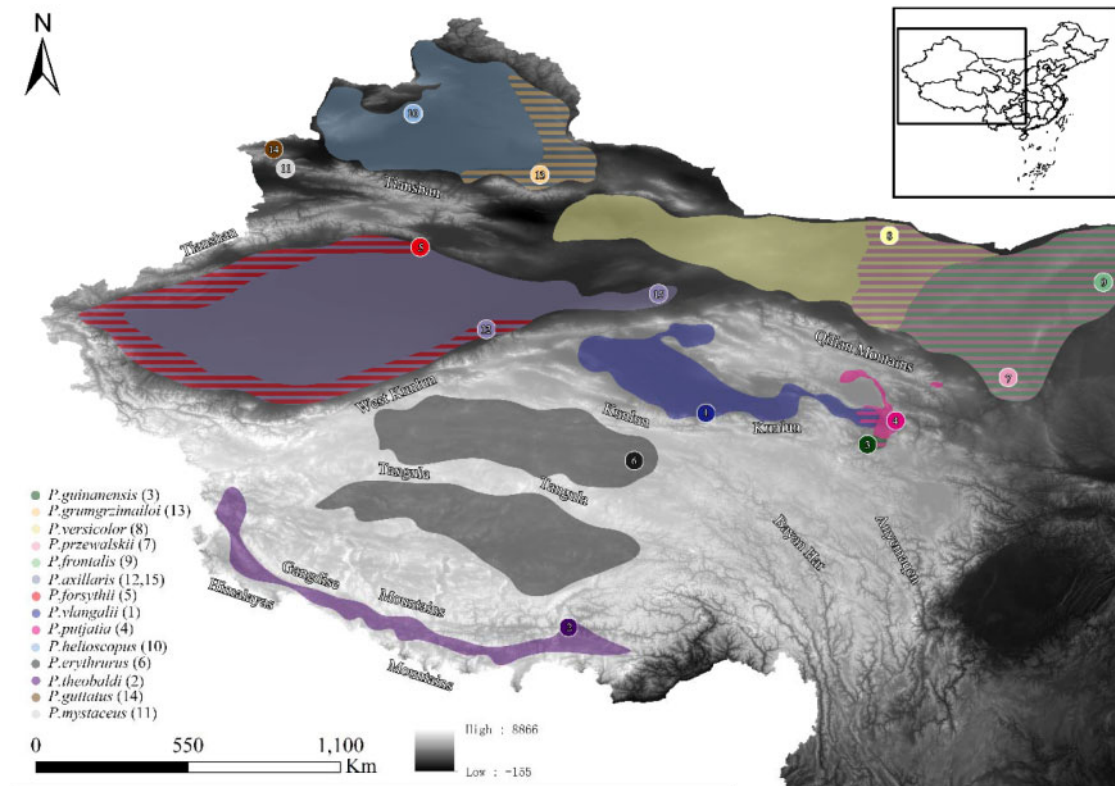


Figure 1. Sampling map of 15 *Phrynocephalus* specimens. Locality of sampled specimens is shown by numbers corresponding to the ones listed in [Supplementary Table S2](#). Species ranges are indicated by solid color when they are not overlapping, and the regions where species ranges overlap are shown as horizontal color stripes. The map was provided by Data Center for Resources and Environmental Sciences, the Chinese Academy of Sciences (<http://www.resdc.cn/>), and edited with ArcGIS 9.3.

recommendations and index codes were added to each sample. Briefly, mRNA was purified from total RNA using poly-T oligo-attached magnetic beads. Fragmentation was carried out using divalent cations under elevated temperature in NEBNext First Strand Synthesis Reaction Buffer (5X). First-strand cDNA was synthesized using random hexamer primer and M-MuLV Reverse Transcriptase (RNase H-). Second strand cDNA synthesis was subsequently performed using DNA Polymerase I and RNase H. Remaining overhangs were converted into blunt ends via exonuclease/polymerase activities. After adenylation of 3' ends of DNA fragments, NEBNext Adaptor with hairpin loop structure was ligated to prepare for hybridization. In order to select cDNA fragments of preferentially 250~300 bp in length, the library fragments were purified with AMPure XP system (Beckman Coulter, Beverly, MA, USA). Then 3 μ L USER Enzyme (NEB, Ipswich, MA, USA) was used with size-selected, adaptor-ligated cDNA at 37°C for 15 min followed by 5 min at 95°C before PCR. Then PCR was performed with Phusion High-Fidelity DNA polymerase, Universal PCR primers, and Index (X) Primer. At last, PCR products were purified (AMPure XP system) and library quality was assessed on the Agilent Bioanalyzer 2100 system.

Libraries were sequenced on an Illumina X Ten platform and 150 bp paired-end reads were generated. A total of around 10 Gb RNA-seq data per specimen were obtained.

Transcriptome assembly and quality control

Raw reads in fastq format were filtered by removing reads containing adapter sequences, reads containing poly-N and low-quality

reads from raw data. De novo transcriptomes were assembled from each specimen using Trinity v2.5 ([Grabherr et al. 2011](#)) with parameters `-seqType fq -max_memory 60 G -CPU 20 -min_kmer_cov 3 -min_contig_length 300 -bfly_opts "-V 20."` We assessed quality of the transcriptome assembly using an in-house perl script modified from <https://github.com/vikas0633/perl/blob/master/CountFasta.pl>. Total transcriptome length ranged from 145 to 338 Mb, with a mean of 174 Mb, and N50 ranged from 2,217 to 2,536, with a mean of 2,320.

