Xiao et al. *Zool. Res.* 2021, 42(4): 502−513
https://doi.org/10.24272/j.issn.2095-8137.2021.096

ZR | Zoological Research

Article Open Access

Genome and population evolution and environmental adaptation of *Glyptosternon maculatum* on the Qinghai-Tibet Plateau

Shi-Jun Xiao[1,2,5,6,#], Zen-Bo Mou[1,#], Rui-Bin Yang[3,#], Ding-Ding Fan[3], Jia-Qi Liu[2], Yu Zou[2], Shi-Lin Zhu[2], Ming Zou[2], Chao-Wei Zhou[1,4,*], Hai-Ping Liu[1,*]

[1] *Institute of Fisheries Science*, *Tibet Academy of Agricultural and Animal Husbandry Sciences*, *Lhasa*, *Tibet* 810000, *China*
[2] *Department of Computer Science*, *Wuhan University of Technology*, *Wuhan*, *Hubei* 430070, *China*
[3] *College of Fisheries*, *Huazhong Agricultural University*, *Wuhan*, *Hubei* 430070, *China*
[4] *Key Laboratory of Freshwater Fish Reproduction and Development (Ministry of Education)*, *College of Fisheries*, *Southwest University*, *Chongqing* 402400, *China*
[5] *College of Plant Protection*, *Jilin Agriculture University*, *Changchun*, *Jilin* 130118, *China*
[6] *Jiaxing Key Laboratory for New Germplasm Breeding of Economic Mycology*, *Jiaxing*, *Zhejiang* 314000, *China*

ABSTRACT

Persistent uplift means the Qinghai-Tibet Plateau (QTP) is an ideal natural laboratory to investigate genome evolution and adaptation within highland environments. However, how paleogeographic and paleoclimatic events influence the genome and population of endemic fish species remains unclear. *Glyptosternon maculatum* is an ancient endemic fish found on the QTP and the only critically endangered species in the Sisoridae family. Here, we found that major transposons in the *G. maculatum* genome showed episodic bursts, consistent with contemporaneous geological and climatic events during the QTP formation. Notably, histone genes showed significant expansion in the *G. maculatum* genome, which may be mediated by long interspersed nuclear elements (LINE) repetitive element duplications. Population analysis showed that ancestral *G. maculatum* populations experienced two significant depressions 2.6 million years ago (Mya) and 10 000 years ago, exhibiting excellent synchronization with Quaternary glaciation and the Younger Dryas, respectively. Thus, we propose that paleogeography and paleoclimate were dominating driving forces for population dynamics in endemic fish on the QTP. Tectonic movements and temperature fluctuation likely destroyed the habitat and disrupted the drainage connectivity among populations. These factors may have caused severe bottlenecks and limited migration among ancestral *G. maculatum* populations, resulting in the low genetic diversity and endangered status of the species today.

Received: 28 March 2021; Accepted: 09 July 2021; Online: 13 July 2021

Foundation items: This project was supported by the Key Research and Development Projects in Tibet: Preservation of Characteristic Biological Germplasm Resources and Utilization of Gene Technology in Tibet (XZ202001ZY0016N), National Natural Science Foundation of China (32072980), and Special Finance of Tibet Autonomous Region (XZNKY-2019-C-053)

[#]Authors contributed equally to this work

*Corresponding authors, E-mail: 442589528@qq.com; luihappying@163.com

Keywords: Qinghai-Tibet Plateau (QTP); *Glyptosternon maculatum*; Genome evolution; Population; High-altitude adaptation

INTRODUCTION

Deciphering how genomes evolve in response to dramatic environmental change is an essential question in evolutionary biology to understand the molecular mechanisms of adaptations and speciation. The Qinghai-Tibet Plateau (QTP) was formed from the collision of the Indian and Eurasian Plates and is the youngest, largest, and highest plateau on Earth, with an average altitude of over 4 000 m a.s.l. (Peng et al., 2006). Continued uplift over the last 50 million years ago (Mya) (Peng et al., 2006) and the extreme climatic environment make the QTP an ideal and unique natural laboratory to investigate how paleogeography and paleoclimate have influenced endemic genome and population evolution. The high-altitude environment of the QTP is characterized by high ultraviolet (UV) radiation, dramatic temperature changes, and nutritional deficiencies (Jiang et al., 2012; Zhang et al., 2016). Consequently, understanding genome and population evolution is essential to gain insight into the molecular mechanisms underlying environmental adaptations and to protect the wild genetic resources of endemic endangered species in the QTP.

Research has elucidated the molecular mechanisms that underlie the high-altitude adaptations of Tibetan people (Petousi & Robbins, 2014; Yi et al., 2010). However, the initial peopling of the QTP by modern humans occurred about 25 000 years ago (Beall, 2007), which is significantly more recent than the long history of several million years for QTP formation. Endemic animals of the plateau represent excellent genome evolution models with long adaptive histories during plateau uplift. In the last decade, studies on genome sequencing in Tibetan animals (Wang et al., 2014), including yaks (Qiu et al., 2012), pigs (Ai et al., 2014), dogs (Gou et al., 2014), antelopes (Ge et al., 2013), and birds (Qu et al., 2013), have broadened our understanding of genomic adaptations to the local environment via the evolution and selection of putatively important functional genes (Jablonski, 2017; Wei et al., 2016). However, few studies have uncovered genome characterizations of endemic animals. Moreover, the relationship between the evolution of endemic species genomes and populations and paleogeographic and paleoclimatic events during the formation of the QTP is still largely unknown.

The dispersal and distribution of fish strictly depend on drainage connections, which are highly influenced by tectonic movements and climate change (Yang et al., 2009). Thus, the genomic evolution of freshwater fish offers an excellent opportunity to contextualize biological evolution to changes in QTP geology and climate (Kang et al., 2017; Xiao et al., 2020; Yang et al., 2009). Although transcriptomic and genomic studies on *Schizothorax o'connori* (Xiao et al., 2020) and *Triplophysa bleekeri* (Yuan et al., 2020) have revealed differential expression and natural selection of key genes in QTP fish adaptations (Ma et al., 2016), research on fish genome and population evolution in relation to environmental change and uplift of the QTP remains limited. The family Sisoridae in the order Siluriformes, which contains glyptosternoid and non-glyptosternoid species, is one of the main taxonomic groups of fish in the basin waters of the QTP (Peng et al., 2006). The evolution and speciation of sisorid catfish are thought to have been profoundly influenced by QTP uplift (Guo et al., 2005; He et al., 2001). *Glyptosternon maculatum* is one of the most ancient species within Glyptosterninae and is distributed at altitudes from 2 800–4 200 m in the Yarlung Zangbo River (Zhang et al., 2010). Consequently, due to its peculiar ecological distribution and evolutionary history (Zhang et al., 2011), *G. maculatum* is an ideal species to explore genome and population evolution during the formation of the QTP and to evaluate how genomic changes contributed to the environmental adaptations of sisorid catfish on the QTP.

To this end, we performed *G. maculatum* genomic analysis and revealed episodic transposon bursts, which may have facilitated functional gene expansions and thus contributed to the speciation and environmental adaptation of *G. maculatum*. Population genetic analyses based on whole-genome variants were performed to address how contemporaneous environmental changes may have influenced ancestral populations. We revealed that both QTP uplift and climatic events likely imposed immense natural stress on *G. maculatum* and limited migration among populations, leading to the low genetic diversity and endangered status of *G. maculatum* today.

MATERIALS AND METHODS

Chromosomal construction and transcriptome assembly for phylogenetic analysis

We used Hi-C techniques for chromosomal assembly of *G. maculatum* based on contig sequences assembled from a previous study (Edgar, 2004). Muscle samples from the same *G. maculatum* fish used for genome sequencing were used for Hi-C library construction following prior research (Gong et al., 2018). The Hi-C libraries were controlled for quality and sequenced on the Illumina Hiseq X Ten platform (Illumina, USA) (Supplementary Table S1). The contig sequences of *G. maculatum* were scaffolded into chromosomes using the bioinformatics method in Gong et al. (2018) (Supplementary Table S2 and Figure S1).

