SCIENCE CHINA
Life Sciences

•RESEARCH PAPER•

https://doi.org/10.1007/s11427-022-2253-7

Enlarged fins of Tibetan catfish provide new evidence of adaptation to high plateau

Liandong Yang[1,3,5†], Ning Sun[1,6†], Honghui Zeng[1†], Ying Wang[1,3,4†], Wenjun Chen[1,6], Zufa Ding[1,6], Yang Liu[1], Jing Wang[1], Minghui Meng[1], Yanjun Shen[1], Jingliang Kang[1], Xiuhui Ma[1], Wenqi Lv[1], Juan Chen[1], Axel Meyer[7*], Baocheng Guo[2,3,8*] & Shunping He[1,3,8*]

[1]*State Key Laboratory of Freshwater Ecology and Biotechnology, Institute of Hydrobiology, Chinese Academy of Sciences, Wuhan 430072, China;*
[2]*Key Laboratory of Zoological Systematics and Evolution, Institute of Zoology, Chinese Academy of Sciences, Beijing 100101, China;*
[3]*Academy of Plateau Science and Sustainability, Qinghai Normal University, Xining 810016, China;*
[4]*Hubei Engineering Research Center for Protection and Utilization of Special Biological Resources in the Hanjiang River Basin, School of Life Sciences, Jianghan University, Wuhan 430056, China;*
[5]*State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming 650223, China;*
[6]*University of Chinese Academy of Sciences, Beijing 100049, China;*
[7]*Department of Biology, University of Konstanz, 78457 Konstanz, Germany;*
[8]*Center for Excellence in Animal Evolution and Genetics, Chinese Academy of Sciences, Kunming 650223, China*

Received October 12, 2022; accepted November 30, 2022; published online February 16, 2023

The uplift of the Tibetan Plateau significantly altered the geomorphology and climate of the Euroasia by creating large mountains and rivers. Fishes are more likely to be affected relative to other organisms, as they are largely restricted to river systems. Faced with the rapidly flowing water in the Tibetan Plateau, a group of catfish has evolved greatly enlarged pectoral fins with more numbers of fin-rays to form an adhesive apparatus. However, the genetic basis of these adaptations in Tibetan catfishes remains elusive. In this study, we performed comparative genomic analyses based on the chromosome-level genome of *Glyptosternum maculatum* in family Sisoridae and detected some proteins with conspicuously high evolutionary rates in particular in genes involved in skeleton development, energy metabolism, and hypoxia response. We found that the *hoxd12a* gene evolved faster and a loss-of-function assay of *hoxd12a* supports a potential role for this gene in shaping the enlarged fins of these Tibetan catfishes. Other genes with amino acid replacements and signatures of positive selection included proteins involved in low temperature (*TRMU*) and hypoxia (*VHL*) responses. Functional assays reveal that the *G. maculatumTRMU* allele generates more mitochondrial ATP than the ancestral allele found in low-altitude fishes. Functional assays of *VHL* alleles suggest that the *G. maculatum* allele has lower transactivation activity than the low-altitude forms. These findings provide a window into the genomic underpinnings of physiological adaptations that permit *G. maculatum* to survive in the harsh environment of the Tibetan Himalayas that mirror those that are convergently found in other vertebrates such as humans.

***Glyptosternum maculatum*, enlarged fins, adaptation, Tibetan Plateau**

Citation: Yang, L., Sun, N., Zeng, H., Wang, Y., Chen, W., Ding, Z., Liu, Y., Wang, J., Meng, M., Shen, Y., et al. (2023). Enlarged fins of Tibetan catfish provide new evidence of adaptation to high plateau. Sci China Life Sci 66, https://doi.org/10.1007/s11427-022-2253-7

†Contributed equally to this work
*Corresponding authors (Axel Meyer, email: axel.meyer@uni-konstanz.de; Baocheng Guo, email: guobaocheng@ioz.ac.cn; Shunping He, email: clad@ihb.ac.cn)

INTRODUCTION

The collision between Indian plate and Eurasian continent around 50 million years ago has dramatically changed the geomorphology, climate, and even biodiversity of Asia (Spicer et al., 2021). As the largest, highest, and also one of the youngest plateaus in the world, the Tibetan Plateau stands 5 km high over a region of approximately 3 million km^2 (Royden et al., 2008). The uplift of the Tibetan Plateau gives rise to many large mountains and rivers, including the Yangtze River, the Yellow River, the Nujiang River, and the Lancang River. The high-altitude environment of the Tibetan Plateau is characterized by low oxygen availability, low temperatures, and strong ultraviolet radiation, thus posing great physiological challenges for endemic organisms (Bickler and Buck, 2007). Relative to terrestrial animals, fishes are more likely to be affected in the face of the uplift of the Tibetan Plateau, as they are largely restricted to river systems.

The inhospitable environments of the Tibetan Plateau have significantly reduced the diversity of fish lineages, including the number of Order and Family. For example, endemic fishes on the Tibetan Plateau mostly belong to three independent fish groups: Tibetan Loaches (family Nemacheilidae; order Cypriniformes), Schizothoracine fishes (family Cyprinidae; order Cypriniformes), and Glyptosternoid fishes (family Sisoridae; order Siluriformes) (Peng et al., 2006; Wu and Wu, 1992). However, the expansion of ecological niches derived from species extinction and uplift of the Tibetan Plateau provides a great opportunity to make the diversification of fish species. Correspondingly, these three lineages of fishes have diversified and widely distributed throughout the Tibetan Plateau and its peripheral regions, including 140 Tibetan Loaches, 76 Schizothoracine fishes, and 71 Glyptosternoid fishes (Chu et al., 1999; Ma et al., 2015; Wang et al., 2016; Yue, 2000). Therefore, these fishes represent an ideal model system to study the genetic basis of adaptation to high altitudes in aquatic animals (Yang et al., 2021).

Recent advances in genomic technologies have enabled many genomes from endemic organisms on the Tibetan Plateau to be decoded, and the genetic basis of adaptation to high altitude is being investigated (Ge et al., 2013; Hao et al., 2019; Li et al., 2018; Qiu et al., 2012; Qu et al., 2020; Qu et al., 2021; Xiong et al., 2020; Yu et al., 2016; Zhu et al., 2018). Although nearly all these studies found high-altitude adaptation is associated with modifications of genetic elements in hypoxia and energy metabolism, multiple genetic routes are involved in different organisms (Hao et al., 2019; Li et al., 2013; Qu et al., 2013; Zhang et al., 2016). Therefore, studies of genetic basis of high-altitude adaptation are needed for more endemic organisms. Moreover, phenotypic adaptations to the Tibetan Plateau are involved with many other aspects, such as adaptation to the fast-water in Glyptosternoid fishes. The uplift of the Tibetan Plateau has been assumed to facilitate speciation and adaptive diversification of the Glyptosternoid fishes (Guo et al., 2005; Peng et al., 2004; Peng et al., 2006). In order to adapt to torrential mountain stream life, Glyptosternoid fishes have evolved highly specialized body shape, with strongly depressed heads and bodies, and greatly enlarged pectoral and pelvic fins modified to form an adhesive apparatus (Guo et al., 2005). However, the genetic basis of these flatted body shape in Glyptosternoid fishes is largely unexplored.

In this study, sisorid catfish species *Glyptosternum maculatum* as a representative species of Glyptosternoid fishes was employed to explore the genetic basis of adaptation to high altitude of the Tibetan Plateau by integrating genomic and functional approaches. The well-established phylogenetic tree shows us this species is the most primitive in family Sisoridae. We first generated a chromosomal assembly of the *G. maculatum* genome. Based on comparing this genome with these of fish species from low altitude, we were able to identify genes showing strong signals of positive selection and rapid evolution. This study provides insights into both new genetic patterns that are crucial to life in this extreme alpine environment, as well as some common adaptive modes between humans and other vertebrates.

RESULTS AND DISCUSSION

Chromosomal genome assembly and annotation

Through analyses, we found many errors in previous version of assembled genome (Liu et al., 2018) (Supplementary Note in Supporting Information), thus we obtained additional PacBio data (106.3 Gb), assembled a new genome with these long reads and constructed a chromosomal level genome using HiC data we generated (Figure 1A). The final assembly resulted in a relatively higher quality genome, with a total size of 690.74 Mb, contig N50 size of 1.3 Mb, and BUSCO of 92.5% (Figure 1B; Tables S1–S8 in Supporting Information). A good chromosomal synteny relationship between genomes was obtained using Illumina HiSeq 2000 reads and PacBio long reads, suggesting high accuracy for this newly assembled genome (Figures S1–S4 in Supporting Information).