RCC genes and assembly factors

Sequences from 84 RCC genes and 41 complex assembly genes were recovered from 2 well-annotated reference genomes of reptile species (lizard: *Anolis carolinensis*; snake: *Ophiophagus hannah*). Transcriptome contigs were aligned to these reference sequences using blastx, and we selected the longest transcript from each *Phrynocephalus* species that aligned to each of the 125 genes to use in this study. We only used a single representative sequence per gene and per species. Therefore, we note that we did not have heterozygous sites in our data, which is a limitation of this study. Orthologous genes were aligned with the Kalign software and edited manually. Each transcript was searched on nonredundant protein sequence (nr) database by BlastX and all potential alternatively spliced exons were deleted from the alignment before performing selection analyses. Finally, we excluded from our analysis genes that were missing in 2 or more species.

Phylogenetic tree construction

The datasets were divided into 3 codon positions for phylogenetic analyses. Maximum likelihood (ML) analyses were carried out using RAxML ver. V8.2.10 on a Linux Debian operating system. The `raxmlHPC-PTHREADS` command was used for the analyses and the GTR + G model was specified for each partition (-m command). An ML tree was identified for the original sequence and additional trees were determined for 2,000 bootstrap-resampled datasets to indicate levels of support for branches (-# option). No outgroup was specified and so the ML tree was rooted with a midpoint root for display purposes.

Selection analyses

Selection analyses were conducted on the (unrooted) nDNA tree that we inferred, and start and stop codons were not included. We used the CODEML program in PAML version 4.8, to detect signatures of selection per codon, and per codon and branch.

Evidence of selection per codon

In PAML version 4.8 (Yang 2007), likelihoods were obtained under M7 and M8 site models, where the ratio of nonsynonymous to synonymous substitutions (ω) follows a beta distribution (M7) and with an additional class of sites under positive selection ($\omega > 1$) (M8). A complete description of the models is available in the PAML manual (<http://abacus.gene.ucl.ac.uk/software/pamlDOC.pdf>). Likelihood ratio tests (LRTs) were used to compare both models. Sites were considered to have experienced positive selection when the LRT of corresponding genes were significant for M7-M8 comparisons. Posterior probabilities for site classes were obtained using a Bayes Empirical Bayesian (BEB) approach (Yang et al. 2005). This allowed identification of positively selected sites under the M8 models. We considered that there was strong evidence for a coding site being under positive selection when the BEB posterior probability was >0.95 and $\omega > 1$. This represents a stringent requirement given that ω is unlikely to average >1 across the entire group.

Selective pressures among different groups of genes (assembly factors and Complexes I–V) were compared using the likelihood ratios of individual genes under the site-model M7-M8.

Evidence of selection per codon, per branch

In the PAML (Yang 2007) branch-site model known as Model A, the branches are divided into foreground and background groups. This model provided more conservative LRT results than those of the branch-based test model. In the foreground lineages, a proportion of the sites are under positive selection ($\omega > 1$), whereas the background lineages feature a class of conserved sites with ω between 0 and 1, and a class of neutral sites with ω fixed at 1. Because it was difficult to select foreground branches a priori, we defined each possible branch as a foreground branch and then tested whether positive selection was present. Selection was tested by comparison of 2 models: null Model A, in which the sites in the foreground lineages are set at $\omega = 1$, against the alternative Model A, in which $\omega > 1$. We used the BEB method to obtain posterior probabilities. The instantaneous rate change of codons i to j depends on parameter π_j (equilibrium frequency of codon j), which was estimated using nucleotide frequencies observed at each of the 3 codon positions within the gene (F3 \times 4 model).

Correlation between number of codons under selection and elevation change in branches

The relationship between selective pressures and elevation changes on different branches was quantified by partial correlation between the number of positively selected sites in each branch and the change in elevation along that branch, whereas controlling for branch length. The number of positively selected sites was calculated as the proportion of positively selected sites multiplied by the amino acid length of each gene. Elevation of each ancestral node was estimated using phytools R package based on the estimated tree (see Figure 2) and the known sampling elevation of each specimen (Supplementary Table S2). Elevation of ancestral node was used to calculate the elevation change along each branch.

Parallel evolution in branches evolving from low to high elevation

We also investigated whether branches evolving from low to high elevation tend to show evidence of selection on the same sites (codons or genes), compared with lineages evolving from high to low elevation. We tested for selection at both gene and codon levels. For this analysis, we only used branches in which at least 1 site (codon or gene) was under positive selection. The criterion for a site to be considered to be under positive selection was that it had significant lnL ratio values and omega values larger than 1. For codon sites, we additionally required BEB posterior probability higher than 0.95. For the codon sites, this included a total of 16 branches, including 8 branches evolving from low to high and 8 branches evolving from high to low elevations. For gene sites, we included 8 branches evolving from low to high and 11 branches evolving from high to low elevations (Figure 2).

We tested whether there is a stronger overlap between sites under selection in branches that transitioned to higher altitude than in other branches. We call the branches associated with transitions towards higher altitude (typically for viviparous species in QTP), “target branches.” More specifically, we test the hypothesis that if a site is under selection in 1 target branch, then it is more likely to be under selection in other target branches, more so than sites that are under selection in nontarget branches.