Transcriptomes of muscle tissue from 16 QTP catfish (Supplementary Tables S3, S4), including 15 Sisoridae fish (*Exostoma labiatum*, *Pareuchiloglanis feae*, *Pareuchiloglanis kamengensis*, *Glyptothorax quadriocellatus*, *Glyptothorax fukiensis honghensis*, *Glyptothorax interspinalum*, *Glyptothorax cavia*, *Glyptothorax dorsalis*, *Glyptothorax laosensis*, *Glyptothorax zainaensis*, *Glyptothorax trilineatus*, *Glyptothorax minimaculatus*, *Bagarius yarrelli*, and *Pseudecheneis sulcatus*) and one Bagridae fish (*Leiocassis longirostris*), were sequenced using the Illumina HiSeq X Ten platform (Illumina, USA). In total, 66 Gb of raw RNA-seq data were generated. The RNA-seq reads of these fish were assembled by Trinity v2.8.6 (Kumar et al., 2016) (Supplementary Table S4).

Phylogenetic analysis among sisorid catfish

The unigenes of the 16 catfish were translated into protein sequences. Protein sequences longer than 50 amino acids were clustered with the whole proteome of *G. maculatum* using OrthoMCL v2.0.9 (Li, 2011). We obtained 167 single-copy genes with a total length of ~232 kb after gene family clustering. MUSCLE v3.8.31 (Li & Durbin, 2009) was used to generate multiple sequence alignments for protein sequences in each single-copy family with default parameters. The alignments of each family were concatenated to a super alignment matrix, which was used for phylogenetic tree reconstruction with RAxML v8 (Li et al., 2003). Divergence time between fish species was estimated using MCMCTree in Phylogenetic Analysis by Maximum Likelihood (PAML) v4 (Lieberman-Aiden et al., 2009), with the options "correlated molecular clock" and "JC69" model. Markov Chain Monte Carlo analysis was run for 20 000 generations using a burn-in of 1 000 iterations. Four calibration points were applied: *Glyptosternon* and *Exostoma* divergence time (5.5–8.8 Mya), *Glyptosternon* and *Glyptothorax* divergence time (7.7–12.2 Mya), *Glyptosternon* and *Pseudecheneis* divergence time (7.7–12.2 Mya), and Sisoridae and Bagridae divergence time (41–143 Mya). These calibration points were traced from TimeTree (http://www.timetree.org/) (Supplementary Figure S2).

Expanded gene families and positively selected gene detection

The phylogenetic tree and divergence times of *G. maculatum* and other species were analyzed using CAFE v5 (De Bie et al., 2006) to infer changes in gene family size using a probabilistic model. Gene Ontology (GO) enrichment analysis was performed on the *G. maculatum* expanded genes using the topGO v2.36.0 package (Alexa et al., 2006). MUSCLE v3.8.31 was used for multi-protein sequence alignment of the *G. maculatum* genes and their orthologs. Conserved CDS alignments of each single-copy gene family were extracted by Gblocks v0.91b (Talavera & Castresana, 2007) and used for further identification of positively selected genes. The ratios of nonsynonymous to synonymous substitutions (K_A/K_S, or ω) were estimated for each single-copy orthologous gene using the codeml program with the branch-site model implemented in PAML v4 (Guindon et al., 2010). A likelihood ratio test was conducted, and false discovery rate (FDR) correction was performed for multiple comparisons. Genes with corrected $P<0.01$ were defined as naturally selected genes.

Phylogeny and burst time estimation of LINE RTE-BovB and L2 type

Repetitive sequences in the *G. maculatum* genome were identified by a combination of homology searching and *ab initio* prediction. For homology-based prediction, we used RepeatMasker v4.1.2 (Bergman & Quesneville, 2007) and ProteinMask to search against RepBase. For *ab initio* prediction, we used Tandem Repeats Finder (TRF) v4.04 (Benson, 1999), LTR_FINDER v1.0 (Xu & Wang, 2007), PILER v1.3.11 (Edgar & Myers, 2005), and RepeatScout v1.0.5 (Price et al., 2005) with default parameters. We found that at least 32% of the *G. maculatum* assembly was composed of repetitive elements (Supplementary Table S5). The identified repeats were annotated using RepBase v23.04 (Jurka et al., 2005) with RepeatMasker v4.1.2 (Bergman & Quesneville, 2007) against this database. Other species, including *Astyanax mexicanus*, *Cyprinus carpio*, *Danio rerio*, *Ictalurus punctatus*, *Oncorhynchus mykiss*, *Oryzias latipes*, *Pelteobagrus fulvidraco*, and *Poecilia formosa*, were also detected using the same pipeline (Supplementary Tables S5, S6 and Figures S3, S4). Interestingly, we found that long interspersed nuclear elements (LINE) content was much more abundant in the *G. maculatum* genome.

The times of transposon bursts were estimated using previously published methods (Albertin et al., 2015). We dated transposable elements longer than 500 bp with RepeatMasker and adjusted the distances for multiple substitutions using the Jukes-Cantor formula:

$$JK = -3/4 \times \log(1 - 4 \times d/3) \quad (1)$$

where d is the distance estimated by RepeatMasker v4.1.2. Using an estimate of 0.0036 JK/myr for synonymous substitutions per million years (Albertin et al., 2015; Xu et al., 2014), the divergence time of the repeats was then estimated.

Population genetic analysis for *G. maculatum* using whole-genome data

Ten living individuals were collected from 2900, 4100-1, and 4100-2 using gill nets. All individuals were narcotized with MS-222 (Solarbio, China) for a few minutes before sampling muscles. For each individual, the muscle samples were immediately frozen in liquid nitrogen after dissection and stored at −80 °C until DNA extraction. Muscle samples from the outgroup, *E. labiatum*, were also collected.

For each sample, genomic DNA was extracted and used for library construction. Sequencing libraries were generated using a Truseq Nano DNA HT Sample Preparation Kit (Illumina, USA) following the manufacturer's recommendations. Fragments with a length of 350 bp were selected, end polished, A-tailed, and ligated with the full-length adapter for Illumina sequencing. After polymerase chain reaction (PCR) amplification, all products were purified (AMPure XP system) and analyzed for size distribution using an Agilent2100 Bioanalyzer (Agilent Technologies, USA) and quantified using real-time PCR. Sequencing was implemented by the Illumina Hiseq 2000 platform (Illumina, USA) and millions of 150 bp paired-end short reads were generated.

Raw data were appraised using FastQC v.0.10.1 and filtered using fastp v0.20.0 (Chen et al., 2018). The remaining high-quality paired-end reads were mapped to the genome using BWA6 v0.7.8 (Li & Durbin, 2010) with the command "mem -t 4 -k 32 –M". To reduce mismatches generated by PCR amplification before sequencing, duplicate reads were removed by SAMtools v1.9 (Li, 2011). Single nucleotide polymorphism (SNP) calling was performed following the Genome Analysis Toolkit (GATK) v4.1.9.0 best practices (McKenna et al., 2010). Annotations were performed using the ANNOVAR package (v2013-05-20) (Wang et al., 2010).

Based on the identified SNPs, a neighbor-joining tree representing the relationships of each individual was constructed using TreeBeST v1.9.2 (http://treesoft.

sourceforge.net/treebest.shtml) with 100 bootstrap resamplings. Based on the same data, the population genetic structure was estimated using the Bayesian computer algorithm with the R package LEA v3.13 (Frichot & François, 2015). Principal component analysis (PCA) was also implemented using R v4.1.0.

Demographic history of *G. maculatum*

The population history of three populations of *G. maculatum* was constructed using the multiple sequentially Markovian coalescent (MSMC) approach v1.1.0 (Schiffels & Durbin, 2014) with a generation time of 6 and mutation rate of 3.51×10^{-9} per year per nucleotide (Yang et al., 2016). MSMC uses a hidden Markov model to analyze patterns of heterozygosity along genome sequences. We also applied the Generalized Phylogenetic Coalescent Sampler (G-PhoCS) v1.2.3 (Gronau et al., 2011). The prior distributions over model parameters were defined by Gamma distributions (Gronau et al., 2011). Markov chain was run for 100 000 iterations and population parameter values were sampled every 10 iterations.

RESULTS

Genome and transcriptome assembly for phylogenetic analysis of sisorid catfish

Based on the contigs assembled from the PacBio long reads (Liu et al., 2018), Hi-C sequencing was applied with the same *G. maculatum* sample collected in the Tibet Plateau to obtain a high-quality chromosomal genome. A 704.8 Mb genome was obtained with a contig and scaffold N50 of 878.4 kb and 27.8 Mb, respectively. More than 91.7% of the sequences were anchored on chromosomes at the base level, resulting into 2 083 genomic sequences longer than 2 kb (Supplementary Table S2 and Figure S1). The GC content of the *G. maculatum* genome was 39.6%. The transcriptomes of 16 sisorid catfish and *Leiocassis longirostris* were sequenced and assembled for the following analyses (Supplementary Table S3).