Using this chromosomal genome assembly of PacBio long reads, we identified a total of 328.76 Mb of repetitive sequences (Table S9, Figures S5 and S6 in Supporting Information), which are predominantly made up of DNA transposons and long interspersed nuclear element (LINE) (Table S10 in Supporting Information), comprising about 40.05% of the *G. maculatum* genome assembly. The percentage of LINE content in *G. maculatum* genome is higher than any other fishes that have so far been sequenced (Figure 1C; Table S11 in Supporting Information). In total, we pre-

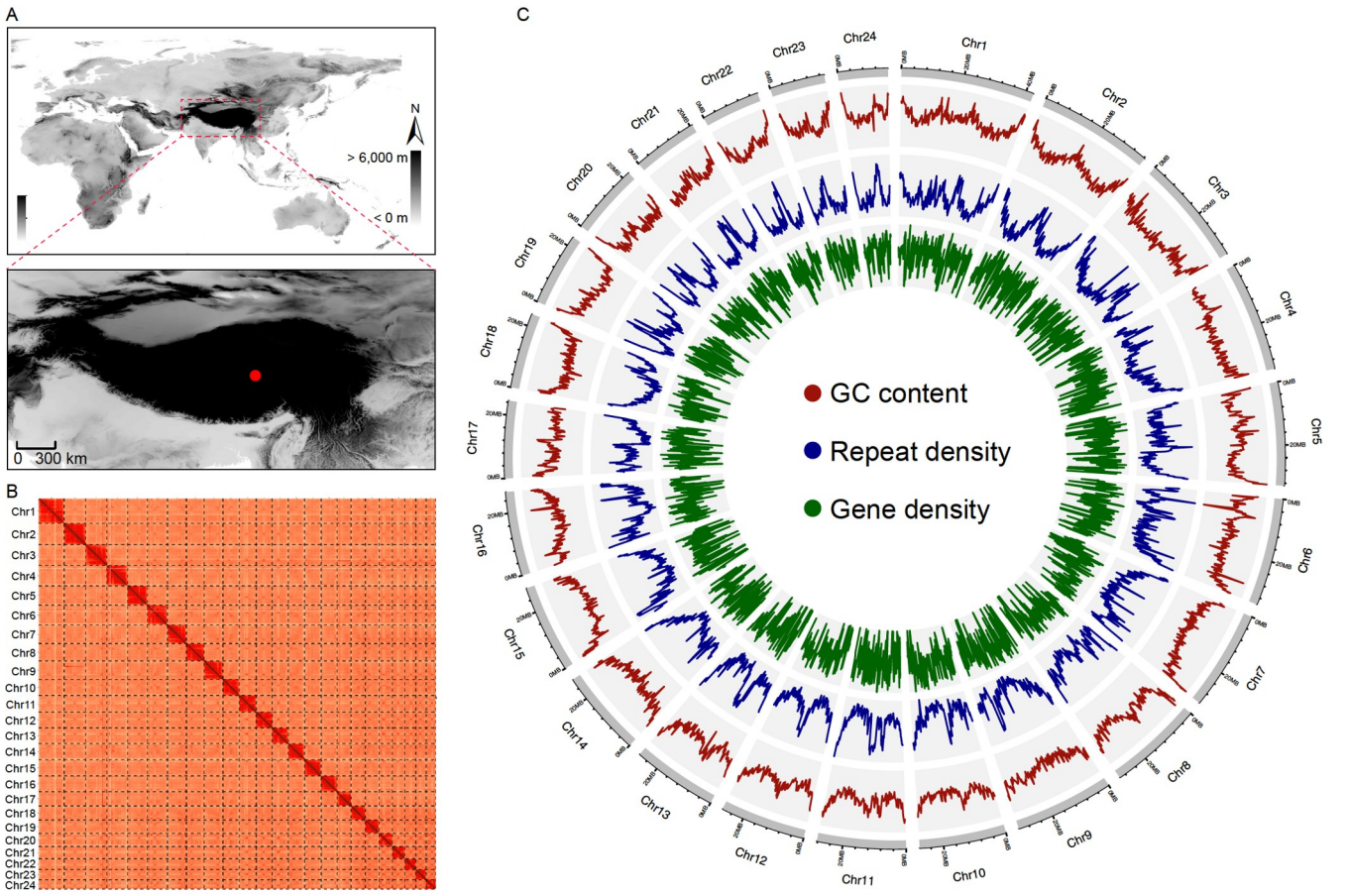

Figure 1 Sampling locality, Hi-C genome assembly, and genomic features. A, Distribution and sampling locality (red circle) of *G. maculatum*. B, Hi-C assembly of *G. maculatum* genome anchored to 24 chromosomes. C, Circos plot for mapped features of *G. maculatum* genome including GC content, Repeat density, gene density (from outside to inside).

dicted 20,911 protein coding genes in its genome based on *de novo*, homology and RNA-seq data from ten tissues (Table S12 and Figure S7 in Supporting Information). More than 96% of the predicted protein coding genes (20,248 genes) were annotated by searching public databases (SwissProt, TrEMBL, and the Kyoto Encyclopedia of Genes and Genomes (KEGG)) (Table S13 in Supporting Information). Finally, we identified sequences for 18,282 noncoding RNAs consisting of ribosomal RNAs, transfer RNAs, microRNAs, and small nuclear RNAs (Table S14 in Supporting Information).

Phylogenomics and evolutionary rates

A high-confidence phylogenetic tree (Figure 2A; Figure S8 in Supporting Information) was constructed for 14 fish species in teleosts, including *G. maculatum*, giant devil catfish, yellow catfish, channel catfish, cave fish, zebrafish, stickleback, fugu, platyfish, medaka, tilapia, seahorse, mudskipper, and spotted gar, using a genome-wide set of 2,110 one-to-one orthologous genes with coding as well as protein sequences. Among them, the giant devil catfish (*Bagarius yarelli*) is included in the same family Sisoridae with the *G. maculatum* but occurs in the low altitude in lower reaches of rivers along Tibetan plateau. All these seven datasets produced identical topologies placing *G. maculatum* as sister to the giant devil catfish (Figure S8 in Supporting Information), consistent with a previous study (Jiang et al., 2019). Estimated divergence times of *G. maculatum* with other fishes calculated using MCMCTree suggest that *G. maculatum* diverged from the giant devil catfish (family Sisoridae) approximately 27.8 million years ago, during the Oligocene (Figure 2A; Figure S9 in Supporting Information), which is generally consistent with an earlier study using mitochondrial genes (Guo et al., 2005).

To explore the demographic history of *G. maculatum*, we inferred historical population size changes using the PSMC model (Li and Durbin, 2011) based on the distribution of heterozygous sites. The demographical history of *G. maculatum* seems to be correlated with the uplifting of the Tibetan Plateau (Figure S10A in Supporting Information). The effective population size of *G. maculatum* underwent a remarkable decline from 5×10^3 to 2×10^3 at about 30 kyr ago, which occurred following one uplift phase movement of the

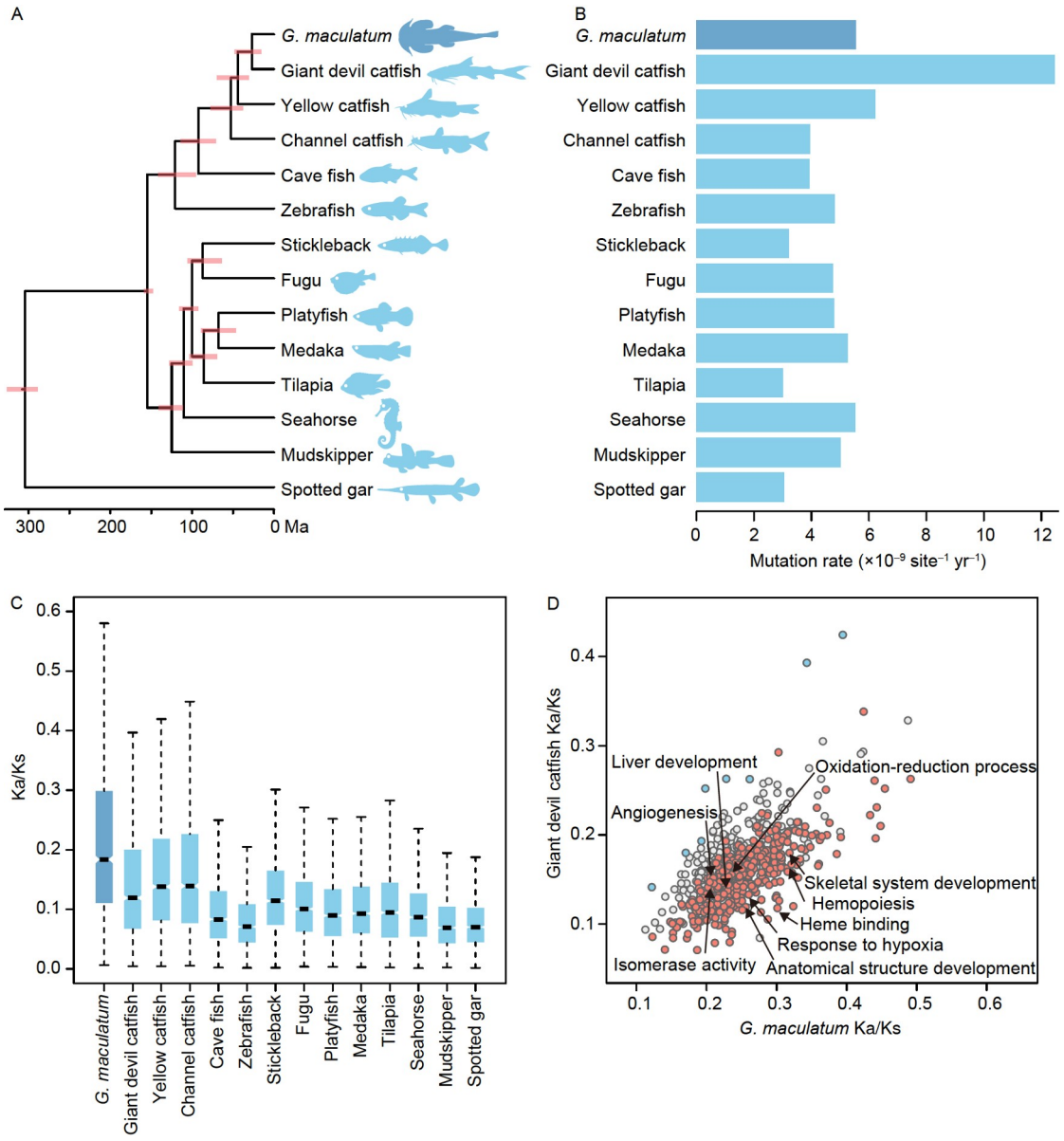

Figure 2 Evolutionary history of *G. maculatum*. A, The phylogenetic tree of 14 ray-finned fishes generated using protein sequences. The branch lengths represent divergence times, while the pink bar at each node indicates the 95% confidence interval. B, Comparison of mutation rates for the 14 ray-finned fish species based on 4D sites. C, Branch-specific Ka/Ks ratios estimated from all orthologs. D, Scatter plot of mean Ka/Ks ratio for each GO category in *G. maculatum* and the giant devil catfish. GO categories with significantly higher mean Ka/Ks ratios in *G. maculatum* (red) and giant devil catfish (blue) are highlighted.