We represented our data as a matrix (m_{si}) with rows corresponding to gene or codon sites and with columns corresponding to branches in the phylogeny. The matrix entries are binary with 1 indicating positive selection and 0 indicating no selection. At each site, s , we take all possible pairs of target branches (i, j) (those evolving from low to high elevations) and count in how many pairs the site is under positive selection in both branches ($m_{si} = m_{sj} = 1$). We define the test statistic γ as the sum of this count over all sites. Formally, in matrix notation

$$\gamma = \sum \{(i, j) : i \text{ and } j \text{ are target branches and } m_{si} = m_{sj} = 1\}.$$

We generated k random permutations of the matrix under the independence assumptions with fixed marginals. We performed permutations using the `permat` package in R (<https://www.rdocumentation.org/packages/vegan/versions/2.4-2/topics/permat>). Three permutation methods (`swap`, `tswap`, and `quasiswap`) were applied (Patefield 1981; Hardy 2008). We condition on the marginals of the matrix in order to take into account the variation in the number of sites or genes under selection among branches and the variation in the number of branches under selection among sites or genes. For each of the permuted matrices, the test statistic γ was computed. We then compared the observed test statistic to the null

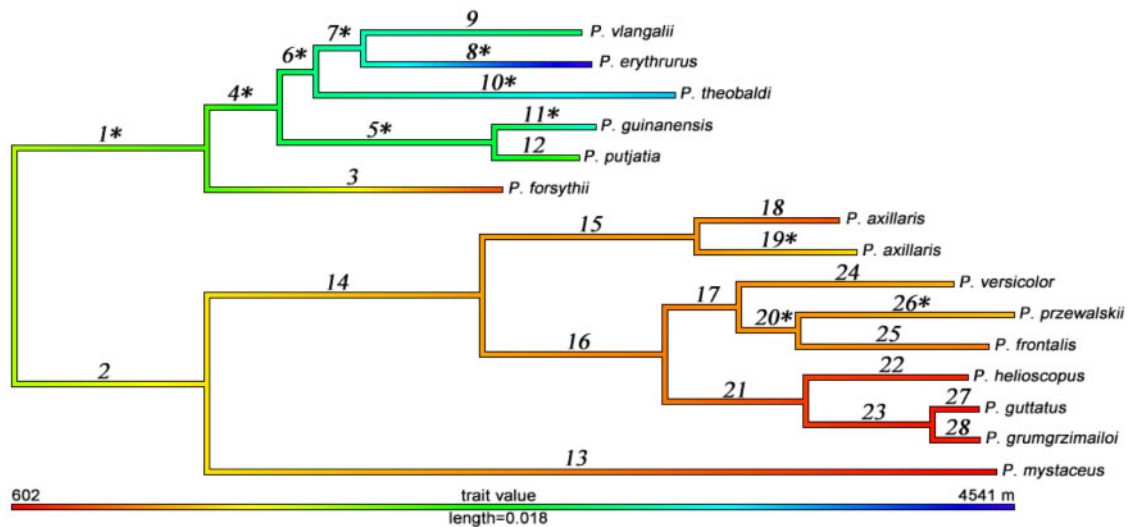


Figure 2. Estimated habitat elevation of species in the *Phrynocephalus* genus (color scale shows altitudes ranging from 602 to 4,541 m). Colors at the tips represent elevation at sampling locations, and elevation at ancestral nodes was inferred as described in “Materials and Methods” section. Branches are numbered, and asterisks indicate 11 branches (No. 1, 4, 5, 6, 7, 8, 10, 11, 19, 20, 26) representing species that evolved from low- to high-elevation habitats. All species descending from branch No. 1 are viviparous, and all species descending from branch No. 2 are oviparous. Delta elevational changes of 28 branches were listed within [Supplementary Table S4](#).

distribution obtained for the k permutations. A 1-tailed test was calculated as the proportion of times the observed value of the test statistic γ is smaller than the test statistic γ of the permuted matrix. We ran every permutation algorithm for 1 million iterations with 100,000 burn-in iterations.

Results

Genes and phylogenetics

After aligning the assembled transcriptomes to 125 RCC gene sequences and deleting the alternative splicing regions (as described in the Materials and Methods), we selected a total of 106 genes having 70,341 bp in combined length, with each gene varying from 87 to 2,526 bp. Detailed information about analyzed genes is listed in [Supplementary Table S3](#).

Phylogenetic analyses using RAxML v8.2.X indicated 2 major clades represented, respectively, by all viviparous or oviparous species ([Figure 3](#)), which is consistent with previously published topologies based mainly on mtDNA datasets ([Jin and Brown 2013](#)). Highest bootstrap values (100) were obtained for all major nodes ([Figure 3](#)).

Comparison of positive selection among complexes and genes

Comparisons of selective pressures among genes in different complexes were made based on the log-likelihood ratios (LLRs) of genes estimated using the M7-M8 model. A high lnL ratio in the M7-M8 model indicates evidence of selection. The log-likelihood (lnL) ratios of all genes are listed in [Supplementary Table S3](#). Using a nonparametric test, we found marginally nonsignificant difference in positive selection pressure between complexes and complex assembly factors (Wilcoxon test $W = 1,505$, $P = 0.1029$, [Figure 4A](#)), with respiratory complex experiencing more selection than complex assembly factors. However, no significant difference among the 5 complexes was found (Kruskal–Wallis $H = 3.3484$, $df = 4$, $P = 0.5013$, [Figure 4B](#)). We point out that the detection of differences in selective pressure

among complexes has the caveat that only 4 genes for Complex II were included, and it does not show evidence of selection. No significant differences were found among Complexes I, III, IV, and V when the test was repeated including those groups only (Kruskal–Wallis $H = 1.2839$, $df = 3$, $P = 0.733$). The 5 genes with the strongest evidence of selection as measured by their large likelihood ratios (>20) are NDUF57 (which is a subunit of Complex I), COX6C (Complex IV), COX5B (Complex IV), NDUF2 (Complex I), and ATP5PD (Complex V).