The phylogenetic relationships of *G. maculatum* to the other 16 sisorid catfish species were constructed using the transcriptome data (Figure 1; Supplementary Table S3), with *L. longirostris* used as the outgroup species (Figure 2A). *Pseudecheneis* was the earlier diverging genus among Sisoridae, and glyptosternoid and non-glyptosternoid fish clearly formed two sister groups. *Glyptosternon* and *Exostoma* were primitive taxa among the glyptosternoid fish, while *Bagarius* was primitive in the non-glyptosternoid group. The *Pseudecheneis* genus diverged from the common ancestor of Sisoridae ~9.7 Mya and glyptosternoid and non-glyptosternoid fish separated ~8.1 Mya.

Protein-coding genes were used to reconstruct a genome-based phylogeny for *G. maculatum* among teleosts (Figure 2B). The analysis supported a close evolutionary relationship between *I. punetaus* and *P. fulvidraco*, consistent with the traditional taxonomy of Siluriformes. From our phylogenetic analysis, *G. maculatum* diverged from their common ancestor with *I. punetaus* and *P. fulvidraco* 37–43 Mya (Supplementary Figure S2), a time that roughly

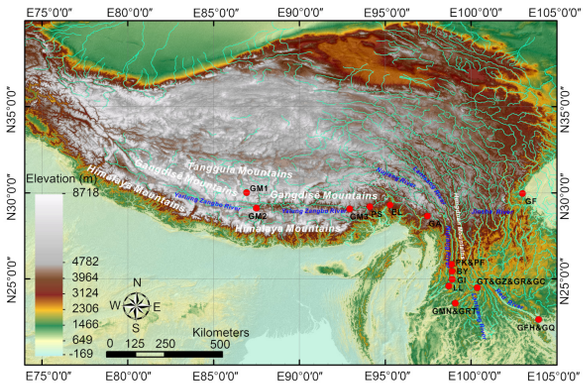

Figure 1 Sampling sites (red spot) for fish species on the Qinghai-Tibet Plateau

Altitude is represented by color bar. Species abbreviations are: *Glyptosternon maculatum* (GM), *Exostoma labiatum* (EL), *Pareuchiloglanis feae* (PF), *Pareuchiloglanis kamengensis* (PK), *Glyptothorax quadriocellatus* (GQ), *Glyptothorax fukiensis honghensis* (GFH), *Glyptothorax interspinalum* (GI), *Glyptothorax cavia* (GC), *Glyptothorax dorsalis* (GD), *Glyptothorax zainaensis* (GZ), *Glyptothorax trilineatus* (GT), *Glyptothorax minimaculatus* (GMN), *Glyptothorax laosensis* (GL), *Bagarius yarrelli* (BY), *Pseudecheneis sulcatus* (PS), and *Leiocassis longirostris* (LL).

corresponds to the collision of the ancient Indian and Eurasian plates in the Cenozoic era (Li et al., 2015).

Repetitive elements in *G. maculatum* genome

More than 30% of the *G. maculatum* genome was composed of repetitive sequences (Supplementary Table S5). The abundance of whole genomic repeats in the *G. maculatum* genome was comparable to that of other teleost fish, but the abundance of long interspersed elements (LINEs) comprised >23.1% and >66.3% of the genome and repetitive elements, respectively, significantly higher than that observed in other teleost fish (Figure 2B; Supplementary Table S4). LINEs (primarily LINE/L1 types) comprise ~18.9% of the human genome (Li et al., 2001), but do not dominate repeat regions in other teleost genomes (Supplementary Table S6). The two most abundant LINE transposons in the *G. maculatum* genome were RTE-BovB (93.4 Mb long and 12.3% of the genome) and L2 (57.9 Mb and 7.6%) (Supplementary Table S6). The abundance of long terminal repeats (LTRs) was also slightly higher in the *G. maculatum* genome relative to the other teleosts, but DNA transposon abundance was lower (Supplementary Table S6).

Timing of LINE expansions in *G. maculatum* genome

The timing of LINE transposon expansions was estimated from sequence comparisons of LINE RTE-BovB and L2. The estimated time for expansion bursts indicated a multi-stage expansion pattern for both types (Figure3; Supplementary Figures S3, S4). The first stage of expansion occurred slowly during the late Eocene and early Oligocene ~30 Mya. The second stage exhibited an even more substantial expansion during the early Miocene ~20 Mya. The third stage showed a sharp expansion starting from 3 Mya in the Pleistocene. We observed remarkable associations of tectonic movements and

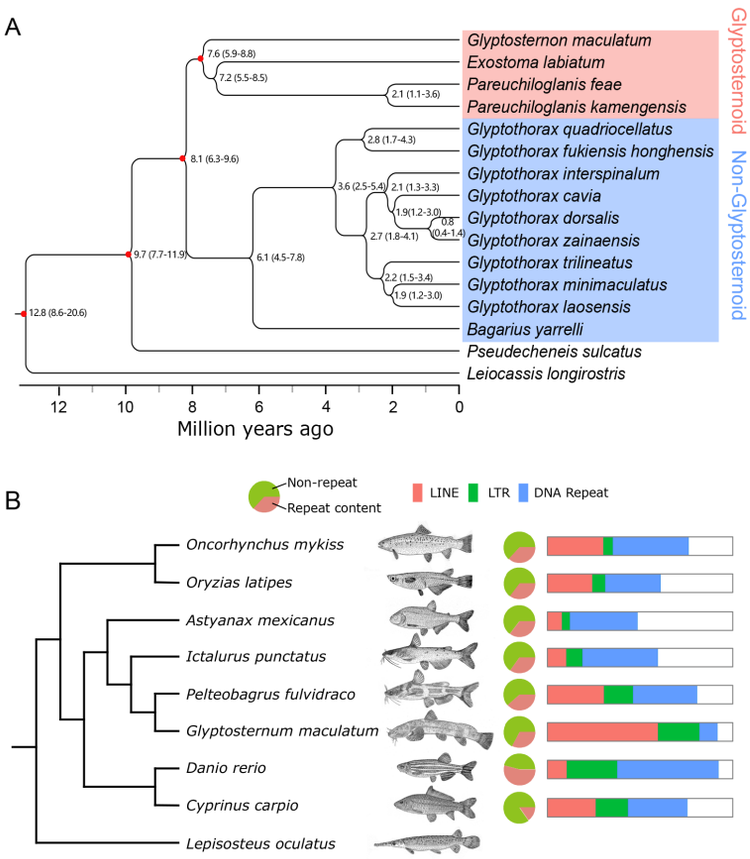

Figure 2 Phylogeny among sisorid catfish on Qinghai-Tibet Plateau and repeat content comparison of *Glyptosternon maculatum* to other teleosts

A: Phylogeny of QTP sisorid fish collected using transcriptome data and *G. maculatum* genome. Species highlighted by colors from glyptosternoid and non-glyptosternoid belong to Sisoridae, while other species are not from the family. Speciation divergence periods are labeled at branches with 95% confidence interval in parentheses. Divergence periods with red points were used for time recalibration. B: Repeat content in genomes of species closely related to *G. maculatum* and detailed repetitive element categories, including LINE (red), LTR (green), DNA transposon (blue), and other (blank).

global surface temperature fluctuations to the historical RTE-BovB and L2 bursts (Figure 3A, B).

Expanded and selected functional genes for high-altitude environmental adaptation

Using whole-genome protein-coding gene clustering of close-related fish species, we identified expanded gene families in the *G. maculatum* genome. The functions of the expanded genes were investigated by enrichment analysis (Supplementary Tables S7, S8). We observed a large-scale expansion of gene families related to chromosome organization, including DNA packing, nucleosome organization, chromatin assembly, nucleosome assembly, DNA conformational changes, and folate metabolism (Figure 4A).

Detailed analysis revealed that genes involved in nucleosome organization were mainly core histone proteins, including H2A, H2B, H3, and H4. In the *G. maculatum* genome, we identified 61, 85, 68, and 81 genes for the H2A, H2B, H3, and H4 histones, respectively, which were significantly higher than their copies in the *P. fulvidraco* and *I. punctaus* genomes (Figure 4B). Interestingly, most duplicated histone genes resided in the LINE/L2 elements. In contrast, few histones in the *P. fulvidraco* and *I. punctaus* genomes were found in LTR/Gypsy, simple repeat, DNA/TcMar-Tc1, and most histones did not overlap with any repetitive elements (Figure 4B). The mosaic structure of the histone with LINE/L2 demonstrated that histone proteins were expanded through the LINE/L2 bursts in the *G. maculatum* genome. Strikingly, based on the timing of the expansion of the LINE/L2 elements with histone genes, we found that those repetitive elements were duplicated ~2–1 Mya (Figure 4C) during the latest wave of transposon bursts, accelerated uplift of the plateau, and sudden drop in temperature (Figure 3B, C).