Tibetan Plateau. This Gonghe movement (~0.15 Ma) uplifted the Tibetan Plateau to an average of above 4,000 m. Considering that it is very implausible to use the PSMC approach to estimate the effective population size when the time is later than 10,000 years ago, we further confirmed this result by obtaining the data from multiple individuals (Xiao et al., 2021) and estimating population history using MSMC approach (Schiffels and Durbin, 2014). These results showed that the population size of *G. maculatum* decreases from 80,000 to 20,000 years ago (Figure S10B in Supporting Information). Therefore, the pattern of population decline in *G. maculatum* is believed to be a response to a common background following this intense uplift phase of the plateau.

The branch length of *G. maculatum* was about one half of the giant devil catfish in the maximum-likelihood tree (Figure S8 in Supporting Information), suggesting a slower evolutionary rate compared with the giant devil catfish. This result was further confirmed by mutation rate analysis using the fourfold degenerate (4D) sites (Figure 2B). To determine whether the protein evolutionary rate of *G. maculatum* is also slower, we compared the Ka/Ks (nonsynonymous-synonymous) substitution ratios (ω) for each fish branch. Interestingly, the *G. maculatum* branch had a significantly higher ratio of Ka/Ks than other fishes used in this study (Figure

2C), independent of the gene sets analyzed (Figure S11 in Supporting Information), demonstrating accelerated protein evolution in the *G. maculatum* lineage. This pattern is very similar to deep-sea snail fish (Wang et al., 2019). Maybe the extreme environment pressure influences the organisms in a very similar mechanism. The environment on the Tibetan Plateau is very inhospitable to live for endemic species, including low oxygen availability, low temperatures and strong ultraviolet radiation. Similar to this, the environment in the hadal zone is even bad with high hydrostatic pressure, darkness, limited food resources, low temperatures and hypoxia. These harsh environments should drive endemic species to evolve faster to adapt to their living environment.

To examine the gene ontology (GO) categories showing accelerated evolution in the *G. maculatum* lineage, we averaged the Ka/Ks ratios for each GO category with more than ten orthologs. The number of GO categories having higher Ka/Ks ratios in *G. maculatum* was significantly larger than the number of GO categories having higher Ka/Ks ratios in the giant devil catfish (619 *vs*. 8), confirming accelerated evolution in *G. maculatum* (Figure 2D). Furthermore, many GO categories involved in skeletal development, energy metabolism, and hypoxia response showed accelerated protein evolution in *G. maculatum*, such as "Skeletal system development", "Oxidation-reduction process", and "Hemopoiesis" (Figure 2D; Table S15 in Supporting Information).

Adaptation to torrent environment

The uplift of the Tibetan Plateau was considered as the most dramatic tectonic event in recent geological history and created many mountain rivers (Liu and Dong, 2013). These swift-flowing streams impose strong challenges for the life of fishes, as they must attempt to avoid being washed away. *G. maculatum* has evolved several specific adaptations to these fast currents, including a strongly flattened head and body, and greatly enlarged wing-like pectoral and pelvic fins, which together form a specialized suction apparatus (Guo et al., 2005). Gene family analyses of the 14 fishes included in our study identified 100 significantly expanded gene families in *G. maculatum* (Table S16 in Supporting Information). Among these expanded gene families, the *CFDP2* (craniofacial development protein 2) gene family grew notably, which might suggest a role in the development of their flattened head. We found several other genes appear to be under positive selection in *G. maculatum*, including *tbx3a* (Govoni et al., 2006) and *gli3* (Hui and Joyner, 1993), which were reported to play an important role in limb development (Table S17 in Supporting Information). Furthermore, several genes involved in fin development were identified as rapidly evolving genes, including *tbx4* and *hoxd12a* (Table S18 in Supporting Information).

To evaluate the functional role of *hoxd12a* in fish fins, we first performed *in situ* hybridization of *hoxd12a* gene from 36–72 hpf (hours post fertilization) in wild-type zebrafish to examine whether *hoxd12a* expression plays a role in the early development of the fin fold. Our results showed that *hoxd12a* is actively expressed in the pectoral fin at 36 and 48 hpf and becomes stronger at 72 hpf (Figure 3A), implying functions in the early development of the pectoral fin (Ahn and Ho, 2008). To further explore the function of *hoxd12a* gene, we inactivated the *hoxd12a* gene from the zebrafish genome using CRISPR/Cas9. A mutant line was obtained with 4 base pair (bp) deletion and 14 bp insertion in the second exon of *hoxd12a* (Figure 3B). In adult fins of zebrafish (~120 dpf (days post fertilization)), we compared the ratio between pectoral fin length and total length and found that the ratio is significantly lower in *hoxd12a*-null zebrafish ($hoxd12a^{-/-}$) compared with the wild-type zebrafish ($hoxd12a^{+/+}$) (each type contains 30 individuals), suggesting shorter pectoral fin in $hoxd12a^{-/-}$ zebrafish (Student's *t*-test, P=0.034) (Figure 3C). We also investigated the number of pectoral fin rays and found that *hoxd12a*-null zebrafish have significantly fewer pectoral fin rays than wild-type zebrafish (Student's *t*-test, P=0.022) (Figure 3D). Taken together, these results show that *hoxd12a* has an important role in the establishment of the number and length of pectoral fin rays in zebrafish and might suggest that an accelerated rate of evolution of *hoxd12a* in *G. maculatum* may be related to the enlarged wing-like pectoral fin morphology adapted to fast waters.

Adaptation to low temperature

With an average elevation of more than 4,000 m above sea level, the mean annual air temperature of the Tibetan Plateau is as low as −1.7°C (Kang et al., 2010), which imposes a severe challenge to the living condition of endemic organisms. Consistent with the low temperature, we found significant expansion of *ZP4* gene (zona pellucida sperm-binding protein 4) in *G. maculatum* (Table S16 in Supporting Information), which is a common phenomenon in cold-adapted organisms, such as the Antarctic notothenioid fishes (Cao et al., 2016). Zona pellucida proteins are constituents of the chorion and have a role in egg-sperm recognition during fertilization, which forms a protective matrix surrounding the egg (Han et al., 2010). Expansion of *ZP4* gene in *G. maculatum* may protect their eggs during fertilization in the cold water of the Tibetan Plateau.

Mitochondria function as the energy generators of the cells (Monaghan and Whitmarsh, 2015). TRMU (tRNA 5-methylaminomethyl-2-thiouridylate methyltransferase), a tRNA modifying enzyme, is responsible for the biosynthesis of the $\tau m^5 s^2 U$ of mitochondrial $tRNA^{Lys}$, $tRNA^{Glu}$ and $tRNA^{Gln}$ (El Yacoubi et al., 2012). We identified three amino acid replacements of TRMU in *G. maculatum* (Figure S12 in

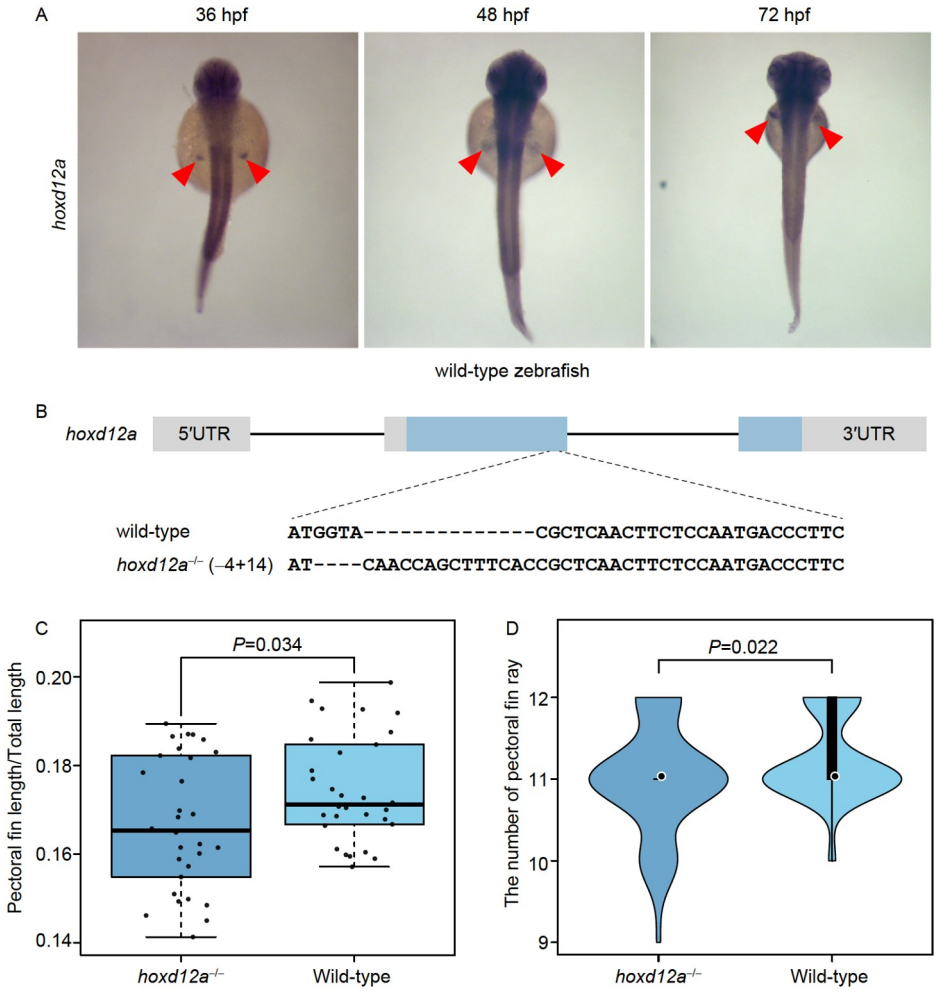

Figure 3 Functional analyses of *hoxd12a* gene in zebrafish. A, Expression patterns of *hoxd12a* gene in pectoral fin of zebrafish from 36 to 72 hpf. Arrowheads indicate the position of pectoral fin. B, Schematic representation of CRISPR/Cas9 target site at exon 2 of *hoxd12a* gene, producing a 4 bp deletion and 14 bp insertion in exon 2. C, Comparison of the ratio between pectoral fin length and total length of *hoxd12a*-null mutant and wild-type zebrafish. D, Comparison of the number of pectoral fin ray of *hoxd12a*-null mutant and wild-type zebrafish.