Positive selection and parallel evolution of branches among elevation gradients

We found a positive correlation between the number of positively selected codons and the absolute change in elevation among 28 branches when average branch length was controlled for (Pearson partial correlation: $r = 0.522$, $P = 0.005$, $df = 25$, [Figure 5A](#)). The correlation remained significant or marginally nonsignificant when we analyzed separately the branches evolving from low to high and the branches from high to low elevation (branches from high to low: $r = -0.635$, $P = 0.036$, $df = 9$, [Figure 5B](#); branches from low to high elevations: $r = 0.506$, $P = 0.054$, $df = 13$, [Figure 5C](#)). This indicates that branches with larger altitudinal changes underwent stronger positive selection. We note that some correlations are marginal and could not be replicated when the strength of selection was measured as the sum of the LLRs of selection in all sites in a branch ([Figure 5D–F](#)). Detailed data including delta elevations, proportions of positive selection codons, sums of likelihood ratios for 106 genes, and mean branch lengths used for correlation analyses were listed in [Supplementary Table S4](#).

Next, we looked for signals of parallel evolution among branches evolving from low to high elevation. Specifically, we used a permutation-based test to determine whether similar codons or genes were showing signals of selection on these branches. For this analysis, we used only branches in which at least 1 site (codon or gene) was under positive selection. The codon and gene-branch matrices used for permutation analyses are listed as [Supplementary](#)

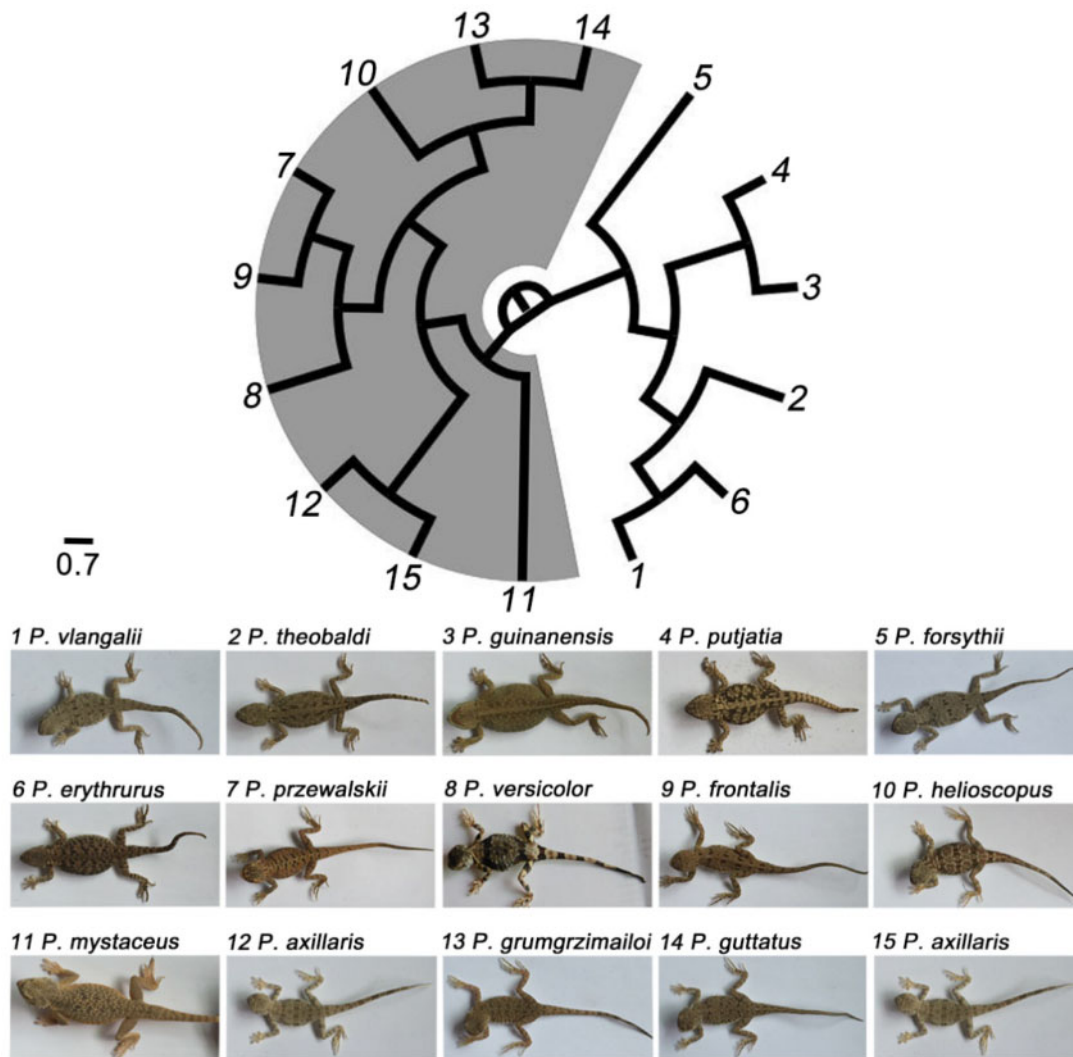


Figure 3. RAXML phylogenetic tree estimated using sequences of 106 RCCs and complex assembling factors genes. The gray shaded clades show the species from low-elevation areas, and the other clades live in high elevations on the Tibetan plateau. Highest bootstrap values (100) were estimated for all nodes.