In the *G. maculatum* genome, many functional genes related to folate absorption and metabolism were expanded genome-wide and positively selected (Figure 4D). Two important genes, i.e., proton-coupled folate transporter (*pcft*) (FDR=0.000 34) and 5-aminoimidazole-4-carboxamide ribonucleotide (AICAR) transformylase/IMP cyclohydrolase (*purh*) (FDR=0.001 2), were positively selected in the likelihood ratio test for the nonsynonymous to synonymous

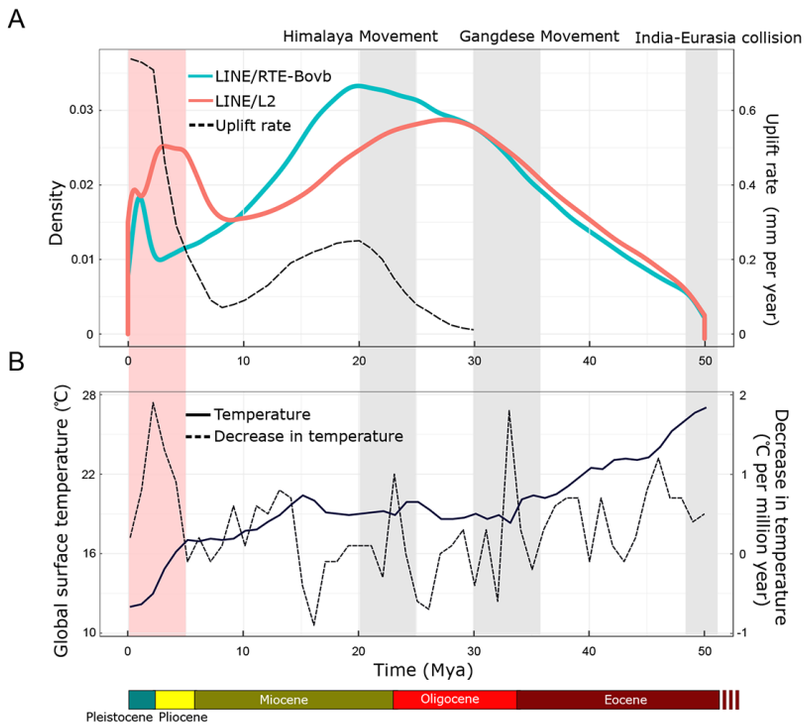

Figure 3 Expansion of LINE/RTE-BovB and LINE/L2 in *Glyptosternon maculatum* genome with tectonic movements and climate change during Qinghai-Tibet Plateau formation

A: Episodic bursts of RTE-BovB (red) and L2 (green) around ~30, ~20, and ~5-1 Mya after India-Eurasia collision, corresponding to Gangdese Movement and two uplift accelerations of QTP ~30 Mya (black dashed line). Recent burst is highlighted in red. Gray shows major tectonic movements. B: Global surface temperature (solid black line) and decrease in temperature per million years (dashed line) after plate collision. Data were collected from a previous paleoclimatology study (Zachos et al., 2001).

substitutions (K_A/K_S) (Anisimova et al., 2001). Remarkably, mitochondrial monofunctional C1-tetrahydrofolate synthase (*mthfd1l*), dihydrofolate reductase (*dhfr*), adenosine kinase (*adk*), phosphodiesterase 3B (*pde3b*), and adenylate cyclase 6 (*adcy6*) genes, which are involved in folate metabolism and purine biosynthesis, were significantly expanded in the *G. maculatum* genome (Figure 4E). We identified three mitochondrial monofunctional C1-tetrahydrofolate synthase (*mthfd1l*) genes with integrated functional domains in the *G. maculatum* genome, but only one in *P. fulvidraco* and *I. punetaus* (Figure 4E). We showed that duplicated *mthfd1l* genes resided in the repeat rich regions, mainly the LINE RTE-BovB and L2 elements. Interestingly, those repetitive elements showed excellent synteny along the genome (Figure 4F), implying that *mthfd1l* genes may be duplicated from segmental duplications after the repeat expansion.

Population genetic diversity for *G. maculatum*

To investigate genetic diversity among *G. maculatum* populations, we collected 30 samples from three sites: two from altitudes of 4 100 m (GM4100, GM2 in Figure 1) and 2 900 m a.s.l. (GM2900, GM3 in Figure 1) in the Yarlung Zangbo River and one from 4 500 m a.s.l. (GM4500, GM1 in Figure 1) in the Dogxung Zangbo River, a main branch of the Yarlung Zangbo River. Angren Lake resides in the Dogxung Zangbo River between the GM4300 and GM4500 populations (Figure 5A).

Using whole-genome resequencing data, we obtained 248 127–257 465 SNPs for species among populations (Supplementary Table S9). Sample clustering and a phylogenetic tree were constructed based on the whole-genome variations (Figure 5B, C). The GM2900, GM4100, and GM4500 populations were grouped into three clades based on PCA (Figure 5B) and phylogenetic analysis (Figure 5C). The genetic structures of the three populations further confirmed the PCA and phylogenetic analysis results (Figure 5D). Using *E. labiatum* as the out-group, phylogenetic analysis based on whole-genome variants implied populations of *G. maculatum* may have originated from habitats with lower altitudes in the southern QTP, consistent with previous fossil studies suggesting that glyptosternoids may have originated in southeastern Tibet and eastern Himalayan areas in Yunnan, China (Ma et al., 2015).

Using whole-genome variants among populations, we investigated the population history for *G. maculatum*. We found that the dynamic profiles of effective population size were similar for different populations (Figure 5E). Three populations experienced a sudden 60-fold drop in effective population size from ~4 000 around 3 Mya to ~60 around 2 Mya (Figure 5E). In addition, we performed a demographic analysis of *G. maculatum* using G-PhoCS (Gronau et al., 2011) to estimate the population split time and migration among populations. We showed that the GM2900 population split from their common ancestor population ~10 thousand

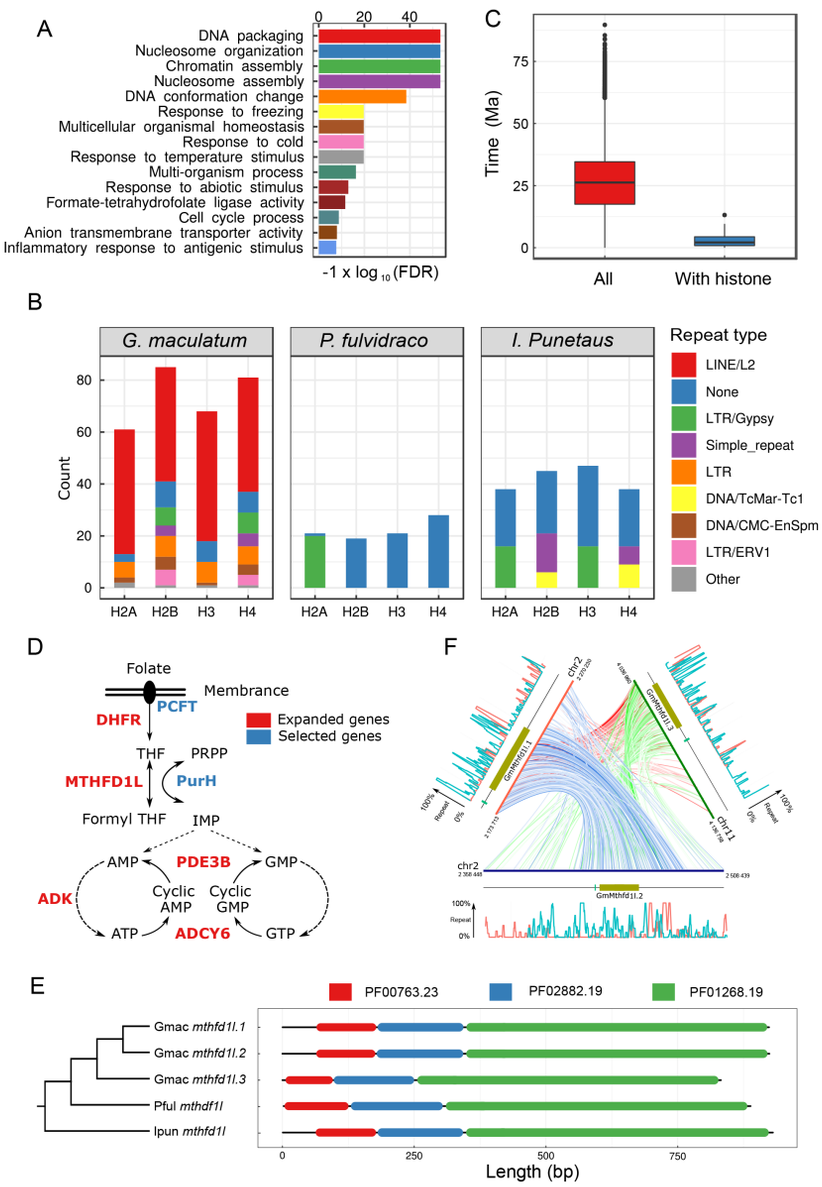

Figure 4 Transposon mediated expansion of histone and folate-related functional genes in *Glyptosternon maculatum* genome