Supporting Information), and the branch-site model implemented in PAML (Yang, 2007) detected significant positive selection in *TRMU* gene in *G. maculatum* (Table S17 in Supporting Information).

To examine the potential functional differences of the TRMU protein in *G. maculatum* compared with all of the lowland fishes, we expressed the *G. maculatum*-specific allele of *TRMU* (pEGFP-N1-*TRMU*), as well as the allele found in all of the lowland fishes including pEGFP-N1 (negative control), the mutants *TRMU* (p.S188T), *TRMU* (p.Q206E), *TRMU* (p.M221I) and *TRMU* (p.S188T+Q206E+M221I), in human embryonic kidney cells (HEK293T) using plasmids. Western blot analysis showed that the endogenous expression of *TRMU* and the *TRMU* mutants was similar to each other (Figure S13 in Supporting Information). Also, all of them were expressed in mitochondria (Figure S14 in Supporting Information). Therefore, the *G. maculatum*-specific allele of *TRMU* and mutant *TRMU* alleles showed similar protein stability and expression location.

To assess whether the amino acid replacement of TRMU in *G. maculatum* affects the steady-state levels and the extent of the 2-thiouridine modification at position 34 in tRNAs, we determined the 2-thiouridylation levels of tRNAs by isolating total RNAs from HEK293T cell line of pEGFP-N1-*TRMU* and the *TRMU* mutants. Our results showed that while the tRNAs for Lys, Glu, and Gln are almost completely modified in HEK293T cells, the mutants *TRMU* (p.M221I) resulted in a severe reduction in the 2-thiouridylation of mitochondrial $tRNA^{Lys}$ and $tRNA^{Gln}$, with residual levels of about 40% of total for $tRNA^{Lys}$ and $tRNA^{Gln}$. Similar results were also found in the *TRMU* mutants (p.S188T+Q206E+M221I) (Figure 4A and B). Therefore, the site M221I was clearly involved in the synthesis of 2-thiouridine at the wobble position of the mitochondrial tRNAs for Lys and Gln.

To further investigate whether amino acid replacement of

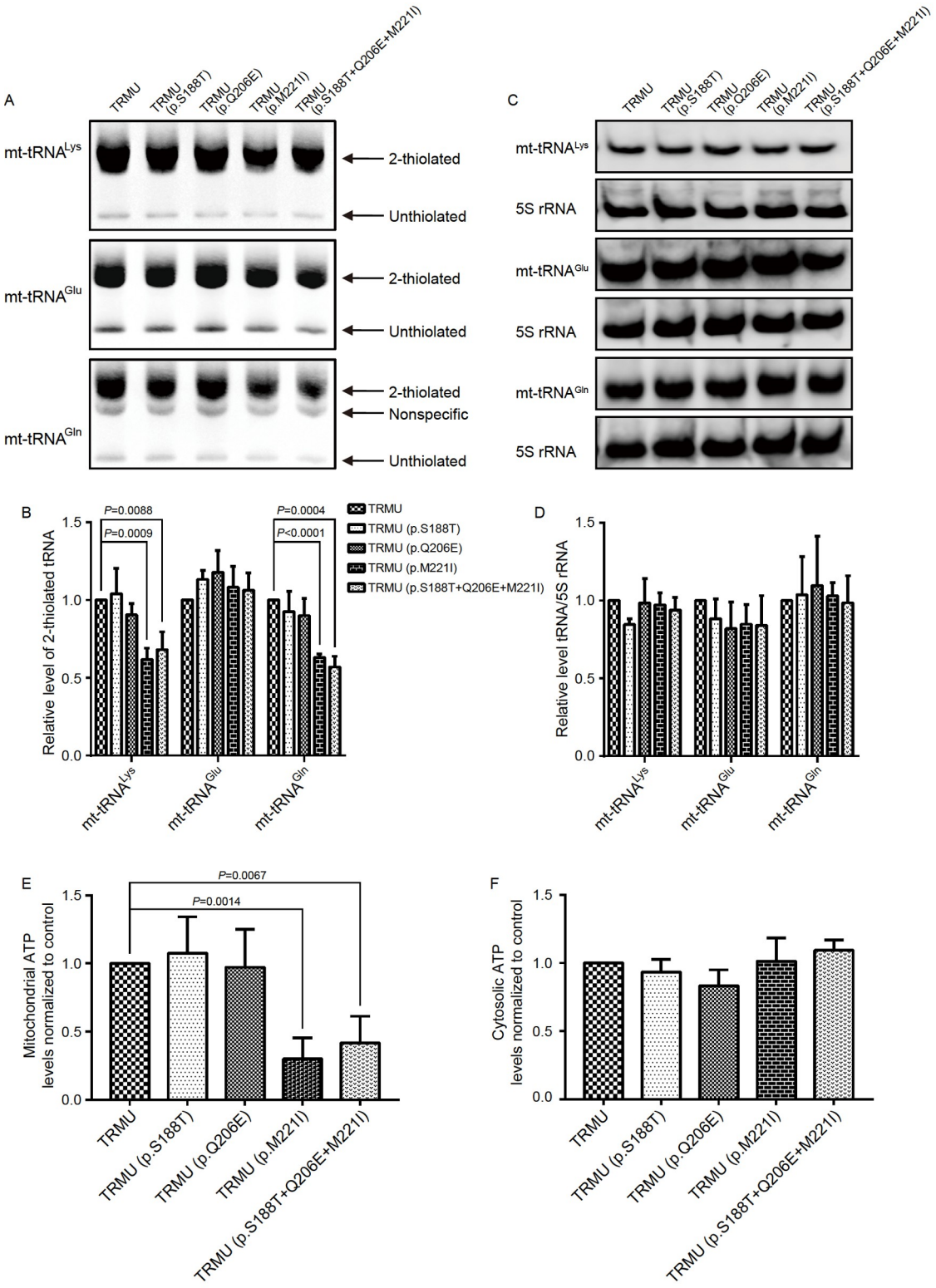

Figure 4 Functional analyses of *TRMU* gene and the amino acid mutants. A, Northern blotting with adding APM to the gels to separate thiolated and unthiolated tRNA was performed and probed for mt-tRNA^{Lys}, mt-tRNA^{Glu,} and mt-tRNA^{Gln} in HEK293T cell lines with *TRMU* and *TRMU* mutants. B, Quantification of the Northern blotting shows the percentage of 2-thiolated tRNA levels. The calculations were based on three independent experiments. The error bars indicate standard errors; *P* indicates the significance, according to Student's *t*-test. C, Northern blotting with a denaturing polyacrylamide gel and probed for mt-tRNA^{Lys}, mt-tRNA^{Glu}, mt-tRNA^{Gln}, and 5S rRNA respectively in HEK293T cell lines with *TRMU* and *TRMU* mutants. D, Quantification of the Northern blotting shows the average relative levels tRNAs content that were normalized to the 5S rRNA respectively. E and F, Measurement of mitochondrial and cytosolic ATP levels. Average cytosolic ATP level (presence of oligomycin for inhibition of the mitochondrial ATP synthesis) and mitochondrial ATP level (subtraction of cytosolic ATP level from total cellular ATP levels) are shown. The calculations were based on three independent experiments.

TRMU in *G. maculatum* ablates the metabolism of tRNA, we subjected total RNAs from mutant and control TRMU in HEK293T cell line to Northern blots and hybridized them with bio-labeled oligodeoxynucleotide probes for mitochondrial $tRNA^{Lys}$, $tRNA^{Glu}$ and $tRNA^{Gln}$ and cytoplasmic $tRNA^{Lys}$. These results showed that amino acid replacement in TRMU caused no significant difference in all three mitochondrial tRNAs (Figure 4C and D). Mitochondria are highly efficient energy-transforming organelles that convert energy stored in nutrients into ATP. The effect of amino acid replacement in TRMU on the capacity of oxidative phosphorylation was further evaluated by measuring the cellular and mitochondrial ATP levels. The levels of mitochondrial ATP in the mutant HEK 293T cell line of *TRMU* (p.M221I) and *TRMU* (p.S188T+Q206E+M221I) were significantly decreased (Student's *t*-test, $P<0.01$) (Figure 4E and F). Taken together, our results demonstrate that M221I may be a key site to increase mitochondrial ATP via regulating the 2-thiouridylation levels of mitochondrial $tRNA^{Lys}$ and $tRNA^{Gln}$, suggesting that this amino acid replacement in *G. maculatum* represents an adaptation to low temperature.

Hypoxia adaption

In high-altitude environments, *G. maculatum* is exposed not only to low temperatures, but also to hypoxia. VHL, a key regulator of HIF (hypoxia-inducible factor), is considered to play a vital role in controlling the hypoxia signaling pathway (Maxwell et al., 1999). We detected three amino acid replacements in VHL of *G. maculatum*, and the branch-site model implemented in PAML (Yang, 2007) identified significant positive selection of the *VHL* gene in *G. maculatum* (likelihood ratio test (LRT), $P=0.017$; Figure 5A). Among these three amino acid replacements, the valine (V) to glutamine (Q) replacement at position 102 was predicted to be "damaging" to function by PolyPhen-2, and the Codeml function of PAML found that this position was likely to be under positive selection for change ($P=0.988$; Figure 5A).