Tables S5 and S6 (see also Methods for more details on the permutation test).

Tests were done using 3 different matrix permutation methods (quasiswap, swap, and tswap) supported that the set of target branches evolving from low to high elevations share more codons under selection than branches evolving from high to low (P -values were zero in all 3 permutation methods). We did not observe this signal at the gene level (P -values were above 0.22 for all 3 permutation methods).

Among the codons experiencing positive selection in more than 1 branch evolving from low to high elevation, only one repeated amino acid state change was discovered, a change from “Ala” to “Gly” in NDUFS7 gene (Supplementary Table S7).

Discussion

This study investigated selection on most of the known nuclear genes composing the RCCs as well as the complex assembly factors. We showed that there is no significant difference in selective pressure among complexes, nor between complex subunits and assembly

factors. Therefore, our results do not support previous studies (Mishmar et al. 2006; Yu et al. 2011; Jin et al. 2018) using only mitochondrial genes, which found higher selection pressure on Complex I than on other complexes of *Phrynocephalus*. Our study is in line with previous studies that showed evidence for positive selection and climate-linked differences in the sequences and expression of RCC genes in a range of animal species with wide biogeographic ranges (Toews et al. 2014; Garvin et al. 2015; Morales et al. 2018).

Results of partial correlation analyses showed that branches evolving under larger elevation changes tend to experience stronger selective pressure, when selective pressure is measured as the number of positively selected codons. Another proxy for strength of selection along a branch is the sum of the LLRs comparing models with (M8) and without selection (M7). We did not find a significant correlation between strength of selection measured by the sum of the LLR and elevation change along a branch. However, our results raise the interesting hypothesis that species that went through larger habitat altitudinal changes in their evolutionary history experience higher levels of selective pressure on RCC genes. This hypothesis could be further tested with more data from *Phrynocephalus* or other species