A: GO functional enrichment for expanded gene families in *G. maculatum* genome. B: Number of histones and their positional relationship with transposons in *G. maculatum* genome, compared with their counterparts in *P. fulvidraco* and *I. punetaus*. Color represents repetitive types. C: Distribution of insertion times for all LINE/L2 and those with histone genes. D: Expanded and naturally selected genes in folate metabolism. E: Expansion of *mthfd1l* gene in *G. maculatum* genome compared to *P. fulvidraco* and *I. punetaus* genomes. F: Synteny of repetitive elements around duplicated *mthfd1l* gene in *G. maculatum* genome. LINE RTE-BovB (red) and L2 (green) content is shown around genes.

years before present (kyr BP) (Figure 5F). Remarkably, we observed that the effective population size of ancestral populations of *G. maculatum* experienced an almost 10 times reduction after the population splits, implying severe bottleneck effects during that period (Figure 5F). Furthermore, the low migration rates among populations indicated limited gene flow among populations (Figure 5F).

DISCUSSION

Our phylogenetic results of sisorid catfish species using whole-genome protein-coding genes are consistent with the results obtained from mitochondrial genes, which suggest that many specialized glyptosternoid genera originated 1.9 Mya in the Pleistocene (Peng et al., 2006; Yu & He, 2012), although we included more *Glyptothorax* species in this work. Our phylogenetic analysis also supported species radiation in the genus *Glyptothorax* ~3 Mya. We showed that *P. feae* and *P. kamengensis* diverged ~2.1 Mya (Figure 2A). We also demonstrated numerous speciation events for *Glyptothorax* over the last 3 Mya (Figure 2A). The times of species radiation for specialized glyptosternoid and non-glyptosternoid species are consistent with the latest sudden uplift of the QTP and

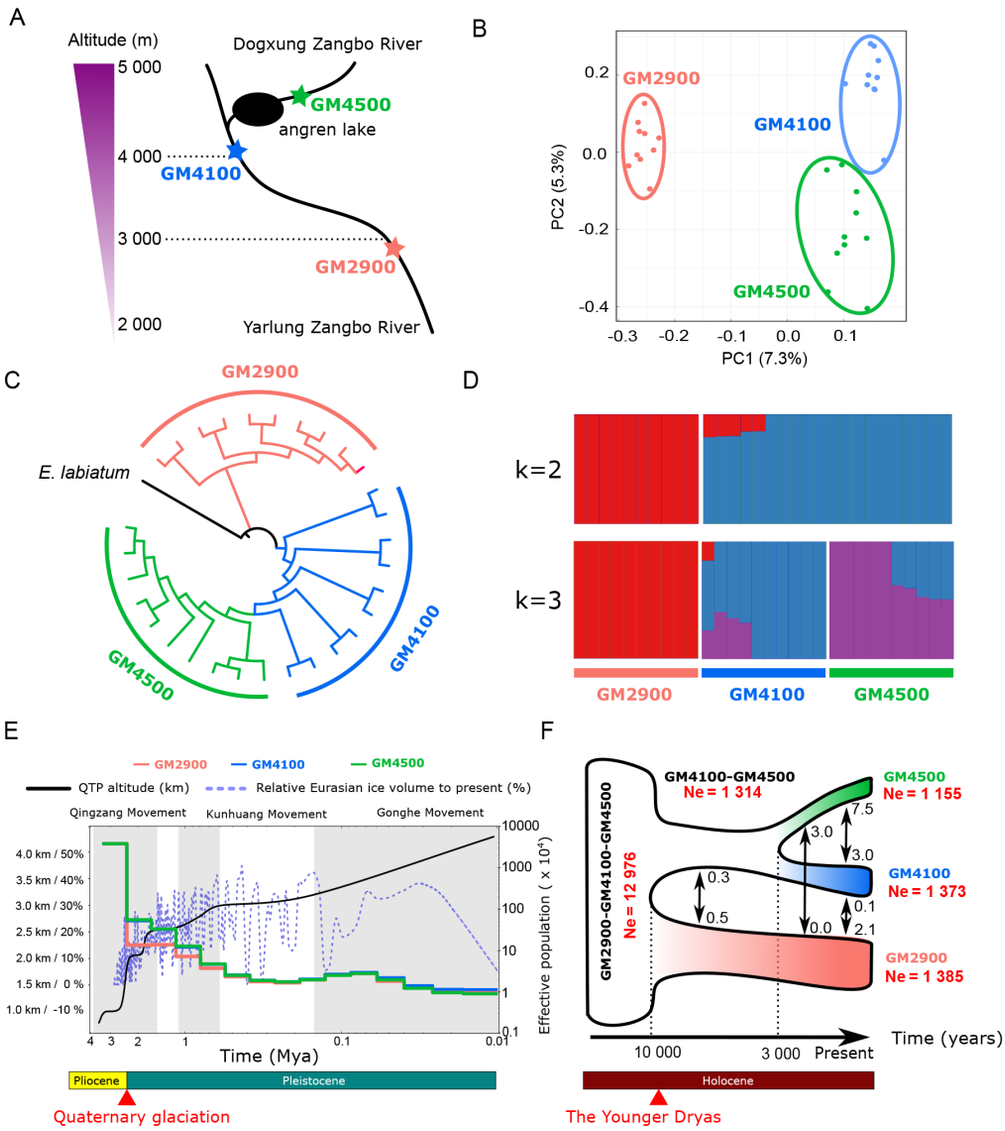

Figure 5 Population structure and history for *Glyptosternon maculatum*

A: Geographical representation of GM2900 (red), GM4100 (blue), and GM4500 (green). B: Principal component analysis (PCA) for all samples using whole-genome variants. C: Phylogenetic relationship among samples based on whole-genome variant data. D: Genetic structures of population samples. E: Average altitude of QTP (left Y axis and black sold line), relative Eurasian ice volume (left Y axis and blue dashed line), and effective population size from MSMC (right Y axis and solid red/blue/green line) in last 4 Mya. Effective population size (*Ne*) profiles were deduced from MSMC. Altitude of QTP and relative Eurasian ice volume were referenced from previous study (Yang et al., 2016). Gray highlights major tectonic movements. F: Demographic scenario for three populations deduced from G-PhoCS. Color scheme for each population is identical to above plot, and width indicates relative effective population size. Branch represents splits among populations. Numbers on arrows show migration rates among populations. Major climate events, including Quaternary glaciation ~2.6 Mya and Younger Dryas ~10 kyr BP, are labeled with red triangles along geological time.

Quaternary glacial period within the last 3 Mya (Jiang & Li, 2014). The coincidence of massive speciation with paleogeographic and paleoclimatic events suggests that species radiation is correlated with environmental fluctuations.

We also observed that the content of LINEs in *G. maculatum* genome, mainly RTE-BovB and L2, was significantly higher than that observed in other teleost fish (Figure 2B). Transposon elements and their associated functions have been found in teleost genomes recently. For example, Tcl-mariner transposons are thought to be involved in the rediploidization of the Atlantic salmon (*Salmo salar*) genome (Lien et al., 2016), and multiple bursts of LINE1, LINE2, CR1, and Deu are reported in the genomes of coelacanth (*Latimeria chalumnae*) (Chalopin et al., 2014) and lungfish (*Neoceratodus forsteri*) (Metcalfe et al., 2012). However, the distribution and roles of RTE-BovB and L2 in teleost evolution remain poorly investigated. RTE-BovB elements comprise 10.7% of the bovine genome (Adelson et

al., 2009), but this type of repeat is not a dominant transposon in most teleost genomes (Supplementary Table S6). Therefore, the expansion of RTE-BovB and L2 type elements in the *G. maculatum* genome is intriguing from an evolutionary perspective. Previous studies have demonstrated the potential function of transposon elements as an evolutionary driving force under rigorous environmental stress (Belyayev, 2014; Craddock, 2016; Platt et al., 2014). LINEs are non-LTR retrotransposons that can use self-transcribed reverse transcriptase to copy and insert their sequences with associated genes into genomes (Kazazian & Goodier, 2002). Consequently, the high accumulation of LINEs in the *G. maculatum* genome may be related to genome evolution and adaptation to the environment during the formation of the QTP.