We also performed three-dimensional (3D) structure simulations to identify the possible effects of the amino acid replacement at position 102 of VHL using PyMOL and found that the Q102V mutation in VHL was located in the functional β domain (Figure 5B). To explore the potential functional effects of this amino acid replacement in VHL, we first examined whether VHL (V102Q) has a different effect in regulating HIF-α protein expression levels compared with VHL. Using three plasmids including HA-VHL (*Glyptosternum* allele), HA-VHL (V102Q) (low altitude allele) and pCMV-HA (negative control), our results indicated that the amino acid replacements in VHL did not affect the protein expression level of HIF-α regardless of normoxia or hypoxia (Figure 5C).

To further examine whether the amino acid replacements in VHL could affect HIF-α at the level of transcriptional activity, we co-transfected HA-VHL or HA-VHL (V102Q) together with HRE-luciferase reporter and CMV-Renilla luciferase reporters into HEK293T cells. Over expression of VHL or VHL (V102Q) significantly inhibited the transcriptional activity of the HIF-α paralogs in normoxia condition. However, both VHL and VHL (V102Q) enhanced the HIF-α transcriptional activity in hypoxia condition. Interestingly, the VHL of *G. maculatum* was less effective at enhancing the transcriptional activity of HIF-α compared with VHL (V102Q) (Figure 5D), suggesting much lower transcriptional activity in the *G. maculatum* protein variant.

In conclusion, this study represents the first example investigating phenotypic adaptation to the Tibetan Plateau by integrating genomic and functional evidences in fishes. Our results have identified several convergent genetic mechanisms in high-altitude adaptation between fishes and other organisms.

MATERIALS AND METHODS

Genome sequencing

A single female *G. maculatum* individual that lived above 3,800 m in the Shigatse City of Tibet in China was used as the template for sequencing. Genomic DNA was isolated from the muscle tissue of the back using Puregene Tissue Core Kit A (Qiagen, USA), which was used to construct ten libraries including short-insert and long-insert libraries. For four short inserts of pair-end libraries, 6 μg genomic DNA was fragmented to the desired insert size, end-repaired and ligated to Illumina paired-end adaptors. Short sizes of 370, 480, and 550 bp of the ligated fragments were selected on agarose gels and were purified by PCR amplification to yield the corresponding libraries. For six long insert sizes (1, 4, 6, 10, 15, and 20 kb) of library construction with mate-pair, 60 μg DNA was used to circularize, and digest linear DNA. A fragmented circularized DNA and purified biotinylated DNA were performed and adaptor ligation. The construction of these libraries was finished on the Illumina HiSeq 2000 sequencing platform at Yunnan Ice Harber Bioinformatics Technology Co., Ltd (Kunming, China).

To obtain chromosomal-level genome assembly, Hi-C libraries were constructed and sequenced using the HiSeq X-Ten platform to obtain 150 bp pair-end reads (76 Gb). Hi-C library construction was performed using the following protocol: the fish muscle was fixed in 1% formaldehyde solution. The nuclear chromatin was obtained and digested using HindIII (NEB, USA). Then, the overhangs resulting from HindIII digestion were blunted with the Klenow enzyme (NEB) and bio-14-dCTP (Invitrogen, USA). After dilution and religation with T4 DNA ligase (NEB), DNA was extracted and sheared to a size of 350–500 bp with a

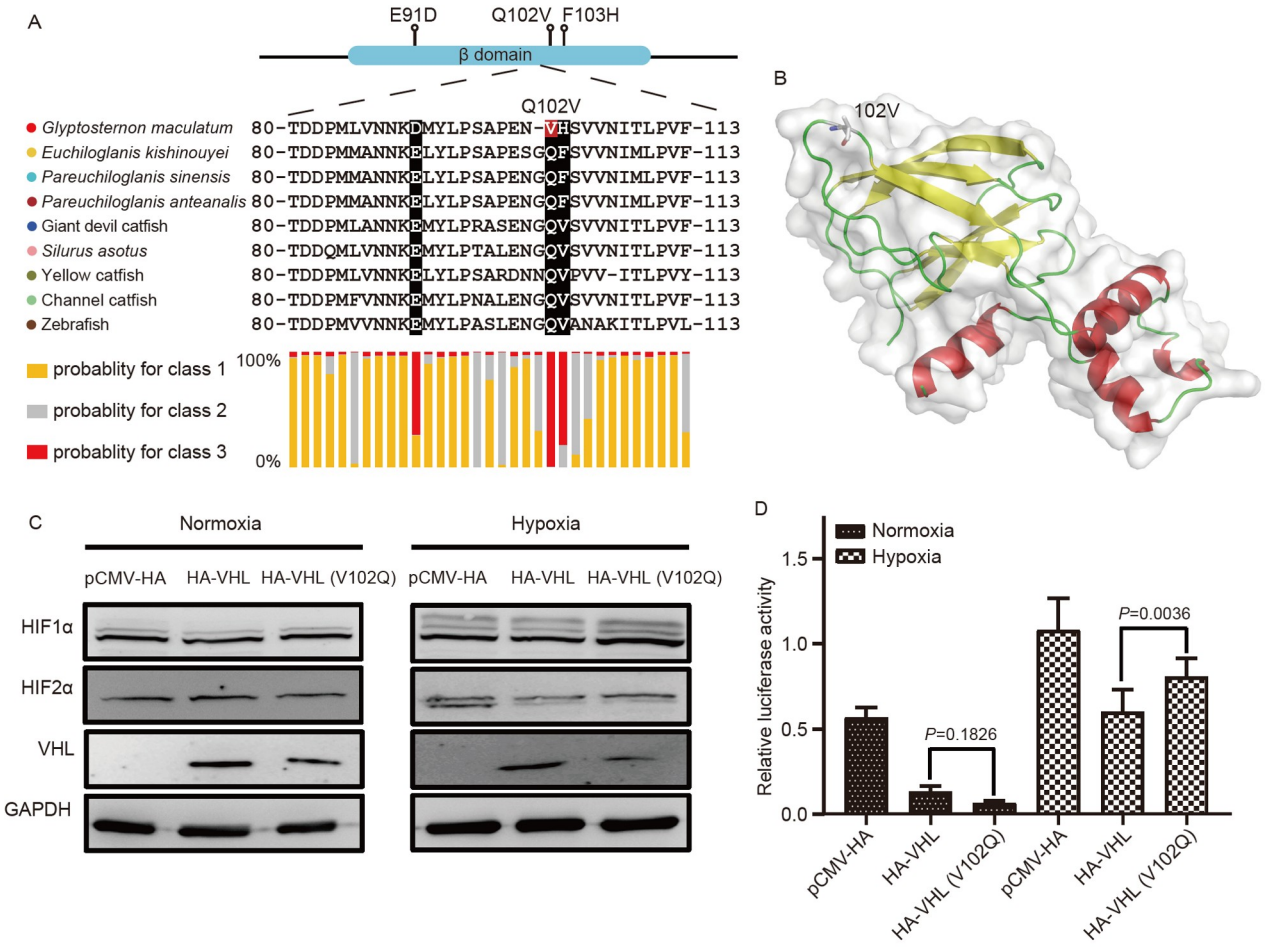

Figure 5 VHL replacements identified in *G. maculatum* and functional assays of the identified variants. A, Alignment of VHL amino acid sequences shows specific amino acid mutations in the *G. maculatum VHL* gene. B, 3D structure simulation of VHL shows the position of *G. maculatum* specific variants. C, Western blot shows the expression of different HIF-αs in HEK293T cells under normoxic and hypoxic conditions regulated between VHL and VHL (V102Q). GAPDH is shown as a loading control. D, Transcriptional activity of HIF-αs under normoxic and hypoxic conditions regulated between VHL and VHL (V102Q).

Bioruptor (Diagenode, Belgium). The biotin-labeled DNA fragments were enriched by utilizing streptavidin beads (Invitrogen) to further finish library preparation.

Total RNAs of seven tissues (heart, liver, kidney, gill, spleen, gonad, brain) were extracted using the TRIzol kit (Life Technologies, USA). Polyadenylated RNA was selected using oligo (dT) purification and reverse transcribed to complementary DNA (cDNA) and sequenced on the Illumina HiSeq 2000 sequencing platform.

De novo assembly

Illumina PCR adapter reads and low-quality reads from the paired-end and mate-pair library were filtered by the step of quality control using an in-house Perl script. For example, reads were filtered based on quality: reads with more than 10% Ns, reads with more than 40% of the read length below Q7 (Phred score⩽7), paired-end reads with more than10 bp overlap between the two ends, reads with more than 10 bp overlap with an adapter sequence and with a maximum of 2 bp mismatches. Meanwhile, all duplicate reads were filtered. After filtering the raw reads, reads were also corrected by the SOAPec (v2.01) software in the SOAP package (Luo et al., 2012) (http://soap.genomics.org.cn) using "Kmer-Freq_HA -k 27 -r 70" and "Corrector_HA -k 27", which 27 mers with low frequency less than 3 were removed in order to minimize the influence of sequencing errors.

All the clean and corrected reads with good quality were assembled using SOAPdenovo 2.04 (Luo et al., 2012) with default parameters. All of the small insert sizes (<1 kb) were assembled into distinct contigs, reads from the long-insert libraries (⩾4 kb) were aligned to the contig sequence according to the sequence overlapping, and then all of the mate-pair libraries were used to construct the contigs into scaffolds from the 4 to 20 kb insert size step by step. All usable reads to form a connection were mapped back to contig sequences to construct scaffolds with at least 5 read pairs for small insert size reads and 7 read pairs for mate-pair

reads required to form a connection. To fill the gaps within the scaffold and improve the contig N50 length, we retrieved read pairs that had one read well aligned on the contigs and the other read located in the gap region, and GapCloser was used to fill them.