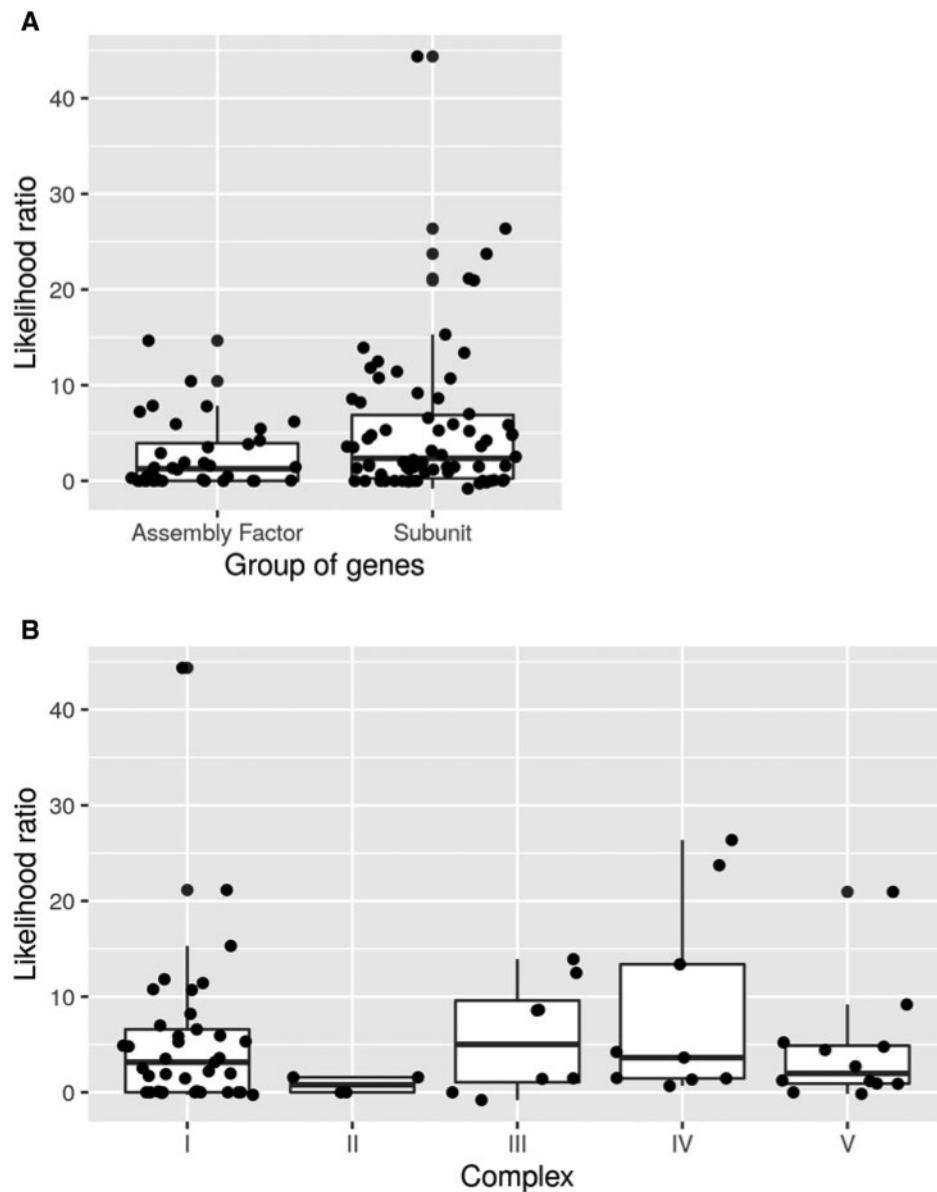


Figure 4. LLRs (lnLR) of selection in the RCC main genes (subunits) and in the RCC assembly factor genes. **(A)** Evidence of positive selection is higher on genes composing the complexes subunits than on complex assembly factors (Wilcoxon test $W = 1,505$, $P = 0.1029$). **(B)** There is no evidence of difference in positive selection among the different complexes (Kruskal–Wallis $H = 3.3484$, $P = 0.5013$).

living in extremely high environments, as well as in other genes that could be under selection due to altitudinal change.

Our codon-branch permutation analyses suggest that branches representing evolving from low to high altitude had more frequent selection on the same codons than branches emerged from high to low altitude. *Phrynocephalus* originated from the lowlands of central Asia and a viviparous subclade (see Figure 3) emerged in QTP along with the uplift of the QTP, 10 million years ago. Our results are consistent with the idea that as *Phrynocephalus* lizards occupied high altitudes, this new environment imposed selection on similar codons of RCCs.

We found only one instance of parallel evolution in amino acid state, in the gene *NDUFS7*. This gene is differentially expressed in pikas at different altitudes in the Himalayas, suggesting a role for altitude adaptation in other groups of animals (Solarí et al. 2018). We speculate that parallel evolution may be rare due to simple

chance, that is, if a certain allele was not available from either shared ancestral variation or recurrent mutation. It could also be rare due to environmental tradeoffs, which could cause an allele to be favorable in one of the high-altitude environments, but not the other. It is also possible that parallel evolution at the codon level is rare, because functional convergence can occur with selection in different genes or codons.

Understanding the genetic adaptation of animals living in extremely high-altitude environments of the QTP has been receiving increasing attention recently (Sun et al. 2018). Here, we analyzed most of the nuclear RCC genes and showed evidence that species living in higher altitudes could undergo stronger positive selection affecting the respiratory chain than closely-related species living in lower altitudes. However, the functional consequences of the genetic adaptation of RCCs showed in this study remain untested. Although studies of metabolism of reptile species living at high altitude remain

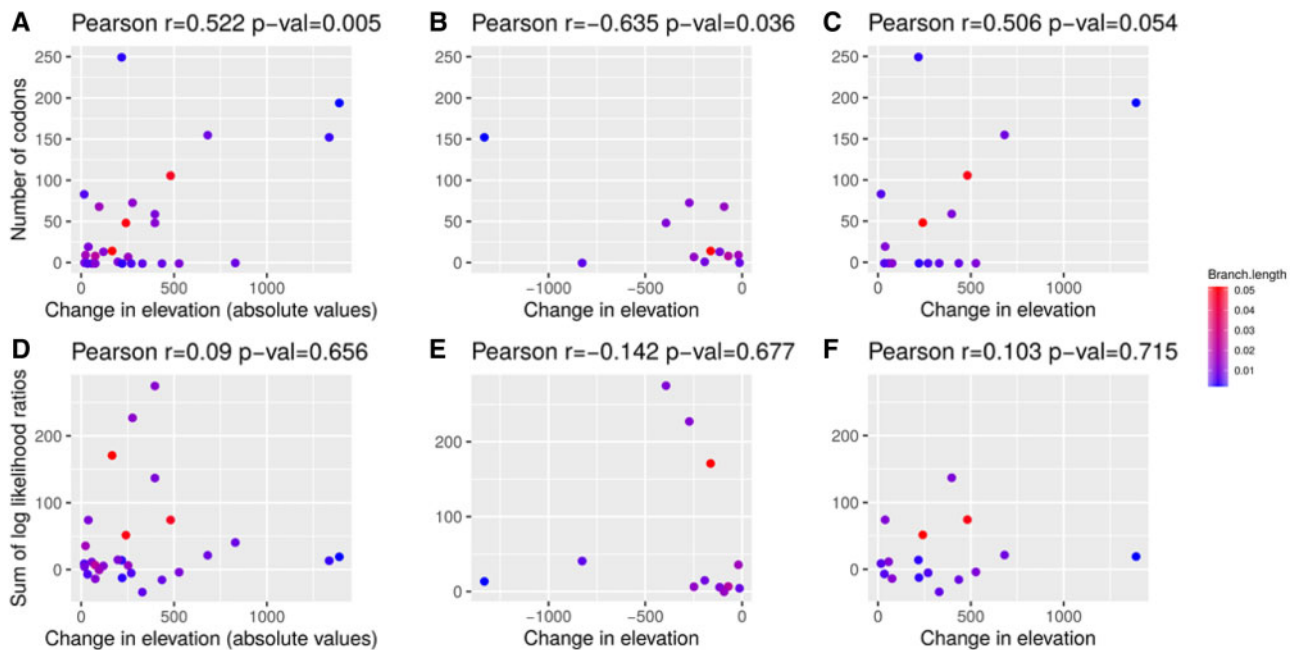


Figure 5. Partial correlation between change in elevation and strength of positive selection, controlling for the effect of branch length. Partial correlation was significant when strength of selection was measured as number of codons under selection (A, B, C), but was not significant when selection was measured as sums of the LLRs of the selections tests (D, E, F). (A, D) All branches included, absolute value of variation in elevation was used. (B, E) Only branches evolving from high to low elevation included (negative variation in elevation). (C, F) Only branches evolving from low to high elevation included (positive variation in elevation).