Interestingly, the time estimations for the expansions of RTE-BovB and L2 type elements showed three episodic transposon bursts, illuminating another coincidence between genome evolution and time of QTP multi-stage uplift (Figure 3). Specifically, the timing of the first RTE-BovB and L2 expansion (~30 Mya) paralleled the Gangdese Movement periods in the QTP (Pan et al., 2012) (Figure 3A) and the divergence of the ancestors of *G. maculatum*, *I. punetaus,* and *P. fulvidraco* (~43–37 Mya) (Supplementary Figure S2). The second RTE-BovB and L2 expansion (~20 Mya) coincided with the even more significant QTP uplift event during the Himalaya Movement at 25–15 Mya (Pan et al., 2012) (Figure 3A). The third expansion (<3 Mya) was concurrent with the sudden accelerated Tibetan uplift during the Qingzang Movement ~4–1 Mya (Li et al., 2015) (Figure 3A). In addition, remarkable decreases in global surface temperature were also associated with each of the RTE-BovB and L2 bursts (Figure 3B). These drops in temperatures could significantly limit the habitat and geographically isolate endemic fish species on the QTP. Nevertheless, the episodic transposon expansions and their coincidence with multi-stage plateau uplift and climate change demonstrate the correlation between genome evolution in endemic fish and major paleogeographic and paleoclimatic events on the QTP.

We also found that the transposon expansions occurred after the ancestral divergence of *G. maculatum*, *I. punetaus,* and *P. fulvidraco* (43–37 Mya; Supplementary Figure S2) and before the speciation of glyptosternoids and non-glyptosternoids (~10 Mya; Figure 3A). Thus, genomic features of earlier transposon expansions (>10 Mya) may be shared among all glyptosternoid and non-glyptosternoid species on the QTP. The most recent transposon burst occurred suddenly over a narrow window that began 4 Mya (Figure 3A). Over this last period, QTP uplift occurred at an unprecedentedly high rate, some 3–4 times that of previous events (Jiang & Li, 2014). The environmental changes ~3 Mya may have exerted severe adaptive selection pressure on endemic fish, as sharp transposon expansion (Figure 3A) and massive speciation (Figure 2A) were both observed during this period.

Genes related to nucleosome organization and folate metabolism were expanded in the *G. maculatum* genome (Figure 4A). The main histone genes, including H2A, H2B, H3, and H4, were significantly expanded in *G. maculatum* compared to *P. fulvidraco* and *I. punetaus* (Figure 4B). Histones are proteins within chromatin, and they play important structural roles in DNA packaging and structural stability (Mariño-Ramírez et al., 2005), as well as functional roles in cold stress (Verleih et al., 2015), antibiotic stimulus (Lü et al., 2014; Noga et al., 2011), UV radiation (Pawlak & Deckert, 2007), and DNA repair (Schild-Poulter et al., 2003). Previous studies have demonstrated the importance of histone protein and chromatin structures to genomic stability (Oberdoerffer & Sinclair, 2007). Our results indicated that LINE L2 bursts mediated the recent and rapid expansion of H2A, H2B, H3, and H4 histone genes in the *G. maculatum* genome, which may be in reaction to extreme environmental adaptations to cold temperatures, food shortages, and UV exposure. Previous studies have also revealed that folate contributes to high-altitude environmental adaptation due to its important role in UV protection and DNA repair (Jablonski & Chaplin, 2010) and folate-related genes exhibit significant signals of high-altitude adaptation selection in Tibetans (Yang et al., 2017). *Mthfd1l*, which is a mitochondrial monofunctional enzyme with 10-formyl-tetrahydrofolate (10-CHO-THF) synthase activity, plays a critical role in the folate cycle and cytoplasmic formate production (Tibbetts & Appling, 2010). In human diseases, *mthfd1l* also contributes to the production and accumulation of NADPH to levels that are sufficient to combat oxidative stress for cell cycle delay and apoptosis, especially in cancer cells (Lee et al., 2017). We observed the mosaic structures of histone genes and *mthfd1l* within repetitive regions, implying that LINE RTE-BovB and L2 bursts may facilitate whole-genome wide expansion of functional genes favorable for environmental adaptation on the QTP.

Glyptosternon maculatum is the only critically endangered Sisoridae species distributed at high altitudes of 2 800 m to 4 500 m a.s.l. on the QTP (Zhang et al., 2010). However, our understanding of the genetic structure and population evolution of this species on the QTP is still not clear. Using whole-genome sequencing data, we observed roughly one SNP per 2.7 kb in the genome of the wild *G. maculatum* populations, which was significantly lower than that found in the *I. punetaus* genome (one SNP per 93 bp) (Liu et al., 2016), thus demonstrating extremely low genetic diversity in *G. maculatum* populations.

Population analysis showed that the effective population size of ancient *G. maculatum* populations experienced a sharp 60-fold decline ~2–3 Mya (Figure 5E), which coincided with the timing of the Quaternary glaciation. Quaternary glaciation, beginning 2.58 Mya, was the last major ice age to occur (Owen et al., 2008). The alternation between glaciation and interglaciation during the Quaternary glaciation period had large impact on the connectivity of the primary drainage system of rivers and lakes on the QTP (Lehmkuhl & Owen, 2005), which likely lead to geographic barriers, and thus to population isolation and speciation of endemic fish on the QTP. We also found that the GM2900, GM4100, and GM4500 populations split ~10 kyr BP (Figure 5F). The sudden cold and dry global climate during the Younger Dryas (11–10 kyr BP) may have contributed to the population split (Gasse et al., 1991) as the drop in temperature could have disrupted or blocked water connectivity and limited gene flow among populations during that period. Diatom records and climate

studies for the Angren Lake also show an extremely low abundance of diatoms ~10 kyr BP (Li et al., 1999), implying that food scarcity may also be an important reason for the population bottleneck.

CONCLUSIONS

We analyzed the genomic features and population diversity of *G. maculatum*, a representative endemic fish on the QTP and the most ancient species within the subfamily Glyptosterninae. We found that whole-genome wide transposons, especially dominant repetitive elements of LINE RTE-BovB and L2, showed episodic bursts, coinciding with the timing of accelerated uplift of the QTP and dramatic climatic fluctuations. We also showed that these transposons mediated functional gene expansions, which may have contributed to the environmental adaptation of *G. maculatum*. Using whole-genome variants, we determined that the ancestral *G. maculatum* group experienced two large-scale population depressions during the Quaternary glaciation ~2.6 Mya and the Younger Dryas ~10 kyr BP. The synchronous tectonic movements and temperature drops during glacial periods would likely have frozen habitats and disrupted drainage connectivity, leading to the historical bottleneck effects for ancient *G. maculatum* populations. For the first time, we revealed the synchronization of the genome and population evolution of *G. maculatum* with historical tectonic movements and climate events during QTP formation. Therefore, we propose that paleogeography and paleoclimate may be dominating driving forces for genome and population evolution of endemic fish on the QTP. The population whole-genome variant data provide a valuable genetic resource and opportunity to study genome and population evolution and to investigate the molecular mechanism underlying the extreme environmental adaptations of endemic fish species on the QTP.

DATA AVAILABILITY

The Hi-C sequencing data for *G. maculatum* were submitted to the National Center for Biotechnology Information (NCBI) BioProject No. PRJNA447978. Genomic and transcriptome sequencing data are available from the NCBI Short Read Archive as SRR7268130–SRR7268162.

SUPPLEMENTARY DATA

Supplementary data to this article can be found online.

COMPETING INTERESTS

The authors declare that they have no competing interests.

AUTHORS' CONTRIBUTIONS

H.P.L. designed the research. H.P.L., S.J.X., Z.B.M., R.B.Y., C.W.Z., and S.J.X. designed and performed the experiments. S.J.X., H.P.L., Z.B.M., J.Q.L., Y.Z., D.D.F., S.L.Z., M.Z., and J.Q.L. performed repetitive element analysis. H.P.L., Z.B.M., and R.B.Y. contributed materials. S.J.X., Y.Z., and J.Q.L. created the figures. S.J.X. and H.P.L. wrote the manuscript, which was reviewed by all authors. All authors read and approved the final version of the manuscript.