For the PacBio reads, we first filtered low-quality reads and the remaining high-quality reads were first corrected using the program Canu (version 1.5) (Koren et al., 2017), and then assembled using the genome assembler Flye (version 2.6) (Kolmogorov et al., 2019) with default parameters. The initial draft assembly was polished for three rounds using the raw PacBio reads with Racon (version 1.2.1) (Vaser et al., 2017), and further corrected for three rounds using the Illumina reads with Pilon (version 1.21) (Walker et al., 2014) to obtain a consensus assembly.

To generate a chromosomal-level assembly of the genome, the filtered Hi-C reads were first aligned to the assembled genome using the bowtie2 (version 2.2.5) (Langmead et al., 2009) end-to-end algorithm, then the assembled contigs were clustered, ordered and directed onto the pseudo-chromosomes using Lachesis (Burton et al., 2013) with default parameters. Finally, the pseudo-chromosomes predicted by Lachesis were cut into bins with equal length of 250 kb and used to construct a heatmap based on the interaction signals.

Genome annotation

We annotated the repeat elements using both *de novo* and homology-based methods. For *de novo* predictions, RepeatModeler (version 1.0.5) (Tarailo-Graovac and Chen, 2009) was used to construct a *de novo* transposable element library, which was then used to predict repeats with RepeatMasker (version 4.0.6) (Tarailo-Graovac and Chen, 2009). For homology-based analysis, transposable elements were identified using RepeatMasker (version 4.0.6) and RepeatProteinMask (version 1.36) with the RepBase transposable element library (Jurka et al., 2005). Finally, we further identified tandem repeats using TRF software (version 4.0.4) (Benson, 1999).

We used three approaches to predict and retrieve the coding genes: *ab initio* gene prediction, homolog prediction, and RNA-seq evidence. *Ab initio* gene prediction methods including Augustus (version 3.2.1) (Stanke et al., 2008), GeneID (version 1.4) (Alioto et al., 2013), GlimmerHMM (version 3.0.4) (Majoros et al., 2004) and SNAP (version 2013-11-29) (Korf, 2004) were used to generate *ab initio* predictions with internal gene models. For homology-based gene prediction, we aligned protein sequences from 13 fishes (Ensembl release 87) against the repeat-masked *G. maculatum* genome using TBLASTN with *E* value less than 1×10^{-5} and Genewise (version 2.4.1) (Birney et al., 2004) for aligning and predicting gene structures. For RNA-seq prediction, the RNA-seq reads were aligned against the reference genome using TopHat (Kim et al., 2013), and the transcripts were assembled using Cufflinks (Trapnell et al., 2012). All the above evidence was merged to form a comprehensive consensus gene set using the EvidenceModeler (EVM) software (Haas et al., 2008). Meanwhile, PASA was used to update the EVM consensus predictions, adding UTR annotations and models for alternatively spliced isoforms.

With a non-redundancy gene dataset, we blasted against the SwissProt and TrEMBL protein databases for the genes annotation, and the motifs and domains were determined by using InterProScan (Jones et al., 2014) against protein databases such as ProDom, PRINTS, Pfam, SMART with *E* values cutoff of 1×10^{-5}. The functional classification of GO categories was performed using Blast2GO (Conesa et al., 2005). KEGG pathway annotation was performed using KOBAS (2.0) (Xie et al., 2011).

Phylogenetic analysis

Protein sequences of 14 fishes (*Glyptosternum maculatum*; giant devil catfish, *Bagarius yarrelli*; yellow catfish, *Pelteobagrus fulvidraco*; channel catfish, *Ictalurus Punctatus*; cave fish, *Astyanax mexicanus*; zebrafish, *Danio rerio*; stickleback, *Gasterosteus aculeatus*; fugu, *Takifugu rubripes*; platyfish, *Xiphophorus maculatus*; medaka, *Oryzias latipes*; tilapia, *Oreochromis niloticus*; seahorse, *Hippocampus comes*; mudskipper, *Boleophthalmus pectinirostris*; and spotted gar, *Lepisosteus oculatus*) were obtained from previous studies and Ensembl database (Release 87). We used OrthoFinder v2.3.4 with default parameters and RBH method to cluster the homologous genes for these 14 fish genomes (Emms and Kelly, 2015). The protein sequences of single-copy orthologous genes were aligned using MUSCLE (Edgar, 2004) with the default parameters. We then filtered the saturated sites and poorly aligned regions using trimAl (Capella-Gutiérrez et al., 2009). We generated various datasets for the following phylogenetic analyses, including the 4D sites, the full amino acid sequences, and the corresponding coding sequence alignments.

The phylogenetic trees were reconstructed using RAxML version 8.1.19 with parameters "-m GTRGAMMA -f a -x 271828 -N 100 -p 12345" (Stamatakis, 2014). For the species tree, we built individual gene trees for each ortholog using RAxML, and these individual gene trees were further summarized by ASTRAL (v.4.7.12) (Stamatakis, 2014). The Bayesian relaxed-molecular clock (BRMC) method, implemented in the MCMCTree program (Yang, 2007), was used to estimate the divergence time between different species. Two calibration time points based on fossil records, *O. latipes*-*T. nigroviridis* (~96.9–150.9 Mya), and *D. rerio*-*G. aculeatus* (~149.85–165.2 Mya) (http://www.fossilrecord.net/dateaclade/index.html), were used as constraints in the MCMCTree estimation.

Demographic history

The demographic history of *G. maculatum* was inferred by applying the Pairwise Sequentially Markovian Coalescence model (Li and Durbin, 2011) to the complete diploid genome sequences of five libraries. The approach of PSMC was based on a Hidden Markov Model (HMM) (Eddy, 1996). The distribution of time to TMRCA (the most recent common ancestor) between two alleles in an individual can be related to the history of population size fluctuation. The estimated generation time (g) was set to 7 years. PSMC modelling was done using a bootstrapping approach, with sampling performed 100 times to estimate the variance of the simulated results.

Expansion and contraction gene families

We used CAFE v3.1 (De Bie et al., 2006) to analyze the expansion and contraction of gene families in *G. maculatum* genome with the results from OrthoFinder. Then *P*-value of each gene family was determined using the results from CAFE. We further filtered the gene families with gene numbers higher than 100. Gene families with a *P*-value lower than 0.01 were taken as having undergone significant expansion or contraction.

Detection of positively selected genes and rapidly evolving genes

Single-copy orthologous genes obtained in the gene family were used in the detection of positively selected genes and rapidly evolving genes. Codeml program in the PAML package (version 4.8) (Yang, 2007) was employed to estimate the lineage specific evolutionary rate for each branch. The branch-site model in Codeml was used to detect positively selected genes with the *G. maculatum* specified as the foreground branch. The LRT was performed to test whether the branch-site model containing positively selected codons (omega>1) fits better than the null model, which only includes the neutral selection or negative selection (omega⩽1). The *P* values for model comparison were computed based on chi-square statistics. The genes with an LRT $P<0.05$ were treated as positively selected genes.

The branch model in Codeml was used to identify rapidly evolving genes using the same orthologous genes as above. The foreground branch (the *G. maculatum* lineage) was found to exhibit a significantly higher omega (regardless of whether it is greater than 1) than the background branch (the other lineages) using the LRT test. The genes with an LRT $P<0.05$ were treated as rapidly evolving genes in *G. maculatum*. GO functional enrichment analyses for both positively selected genes and rapidly evolving genes were carried out by DAVID (Huang et al., 2009).

Functional assay of *Hoxd12a*

Experimental fish and maintenance

AB wild-types of zebrafish were raised, maintained, reproduced, and staged according to standard protocol. All zebrafish were maintained under standard conditions at 28.5°C in the system water with a circulated water system with a 14 h light and 10 h dark cycle and fed newly hatched brine shrimp (*Artemia salina*). All animal experiments were approved by the ethical board of the Animal Care and Use Committee of the Institute of Hydrobiology, Chinese Academy of Sciences (protocol number IHB-2017001).

In situ hybridization

Digoxigenin-labeled antisense RNA probes were synthesized by *in vitro* transcription. Probes were synthesized with digoxigenin (DIG)-labeled antisense RNA probes specific to zebrafish *Hoxd12a* (forward primer: GTGGCTACTCTCAGCCGTTT; reverse primer: TAATACGACTCACTATAGGGCTGGGCGCTCAGTGACGGC). Whole mount *in situ* hybridization (WISH) was performed as described (Thisse and Thisse, 2008).

Generation of Hoxd12a-null zebrafish

Disruption of *Hoxd12a* in zebrafish was accomplished via CRISPR/Cas9 technologies. Guide RNA (sgRNA) generated against the *Hoxd12a* target sequence in the second exon, was designed using the CRISPR design tool (http://crispr.mit.edu). The zebrafish codon-optimized Cas9 plasmid was digested with XbaI and then purified and transcribed using the T7 mMessage mMachine Kit (Ambion, USA). pMD49T gRNA vector was used for amplifying the sgRNA template. The primers for amplifying *Hoxd12a* gRNA are 5′-TTGTAATACGACTCACTATAGGAGAAGTTGAGCGTACCATGTTTTAGAGCTAGAAATAGC-3′ and 5′-AAAAAAAGCACCGACTCGGTGCC- ACT-3′. sgRNA was synthesized using the Transcript Aid T7 High Yield Transcription Kit (Fermentas, USA). 500 pg Cas9 mRNA and 50 pg gRNAs were co-injected at a one-cell stage for each embryo. The mutations were detected by sequencing of targeting sites. The heterozygous males and females in the F1 population were crossed to generate an F2 population that contained *Hoxd12a* homozygotes.