limited, a recent study indicated lower respiratory rate and higher ATP production efficiency of mitochondria in *P. erythrurus* from high elevation areas of the QTP than that of *P. przewalskii* from low elevations outside of the QTP, when both species were maintained at pressures of oxygen similar to that of their respective habitats (Tang et al. 2013). These could be functional consequences of the positively selected codon changes discovered in this study.

In summary, our research shows that elevation leads to important differences in selection on mitochondrial RCCs of *Phrynocephalus* lizards from the Tibetan plateau. Groups that have colonized high from lower elevations appear to have experienced stronger selective forces than those that have moved in the opposite direction. Increased selection seems to be associated with greater shifts in elevation. Finally, selection pressures are similar among genes encoding RCC subunits and RCC assembly factors.

Acknowledgments

We wish to thank Kailong Zhang for help of specimen collection. We are especially grateful to Rasmus Nielsen in University of California, Berkeley, for his kind suggestions on the analyses of this work.

Funding

This work was supported by the National Natural Science Foundation of China (Grant No. 31772447) and the Second Tibetan Plateau Scientific Expedition and Research Program (STEP), Grant No. 2019 QZKK0501. V.S. did this research within the framework of the HSE University Basic Research Program.

Data Availability Statement

Data on geographical sampling locations and all measured dorsal and substrate color are deposited in FigShare (doi:10.6084/m9.figshare.10308572). The corresponding authors are responsible for any personal requirement for the materials.

Supplementary Material

Supplementary material can be found at <https://academic.oup.com/cz>.

References

- Acuña-Castroviejo D, Martin M, Macias M, Escames G, Leon J et al., 2001. Melatonin, mitochondria, and cellular bioenergetics. *J Pineal Res* 30:65–74.
- Cai Q, Qian X, Lang Y, Luo Y, Xu J et al., 2013. Genome sequence of ground tit *pseudopodoces humilis* and its adaptation to high altitude. *Genome Biol* 14:R29.
- Camus MF, Wolff JN, Sgrò CM, Dowling DK, 2017. Experimental support that natural selection has shaped the latitudinal distribution of mitochondrial haplotypes in Australian *Drosophila melanogaster*. *Mol Biol Evol* 34:2600–2612.
- Fiedorczuk K, Letts J, Degliesposti G, Kaszuba K, Skehel M et al., 2016. Atomic structure of the entire mammalian mitochondrial complex I. *Nature* 538:406–410.
- Fonseca RRD, Johnson WE, Obrien SJ, Ramos MJ, Antunes A, 2008. The adaptive evolution of the mammalian mitochondrial genome. *BMC Genom* 9:119–119.
- Foote AD, Morin PA, Durban JW, Pitman RL, Wade PR et al., 2011. Positive selection on the killer whale mitogenome. *Biol Lett* 7:116–118.
- Garvin MR, Thorgaard GH, Narum SR, 2015. Differential expression of genes that control respiration contribute to thermal adaptation in redband trout (*Oncorhynchus mykiss gairdneri*). *Genome Biol Evol* 7:1404–1414.
- Grabherr M, Haas BJ, Yassour M, Levin JZ, Thompson DA et al., 2011. Full-length transcriptome assembly from RNA-seq data without a reference genome. *Nat Biotechnol* 29:644–652.
- Hardy OJ, 2008. Testing the spatial phylogenetic structure of local communities: statistical performances of different null models and test statistics on a locally neutral community. *J Ecol* 96:914–926.
- Hassanin A, Ropiquet A, Couloux A, Cruaud C, 2009. Evolution of the mitochondrial genome in mammals living at high altitude: new insights from a study of the tribe *Caprini* (Bovidae, Antilopinae). *J Mol Evol* 68:293–310.
- Hayes J, O'Connor CS, 1999. Natural selection on thermogenic capacity of high-altitude deer mice. *Evolution* 53:1280–1287.