REFERENCES

Adelson DL, Raison JM, Edgar RC. 2009. Characterization and distribution of retrotransposons and simple sequence repeats in the bovine genome. *Proceedings of the National Academy of Sciences of the United States of America*, **106**(31): 12855−12860.

Ai HS, Yang B, Li J, Xie XH, Chen H, Ren J. 2014. Population history and genomic signatures for high-altitude adaptation in tibetan pigs. *BMC Genomics*, **15**(1): 834.

Albertin CB, Simakov O, Mitros T, Wang ZY, Pungor JR, Edsinger-Gonzales E, et al. 2015. The octopus genome and the evolution of cephalopod neural and morphological novelties. *Nature*, **524**(7564): 220−224.

Alexa A, Rahnenführer J, Lengauer T. 2006. Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. *Bioinformatics*, **22**(13): 1600−1607.

Anisimova M, Bielawski JP, Yang ZH. 2001. Accuracy and power of the likelihood ratio test in detecting adaptive molecular evolution. *Molecular Biology and Evolution*, **18**(8): 1585−1592.

Beall CM. 2007. Two routes to functional adaptation: Tibetan and andean high-altitude natives. *Proceedings of the National Academy of Sciences of the United States of America*, **104**(S1): 8655−8660.

Belyayev A. 2014. Bursts of transposable elements as an evolutionary driving force. *Journal of Evolutionary Biology*, **27**(12): 2573−2584.

Benson G. 1999. Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Research*, **27**(2): 573−580.

Bergman CM, Quesneville H. 2007. Discovering and detecting transposable elements in genome sequences. *Briefings in Bioinformatics*, **8**(6): 382−392.

Chalopin D, Fan SH, Simakov O, Meyer A, Schartl M, Volff JN. 2014. Evolutionary active transposable elements in the genome of the coelacanth. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution*, **322**(6): 322−333.

Chen SF, Zhou YQ, Chen YR, Gu J. 2018. Fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*, **34**(17): i884−i890.

Craddock EM. 2016. Profuse evolutionary diversification and speciation on volcanic islands: transposon instability and amplification bursts explain the genetic paradox. *Biology Direct*, **11**(1): 44.

De Bie T, Cristianini N, Demuth JP, Hahn MW. 2006. CAFE: a computational tool for the study of gene family evolution. *Bioinformatics*, **22**(10): 1269−1271.

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, **32**(5): 1792−1797.

Edgar RC, Myers EW. 2005. Piler: identification and classification of genomic repeats. *Bioinformatics*, **21**(S1): i152−i158.

Frichot E, François O. 2015. LEA: an R package for landscape and ecological association studies. *Methods in Ecology and Evolution*, **6**(8): 925−929.

Gasse F, Arnold M, Fontes JC, Fort M, Gibert E, Huc A, et al. 1991. A 13,000-year climate record from western Tibet. *Nature*, **353**(6346): 742−745.

Ge RL, Cai QL, Shen YY, San A, Ma L, Zhang Y, et al. 2013. Draft genome sequence of the Tibetan antelope. *Nature Communications*, **4**: 1858.

Gong GR, Dan C, Xiao SJ, Guo WJ, Huang PP, Xiong Y, et al. 2018. Chromosomal-level assembly of yellow catfish genome using third-

generation DNA sequencing and Hi-C analysis. *GigaScience*, **7**(11): giy120.
Gou X, Wang Z, Li N, Qiu F, Xu Z, Yan DW, et al. 2014. Whole-genome sequencing of six dog breeds from continuous altitudes reveals adaptation to high-altitude hypoxia. *Genome Research*, **24**(8): 1308−1315.
Gronau I, Hubisz MJ, Gulko B, Danko CG, Siepel A. 2011. Bayesian inference of ancient human demography from individual genome sequences. *Nature Genetics*, **43**(10): 1031−1034.
Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of phyml 3.0. *Systematic Biology*, **59**(3): 307−321.
Guo XG, He SP, Zhang YG. 2005. Phylogeny and biogeography of chinese sisorid catfishes re-examined using mitochondrial cytochrome *b* and 16S rRNA gene sequences. *Molecular Phylogenetics and Evolution*, **35**(2): 344−362.
He SP, Cao WX, Chen YY. 2001. The uplift of Qinghai-Xizang (Tibet) plateau and the vicariance speciation of Glyptosternoid fishes (siluriformes: Sisoridae). *Science in China Series C: Life Sciences*, **44**(6): 644−651.
Jablonski NG. 2017. Genes for the high life: new genetic variants point to positive selection for high altitude hypoxia in Tibetans. *Zoological Research*, **38**(3): 117.
Jablonski NG, Chaplin G. 2010. Human skin pigmentation as an adaptation to UV radiation. *Proceedings of the National Academy of Sciences of the United States of America*, **107**(S2): 8962−8968.
Jiang CN, Yu GR, Li YN, Cao GM, Yang ZP, Sheng WP, et al. 2012. Nutrient resorption of coexistence species in alpine meadow of the Qinghai-Tibetan Plateau explains plant adaptation to nutrient-poor environment. *Ecological Engineering*, **44**: 1−9.
Jiang XD, Li ZX. 2014. Seismic reflection data support episodic and simultaneous growth of the Tibetan Plateau since 25 MYR. *Nature Communications*, **5**: 5453.
Jurka J, Kapitonov VV, Pavlicek A, Klonowski P, Kohany O, Walichiewicz J. 2005. Repbase update, a database of eukaryotic repetitive elements. *Cytogenetic and Genome Research*, **110**(1−4): 462−467.
Kang JL, Ma XH, He SP. 2017. Evidence of high-altitude adaptation in the Glyptosternoid fish, *Creteuchiloglanis macropterus* from the nujiang river obtained through transcriptome analysis. *BMC Evolutionary Biology*, **17**(1): 229.
Kazazian Jr HH, Goodier JL. 2002. LINE drive: retrotransposition and genome instability. *Cell*, **110**(3): 277−280.
Kumar S, Stecher G, Tamura K. 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution*, **33**(7): 1870−1874.
Lee D, Xu IMJ, Chiu DKC, Lai RKH, Tse APW, Li LL, et al. 2017. Folate cycle enzyme mthfd1l confers metabolic advantages in hepatocellular carcinoma. *The Journal of Clinical Investigation*, **127**(5): 1856−1872.
Lehmkuhl F, Owen LA. 2005. Late quaternary glaciation of tibet and the bordering mountains: a review. *Boreas*, **34**(2): 87−100.
Li H. 2011. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics*, **27**(21): 2987−2993.
Li H, Durbin R. 2009. Fast and accurate short read alignment with burrows-wheeler transform. *Bioinformatics*, **25**(14): 1754−1760.
Li H, Durbin R. 2010. Fast and accurate long-read alignment with burrows-wheeler transform. *Bioinformatics*, **26**(5): 589−595.
Li JJ, Zhou SZ, Zhao ZJ, Zhang J. 2015. The Qingzang movement: the major uplift of the Qinghai-Tibetan Plateau. *Science China Earth Sciences*, **58**(11): 2113−2122.
Li L, Stoeckert Jr CJ, Roos DS. 2003. OrthoMCL: Identification of ortholog groups for eukaryotic genomes. *Genome Research*, **13**(9): 2178−2189.
Li SF, Wang FB, Zhang J, Shen XH, Zeng ZQ. 1999. Diato Diatom-based reconstruction of Holocene environmental changes in Angren Lake, Southern Tibet. *Chinese Science Bulletin*, **44**(12): 1123−1126.
Li WH, Gu ZL, Wang HD, Nekrutenko A. 2001. Evolutionary analyses of the human genome. *Nature*, **409**(6822): 847−849.
Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, et al. 2009. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science*, **326**(5950): 289−293.
Lien S, Koop BF, Sandve SR, Miller JR, Kent MP, Nome T, et al. 2016. The atlantic salmon genome provides insights into rediploidization. *Nature*, **533**(7602): 200−205.
Liu HP, Liu QY, Chen ZQ, Liu YC, Zhou CW, Liang QQ, et al. 2018. Draft genome of *Glyptosternon maculatum*, an endemic fish from tibet plateau. *GigaScience*, **7**(9): giy104.
Liu ZJ, Liu SK, Yao J, Bao LS, Zhang JR, Li Y, et al. 2016. The channel catfish genome sequence provides insights into the evolution of scale formation in teleosts. *Nature Communications*, **7**: 11757.
Lü AJ, Hu XC, Wang Y, Shen XJ, Li X, Zhu AH, et al. 2014. ITRAQ analysis of gill proteins from the zebrafish (*Danio rerio*) infected with *Aeromonas hydrophila*. *Fish & Shellfish Immunology*, **36**(1): 229−239.
Ma XH, Dai W, Kang JL, Yang LD, He SP. 2016. Comprehensive transcriptome analysis of six catfish species from an altitude gradient reveals adaptive evolution in Tibetan fishes. *G3 Genes|Genomes|Genetics*, **6**(1): 141−148.
Ma XH, Kang JL, Chen WT, Zhou CJ, He SP. 2015. Biogeographic history and high-elevation adaptations inferred from the mitochondrial genome of Glyptosternoid fishes (sisoridae, siluriformes) from the southeastern Tibetan Plateau. *BMC Evolutionary Biology*, **15**: 233.
Mariño-Ramírez L, Kann MG, Shoemaker BA, Landsman D. 2005. Histone structure and nucleosome stability. *Expert Review of Proteomics*, **2**(5): 719−729.
McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. 2010. The genome analysis toolkit: a mapreduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, **20**(9): 1297−1303.
Metcalfe CJ, Filée J, Germon I, Joss J, Casane D. 2012. Evolution of the australian lungfish (*Neoceratodus forsteri*) genome: a major role for cr1 and l2 LINE elements. *Molecular Biology and Evolution*, **29**(11): 3529−3539.
Noga EJ, Borron PJ, Hinshaw J, Gordon WC, Gordon LJ, Seo JK. 2011. Identification of histones as endogenous antibiotics in fish and quantification in rainbow trout (*Oncorhynchus mykiss*) skin and gill. *Fish Physiology and Biochemistry*, **37**(1): 135−152.
Oberdoerffer P, Sinclair DA. 2007. The role of nuclear architecture in genomic instability and ageing. *Nature Reviews Molecular Cell Biology*, **8**(9): 692−702.
Owen LA, Caffee MW, Finkel RC, Seong YB. 2008. Quaternary glaciation of the himalayan–Tibetan orogen. *Journal of Quaternary Science*, **23**(6-7): 513−531.
Pan GT, Wang LQ, Li RS, Yuan SH, Ji WH, Yin FG, et al. 2012. Tectonic evolution of the Qinghai-Tibet Plateau. *Journal of Asian Earth Sciences*, **53**: 3−14.