Bone staining and observation with light microscopy

The staining procedure was performed according to a standard method. Briefly, three groups of adult young zebrafish (three-month-old zebrafish) were collected and fixed in 4% paraformaldehyde in phosphate buffered saline (PBS) at 4°C overnight. After washing with milli-Q water and Tris-Buffered Saline and Tween20 (TBST), Cartilage was stained with 0.02% alcian blue in 30% acetic acid/70% EtOH for 6 h, then washed with Ethanol solution with different con-

centrations and milli-Q water. For bone staining by Alizarin, bleaching in 1% H_2O_2/1% KOH for 20 min and digesting in 1 mg mL^{-1} trypsin in 60% saturated sodium borate for 20 min, then staining with freshly prepared 0.04 mg mL^{-1} Alizarin solution overnight. Samples were washed briefly in 1% KOH for 30 s for four times before being photographed with a light microscope.

Measurement of the pectoral fin length and distal radials number

Pectoral-fin length and body length of the individuals were measured (to the nearest 0.01 mm) and relative length was calculated as pectoral-fin length to body length ratio. The number of distal radials was counted with a light microscope.

Functional assay of *TRMU*

Plasmid constructs and mutants

G. maculatum gene *TRMU* was PCR-amplified from *G. maculatum* cDNA using PCR. Amplified genes were subcloned into pEGFP-N1 (Clontech, USA) vetors to generate pEGFP-N1-*TRMU*. The mutants *TRMU* (p.S188T), *TRMU* (p.Q206E), *TRMU* (p.M221I) and *TRMU* (p.S188T+Q206E+M221I) were amplified by PCR and subcloned into pEGFP-N1 (Clontech).

Cell culture and transfection

HEK293T cells were originally obtained from American Type Culture Collection (ATCC) and maintained in Dulbecco's modified Eagle's medium (HyClone, USA) supplemented with 10% fetal bovine serum (FBS; HyClone) and 1% penicillin streptomycin solution (HyClone) at 37°C in a humidified atmosphere incubator containing 5% CO_2. HEK293T cells were transfected with the constructed plasmids using VigoFect (Vigorous Biotechnology, Beijing, China) following the manufacturer's instructions.

Antibodies and Western blot analysis

The antibodies used were as follows: anti-eGFP antibody (1:2,000 for IB analysis), anti-GAPDH (1:5,000 for IB analysis). Western blot was performed as previously described. HEK293T cell lysates were homogenized in RIPA reagent (Invitrogen) using a homogenizer. Twenty μg of total cellular proteins were electrophoresed through 12% sodium dodecyl sulfate (SDS)-polyacrylamide gels and then transferred to a polyvinyl difluoride (PVDF) membrane. After transfer, the membrane was incubated in TBST and 5% nonfat dry milk for 1 h. Then, the membrane was incubated with primary antibody overnight at 4°C then with secondary antibody. We used a Fujifilm LAS4000 mini-luminescent image analyzer to photograph the blots.

Subcellular localization of TRMU

To determine the subcellular localization of *TRMU* and the mutants *TRMU* (p.S188T), *TRMU* (p.Q206E), *TRMU* (p.M221I) and *TRMU* (p.S188T+Q206E+M221I), pEGFP-N1-*TRMU*, expressing *TRMU*-EGFP fusion protein, was transfected into the HEK293T cell line for the transient expression. MitoTracker®Red CMXRos (Invitrogen) was used to label mitochondria. HEK293T cells were then viewed under a Leica confocal fluorescence microscope.

Mitochondrial tRNA Northern analysis

Total RNAs were isolated from fish using Total RNA Kit II (Omega, USA). The presence of thiouridine modification in the tRNAs was verified by the retardation of electrophoretic mobility in a polyacrylamide gel that contains 0.1 mg mL^{-1} (N-acryloylaminophenyl) mercuric chloride (APM). Ten μg of total cellular RNAs were electrophoresed through a 15% polyacrylamide/7.5 mol L^{-1} urea gel in Tris-borate-ethylenediaminetetraacetic acid (EDTA) buffer (TBE) (after heating the sample at 65°C for 10 min) and then electroblotted onto a positively charged nylon membrane (GE Healthcare, USA) for the hybridization analysis with oligodeoxynucleotide probes. Oligodeoxynucleosides used for biotin-labeled probes (Table S19 in Supporting Information) were mitochondrial $tRNA^{Lys}$, $tRNA^{Glu}$, $tRNA^{Gln}$, cytoplasmic $tRNA^{Lys}$ and 5S rRNA. The hybridization reactions were carried out in a mixture of 5× SSC (SSC is standard saline citrate; 1× SSC is 150 mmol L^{-1} NaCl plus 15 mmol L^{-1} trisodium citrate), 2×Denhardt's solution, 20 mmol L^{-1} Na_2HPO_4 (PH 7.2) and 7% SDS, for 16 h at 55°C. After hybridization, the samples were washed two times for 20 min in 2× SSC-0.1% SDS at 50°C. Quantification of the radioactive signal was performed with the ImageJ software from Molecular Dynamics.

Measurement of ATP generation

ATP generation was measured using an ATP assay kit (Beyotime, Shanghai, China) and a luminometer (Promega, USA) according to the manufacturer's instructions. Briefly, HEK293T cells were treated with the ATP synthase inhibitor oligomycin to determine cytosolic ATP generation and the same number of cells in the same circumstances were incubated in DMEM for total ATP levels measurement. After treatment, all HEK293T cells were homogenized and centrifuged at 12,000×*g* for 5 min at 4°C respectively. Mitochondrial ATP levels=(total cellular ATP levels) −(oligomycin-resistant ATP levels, cytosolic ATP).

Statistical analysis

Statistical analysis was carried out using the unpaired, two-tailed Student's *t*-test contained. Differences were considered significant at a $P<0.05$.

Functional assay of *VHL*

Plasmid constructs and mutants

G. maculatum gene *VHL* as well as *VHL* mutant (V102Q) were biosynthetic that were subcloned into pCMV-HA vetors (CloneTech, USA).

Cell culture and transfection

HEK293T cells were originally obtained from ATCC and maintained in Dulbecco's modified Eagle's medium (HyClone) supplemented with 10% FBS (HyClone) and 1% penicillin streptomycin solution (HyClone) at 37°C in a humidified atmosphere incubator containing 5% CO_2. HEK293T cells were transfected with the constructed plasmids using VigoFect (Vigorous Biotechnology, Beijing, China) following the manufacturer's instructions.

Hypoxia treatment

The Ruskinn INVIVO2 I-400 workstation was used for hypoxia treatment on cells. The O_2 concentration was adjusted to the appropriate value (2%) prior to experimentation. Transfected cells were incubated in fresh medium and divided into two groups and cultured in normoxic or hypoxic chambers for 14 h.

Luciferase reporter assay

HEK293T cells were seeded in 24-well plates for 12 h under normoxia (21% O_2) and transfected with the indicated plasmids together with Hypoxia Response Element luciferase reporter (HRE-Luc.) following the instructions recommended in the manufacturer's protocol. and pCMV-Renilla as an internal control. Luciferase activity was measured 24 h after transfection using the Dual-luciferase Reporter Assay System (Promega). Data were normalized to Renilla luciferase.

Data availability

Raw sequencing reads have been deposited in NCBI with the BioProject accession PRJNA820898.

Compliance and ethics *The author(s) declare that they have no conflict of interest.*

Acknowledgements *This work was supported by the Strategic Priority Research Program of Chinese Academy of Sciences (XDB31000000), the National Natural Science Foundation of China (32170480, 31972866, 31702016, 31601858, 32022009), Chinese Academy of Sciences (Youth Innovation Promotion Association, Chinese Academy of Sciences (http://www.yicas.cn), the Pioneer Hundred Talents Program, and ZDBS-LY-SM005), the Second Tibetan Plateau Scientific Expedition and Research Program (STEP, 2019QZKK0501), State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences (GREKF21-04), and the Young Top-notch Talent Cultivation Program of Hubei Province. This work was supported by the Wuhan Branch, Supercomputing Center, Chinese Academy of Sciences, China. We are grateful to Prof. Igor Schneider from Instituto de Ciencias Biologicas, Universidade Federal do Para for help with English editing. We thank Prof. Le Kang for the comments of this paper.*

References

Ahn, D., and Ho, R.K. (2008). Tri-phasic expression of posterior *Hox* genes during development of pectoral fins in zebrafish: implications for the evolution of vertebrate paired appendages. Dev Biol 322, 220–233.

Alioto, T., Picardi, E., Guigó, R., and Pesole, G. (2013). ASPic-GeneID: a lightweight pipeline for gene prediction and alternative isoforms detection. Biomed Res Int 2013, 1–11.

Benson, G. (1999). Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res 27, 573–580.

Bickler, P.E., and Buck, L.T. (2007). Hypoxia tolerance in reptiles, amphibians, and fishes: life with variable oxygen availability. Annu Rev Physiol 69, 145–170.

Birney, E., Clamp, M., and Durbin, R. (2004). GeneWise and genomewise. Genome Res 14, 988–995.

Burton, J.N., Adey, A., Patwardhan, R.P., Qiu, R., Kitzman, J.O., and Shendure, J. (2013). Chromosome-scale scaffolding of *de novo* genome assemblies based on chromatin interactions. Nat Biotechnol 31, 1119–1125.

Cao, L., Huang, Q., Wu, Z., Cao, D.D., Ma, Z., Xu, Q., Hu, P., Fu, Y., Shen, Y., Chan, J., et al. (2016). Neofunctionalization of zona pellucida proteins enhances freeze-prevention in the eggs of Antarctic notothenioids. Nat Commun 7, 12987.