- Jin YT, Brown RP, 2013. Species history and divergence times of viviparous and oviparous Chinese toad-headed sand lizards *Phrynocephalus* on the Qinghai-Tibetan Plateau. *Mol Phylogenet Evol* 68:259–268.
- Jin YT, Brown RP, Liu NF, 2008. Cladogenesis and phylogeography of the lizard *Phrynocephalus vlangalii* (agamidae) on the Tibetan plateau. *Mol Ecol* 17:1971–1982.
- Jin YT, Li JQ, Liu NF, 2013. Elevation-related variation in life history traits among *Phrynocephalus* lineages on the tibetan plateau: do they follow typical squamate ecogeographic patterns?. *J Zool* 290:293–301.
- Jin Y, Wo Y, Tong H, Song S, Zhang L et al., 2018. Evolutionary analysis of mitochondrially encoded proteins of toad-headed lizards *Phrynocephalus* along an altitudinal gradient. *BMC Genom* 19:185.
- Lachance J, Tishkoff SA, 2013. Population genomics of human adaptation. *Annu Rev Ecol Evol Syst* 44:123–143.
- Letts J, Fiedorczuk K, Sazanov L, 2016. The architecture of respiratory super-complexes. *Nature* 537:644–648.
- Li Y, Ren Z, Shedlock AM, Wu J, Sang L et al., 2013. High altitude adaptation of the schizothoracine fishes (Cyprinidae) revealed by the mitochondrial genome analyses. *Gene* 517:169–178.
- Lim MCW, Witt CC, Graham CH, Dávalos LM, 2019. Parallel molecular evolution in pathways, genes, and sites in high-elevation hummingbirds revealed by comparative transcriptomics. *Genome Biol Evol* 11:1573–1585.
- Lobo-Jarne T, Ugalde C, 2018. Respiratory chain supercomplexes: structures, function and biogenesis. *Semin Cell Dev Biol* 76:179–190.
- Luo Y, Chen Y, Liu F, Yuqi G, 2003. Mitochondrial genome of Tibetan wild ass *Equus kiang* reveals substitutions in nadh which may reflect evolutionary adaptation to cold and hypoxic conditions. *Asia Life Sci* 21:1–11.
- Mishmar D, Ruiz-Esini E, Mondragon-Palomino M, Procaccio V, Gaut B et al., 2006. Adaptive selection of mitochondrial complex I subunits during primate radiation. *Gene* 378:11–18.
- Morales HE, Pavlova A, Amos N, Major R, Kilian A et al. 2018. Concordant divergence of mitogenomes and a mitonuclear gene cluster in bird lineages inhabiting different climates. *Nature Ecology & Evolution* 2:1258–1267. doi:10.1038/s41559-018-0606-3.
- Ngatia JN, Lan TM, Dinh TD, Zhang L, Ahmed AK et al., 2019. Signals of positive selection in mitochondrial protein-coding genes of woolly mammoth: adaptation to extreme environments?. *Ecol Evol* 9:6821–6832.
- Patefield WM, 1981. Algorithm as 159: an efficient method of generating random $r \times c$ tables with given row and column totals. *J R Stat Soc* 30:91–97.
- Rezende EL, Dinizfilho JAF, 2012. Phylogenetic analyses: comparing species to infer adaptations and physiological mechanisms. *Compr Physiol* 2:639–674.
- Schmidt TR, Wu W, Goodman M, Grossman LI, 2001. Evolution of nuclear- and mitochondrial-encoded subunit interaction in cytochrome c oxidase. *Mol Biol Evol* 18:563–569.
- Solari K, Ramakrishnan U, Hadly EA, 2018. Gene expression is implicated in the ability of pikas to occupy himalayan elevational gradient. *PLoS ONE* 13:e0207936.
- Solovyeva EN, Lebedev VS, Dunayev EA, Nazarov RA, Bannikova AA et al., 2018. Cenozoic aridization in Central Eurasia shaped diversification of toad-headed agamas (*Phrynocephalus*; Agamidae, Reptilia). *PeerJ* 6:e4543.
- Sun Y, Fu T, Jin J, Murphy RW, Hillis DM et al., 2018. Species groups distributed across elevational gradients reveal convergent and continuous genetic adaptation to high elevations. *Proc Natl Acad Sci USA* 115:e10634–e10641.
- Tang X, Xin Y, Wang H, Li W, Zhang Y et al., 2013. Metabolic characteristics and response to high altitude in *Phrynocephalus erythrurus* (Lacertilia: Agamidae), a lizard dwell at altitudes higher than any other living lizards in the world. *PLoS ONE* 8:e71976.
- Toews DPL, Mandic M, Richards JG, Irwin DE, 2014. Migration, mitochondria, and the yellow-rumped warbler. *Evolution* 68:241–255.
- Yang Z, 2007. Paml 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol* 24:1586–1591.
- Yang Z, Wong WSW, Nielsen R, 2005. Bayes empirical bayes inference of amino acid sites under positive selection. *Mol Biol Evol* 22:1107–1118.
- Yu L, Wang X, Ting N, Zhang Y, 2011. Mitogenomic analysis of Chinese snub-nosed monkeys: evidence of positive selection in nadh dehydrogenase genes in high-altitude adaptation. *Mitochondrion* 11:497–503.
- Yuan ML, Zhang QL, Zhang L, Jia CL, Li XP et al., 2018. Mitochondrial phylogeny, divergence history and high-altitude adaptation of grassland caterpillars (Lepidoptera: Lymantriinae: *Gynaephora*) inhabiting the Tibetan Plateau. *Mol Phylogenet Evol* 122:116–124.
- Zhao KT, 1999. *Phrynocephalus* (Agamidae). In: Zhao EM, Zhao KT, Zhou KY, editors. *Fauna Sinica, Reptilia (Squamata: Lacertilia)*. Vol. 2. Beijing: Science Press. 151–193.
- Zhou T, Shen X, Irwin DM, Shen Y, Zhang Y, 2014. Mitogenomic analyses propose positive selection in mitochondrial genes for high-altitude adaptation in Galliform birds. *Mitochondrion* 18:70–75.
- Zink RM, 2005. Natural selection on mitochondrial DNA in parus and its relevance for phylogeographic studies. *Proc Biol Sci* 272:71–78.