Pawlak S, Deckert J. 2007. Histone modifications under environmental stress. *Biological Letters*, **44**(2): 65−73.

Peng ZG, Ho SYW, Zhang YG, He SP. 2006. Uplift of the Tibetan Plateau: evidence from divergence times of Glyptosternoid catfishes. *Molecular Phylogenetics and Evolution*, **39**(2): 568−572.

Petousi N, Robbins PA. 2014. Human adaptation to the hypoxia of high altitude: the Tibetan paradigm from the pregenomic to the postgenomic era. *Journal of Applied Physiology*, **116**(7): 875−884.

Platt II RN, Vandewege MW, Kern C, Schmidt CJ, Hoffmann FG, Ray DA. 2014. Large numbers of novel miRNAs originate from DNA transposons and are coincident with a large species radiation in bats. *Molecular Biology and Evolution*, **31**(6): 1536−1545.

Price AL, Jones NC, Pevzner PA. 2005. *De novo* identification of repeat families in large genomes. *Bioinformatics*, **21**(S1): i351−i358.

Qiu Q, Zhang GJ, Ma T, Qian WB, Wang JY, Ye ZQ, et al. 2012. The yak genome and adaptation to life at high altitude. *Nature Genetics*, **44**(8): 946−949.

Qu YH, Zhao HW, Han NJ, Zhou GY, Song G, Gao B, et al. 2013. Ground tit genome reveals avian adaptation to living at high altitudes in the Tibetan Plateau. *Nature Communications*, **4**: 2071.

Schiffels S, Durbin R. 2014. Inferring human population size and separation history from multiple genome sequences. *Nature Genetics*, **46**(8): 919−925.

Schild-Poulter C, Shih A, Yarymowich NC, Haché RJG. 2003. Down-regulation of histone H2B by DNA-dependent protein kinase in response to DNA damage through modulation of octamer transcription factor 1. *Cancer Research*, **63**(21): 7197−7205.

Talavera G, Castresana J. 2007. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Systematic Biology*, **56**(4): 564−577.

Tibbetts AS, Appling DR. 2010. Compartmentalization of mammalian folate-mediated one-carbon metabolism. *Annual Review of Nutrition*, **30**: 57−81.

Verleih M, Borchel A, Krasnov A, Rebl A, Korytář T, Kühn C, et al. 2015. Impact of thermal stress on kidney-specific gene expression in farmed regional and imported rainbow trout. *Marine Biotechnology*, **17**(5): 576−592.

Wang GD, Fan RX, Zhai WW, Liu F, Wang L, Zhong L, et al. 2014. Genetic convergence in the adaptation of dogs and humans to the high-altitude environment of the Tibetan Plateau. *Genome Biology and Evolution*, **6**(8): 2122−2128.

Wang K, Li MY, Hakonarson H. 2010. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Research*, **38**(16): e164.

Wei CH, Wang HH, Liu G, Zhao FP, Kijas JW, Ma YJ, et al. 2016. Genome-wide analysis reveals adaptation to high altitudes in Tibetan sheep. *Scientific Reports*, **6**: 26770.

Xiao SJ, Mou ZB, Fan DD, Zhou H, Zou M, Zou Y, et al. 2020. Genome of tetraploid fish *Schizothorax o'connori* provides insights into early re-diploidization and high-altitude adaptation. *iScience*, **23**(9): 101497.

Xu P, Zhang XF, Wang XM, Li JT, Liu GM, Kuang YY, et al. 2014. Genome sequence and genetic diversity of the common carp, *Cyprinus carpio*. *Nature Genetics*, **46**(11): 1212−1219.

Xu Z, Wang H. 2007. LTR_Finder: an efficient tool for the prediction of full-length ltr retrotransposons. *Nucleic Acids Research*, **35**(S1): W265−W268.

Yang J, Jin ZB, Chen J, Huang XF, Li XM, Liang YB, et al. 2017. Genetic signatures of high-altitude adaptation in Tibetans. *Proceedings of the National Academy of Sciences of the United States of America*, **114**(16): 4189−4194.

Yang JX, Chen XL, Bai J, Fang DM, Qiu Y, Jiang WS, et al. 2016. The *Sinocyclocheilus* cavefish genome provides insights into cave adaptation. *BMC Biology*, **14**: 1.

Yang SJ, Dong HL, Lei FM. 2009. Phylogeography of regional fauna on the Tibetan Plateau: a review. *Progress in Natural Science*, **19**(7): 789−799.

Yi X, Liang Y, Huerta-Sanchez E, Jin X, Cuo ZX, Pool JE, et al. 2010. Sequencing of 50 human exomes reveals adaptation to high altitude. *Science*, **329**(5987): 75−78.

Yu ML, He SP. 2012. Phylogenetic relationships and estimation of divergence times among sisoridae catfishes. *Science China Life Sciences*, **55**(4): 312−320.

Yuan DY, Chen XH, Gu HR, Zou M, Zou Y, Fang J, et al. 2020. Chromosomal genome of *Triplophysa bleekeri* provides insights into its evolution and environmental adaptation. *GigaScience*, **9**(11): giaa132.

Zachos J, Pagani M, Sloan L, Thomas E, Billups K. 2001. Trends, rhythms, and aberrations in global climate 65 Ma to present. *Science*, **292**(5517): 686−693.

Zhang HJ, Xie CX, Li DP, Liu HP, Yang XF. 2011. Blood cells of a sisorid catfish *Glyptosternum maculatum* (siluriformes: Sisoridae), in Tibetan Plateau. *Fish Physiology and Biochemistry volume*, **37**(1): 169−176.

Zhang HJ, Xie CX, Li DP, Xiong DM, Liu HP, Suolang SZ, et al. 2010. Haematological and blood biochemical characteristics of *Glyptosternum maculatum* (siluriformes: Sisoridae) in Xizang (Tibet). *Fish Physiology and Biochemistry*, **36**(3): 797−801.

Zhang Q, Gou WY, Wang XT, Zhang YW, Ma J, Zhang HL, et al. 2016. Genome resequencing identifies unique adaptations of tibetan chickens to hypoxia and high-dose ultraviolet radiation in high-altitude environments. *Genome Biology and Evolution*, **8**(3): 765−776.