Capella-Gutiérrez, S., Silla-Martínez, J.M., and Gabaldón, T. (2009). trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25, 1972–1973.

Chu, X.L., Zheng, B.S., and Dai, D.Y. (1999). Fauna Sinica (Osteichthyes: Siluriformes) (in Chinese). Beijing: Science Press.

Conesa, A., Gotz, S., Garcia-Gomez, J.M., Terol, J., Talon, M., and Robles, M. (2005). Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21, 3674–3676.

De Bie, T., Cristianini, N., Demuth, J.P., and Hahn, M.W. (2006). CAFE: a computational tool for the study of gene family evolution. Bioinformatics 22, 1269–1271.

Eddy, S.R. (1996). Hidden Markov models. Curr Opin Struct Biol 6, 361–365.

Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32, 1792–1797.

El Yacoubi, B., Bailly, M., and de Crécy-Lagard, V. (2012). Biosynthesis and function of posttranscriptional modifications of transfer RNAs. Annu Rev Genet 46, 69–95.

Emms, D.M., and Kelly, S. (2015). OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. Genome Biol 16, 157.

Ge, R.L., Cai, Q., Shen, Y.Y., San, A., Ma, L., Zhang, Y., Yi, X., Chen, Y., Yang, L., Huang, Y., et al. (2013). Draft genome sequence of the Tibetan antelope. Nat Commun 4, 1858.

Govoni, K.E., Lee, S.K., Chadwick, R.B., Yu, H., Kasukawa, Y., Baylink, D.J., and Mohan, S. (2006). Whole genome microarray analysis of growth hormone-induced gene expression in bone: T-box3, a novel transcription factor, regulates osteoblast proliferation. Am J Physiol Endocrinol Metab 291, E128–E136.

Guo, X., He, S., and Zhang, Y. (2005). Phylogeny and biogeography of Chinese sisorid catfishes re-examined using mitochondrial cytochrome b and 16S rRNA gene sequences. Mol Phylogenet Evol 35, 344–362.

Haas, B.J., Salzberg, S.L., Zhu, W., Pertea, M., Allen, J.E., Orvis, J., White, O., Buell, C.R., and Wortman, J.R. (2008). Automated eukaryotic gene structure annotation using EVidenceModeler and the Program to Assemble Spliced Alignments. Genome Biol 9, R7.

Han, L., Monné, M., Okumura, H., Schwend, T., Cherry, A.L., Flot, D.,

Matsuda, T., and Jovine, L. (2010). Insights into egg coat assembly and egg-sperm interaction from the X-ray structure of full-length ZP3. Cell 143, 404–415.
Hao, Y., Xiong, Y., Cheng, Y., Song, G., Jia, C., Qu, Y., and Lei, F. (2019). Comparative transcriptomics of 3 high-altitude passerine birds and their low-altitude relatives. Proc Natl Acad Sci USA 116, 11851–11856.
Huang, D.W., Sherman, B.T., and Lempicki, R.A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 4, 44–57.
Hui, C., and Joyner, A.L. (1993). A mouse model of Greig cephalo-polysyndactyly syndrome: the *extra-toes*^{J} mutation contains an intragenic deletion of the *Gli3* gene. Nat Genet 3, 241–246.
Jiang, W., Lv, Y., Cheng, L., Yang, K., Bian, C., Wang, X., Li, Y., Pan, X., You, X., Zhang, Y., et al. (2019). Whole-genome sequencing of the giant devil catfish, *Bagarius yarrelli*. Genome Biol Evol 11, 2071–2077.
Jones, P., Binns, D., Chang, H.Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen, J., Mitchell, A., Nuka, G., et al. (2014). InterProScan 5: genome-scale protein function classification. Bioinformatics 30, 1236–1240.
Jurka, J., Kapitonov, V.V., Pavlicek, A., Klonowski, P., Kohany, O., and Walichiewicz, J. (2005). Repbase Update, a database of eukaryotic repetitive elements. Cytogenet Genome Res 110, 462–467.
Kang, S., Xu, Y., You, Q., Flügel, W.A., Pepin, N., and Yao, T. (2010). Review of climate and cryospheric change in the Tibetan Plateau. Environ Res Lett 5, 015101.
Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S. L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol 14, R36.
Kolmogorov, M., Yuan, J., Lin, Y., and Pevzner, P.A. (2019). Assembly of long, error-prone reads using repeat graphs. Nat Biotechnol 37, 540–546.
Koren, S., Walenz, B.P., Berlin, K., Miller, J.R., Bergman, N.H., and Phillippy, A.M. (2017). Canu: scalable and accurate long-read assembly via adaptive *k*-mer weighting and repeat separation. Genome Res 27, 722–736.
Korf, I. (2004). Gene finding in novel genomes. BMC Bioinformatics 5, 59.
Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol 10, R25.
Li, H., and Durbin, R. (2011). Inference of human population history from individual whole-genome sequences. Nature 475, 493–496.
Li, J.T., Gao, Y.D., Xie, L., Deng, C., Shi, P., Guan, M.L., Huang, S., Ren, J.L., Wu, D.D., Ding, L., et al. (2018). Comparative genomic investigation of high-elevation adaptation in ectothermic snakes. Proc Natl Acad Sci USA 115, 8406–8411.
Li, M., Tian, S., Jin, L., Zhou, G., Li, Y., Zhang, Y., Wang, T., Yeung, C.K. L., Chen, L., Ma, J., et al. (2013). Genomic analyses identify distinct patterns of selection in domesticated pigs and Tibetan wild boars. Nat Genet 45, 1431–1438.
Liu, H., Liu, Q., Chen, Z., Liu, Y., Zhou, C., Liang, Q., Ma, C., Zhou, J., Pan, Y., Chen, M., et al. (2018). Draft genome of *Glyptosternon maculatum*, an endemic fish from Tibet Plateau. Gigascience 7.
Liu, X.D., and Dong, B.W. (2013). Influence of the Tibetan Plateau uplift on the Asian monsoon-arid environment evolution. Chin Sci Bull 58, 4277–4291.
Luo, R., Liu, B., Xie, Y., Li, Z., Huang, W., Yuan, J., He, G., Chen, Y., Pan, Q., Liu, Y., et al. (2012). SOAPdenovo2: an empirically improved memory-efficient short-read *de novo* assembler. Gigascience 1, 18.
Ma, X., Kang, J., Chen, W., Zhou, C., and He, S. (2015). Biogeographic history and high-elevation adaptations inferred from the mitochondrial genome of Glyptosternoid fishes (Sisoridae, Siluriformes) from the southeastern Tibetan Plateau. BMC Evol Biol 15, 233.
Majoros, W.H., Pertea, M., and Salzberg, S.L. (2004). TigrScan and GlimmerHMM: two open source *ab initio* eukaryotic gene-finders. Bioinformatics 20, 2878–2879.
Maxwell, P.H., Wiesener, M.S., Chang, G.W., Clifford, S.C., Vaux, E.C., Cockman, M.E., Wykoff, C.C., Pugh, C.W., Maher, E.R., and Ratcliffe, P.J. (1999). The tumour suppressor protein VHL targets hypoxia-inducible factors for oxygen-dependent proteolysis. Nature 399, 271–275.
Monaghan, R.M., and Whitmarsh, A.J. (2015). Mitochondrial proteins moonlighting in the nucleus. Trends Biochem Sci 40, 728–735.
Peng, Z., He, S., and Zhang, Y. (2004). Phylogenetic relationships of glyptosternoid fishes (Siluriformes: Sisoridae) inferred from mitochondrial cytochrome *b* gene sequences. Mol Phylogenet Evol 31, 979–987.
Peng, Z., Ho, S.Y.W., Zhang, Y., and He, S. (2006). Uplift of the Tibetan plateau: evidence from divergence times of glyptosternoid catfishes. Mol Phylogenet Evol 39, 568–572.
Qiu, Q., Zhang, G., Ma, T., Qian, W., Wang, J., Ye, Z., Cao, C., Hu, Q., Kim, J., Larkin, D.M., et al. (2012). The yak genome and adaptation to life at high altitude. Nat Genet 44, 946–949.
Qu, Y., Chen, C., Chen, X., Hao, Y., She, H., Wang, M., Ericson, P.G.P., Lin, H., Cai, T., Song, G., et al. (2021). The evolution of ancestral and species-specific adaptations in snowfinches at the Qinghai-Tibet Plateau. Proc Natl Acad Sci USA 118, e2012398118.
Qu, Y., Zhao, H., Han, N., Zhou, G., Song, G., Gao, B., Tian, S., Zhang, J., Zhang, R., Meng, X., et al. (2013). Ground tit genome reveals avian adaptation to living at high altitudes in the Tibetan Plateau. Nat Commun 4, 2071.
Qu, Y., Chen, C., Xiong, Y., She, H., Zhang, Y.E., Cheng, Y., DuBay, S., Li, D., Ericson, P.G.P., Hao, Y., et al. (2020). Rapid phenotypic evolution with shallow genomic differentiation during early stages of high elevation adaptation in Eurasian Tree Sparrows. Natl Sci Rev 7, 113–127.
Royden, L.H., Burchfiel, B.C., and van der Hilst, R.D. (2008). The geological evolution of the Tibetan Plateau. Science 321, 1054–1058.
Schiffels, S., and Durbin, R. (2014). Inferring human population size and separation history from multiple genome sequences. Nat Genet 46, 919–925.
Spicer, R.A., Su, T., Valdes, P.J., Farnsworth, A., Wu, F.X., Shi, G., Spicer, T.E.V., and Zhou, Z. (2021). Why 'the uplift of the Tibetan Plateau' is a myth. Natl Sci Rev 8.
Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312–1313.
Stanke, M., Diekhans, M., Baertsch, R., and Haussler, D. (2008). Using native and syntenically mapped cDNA alignments to improve *de novo* gene finding. Bioinformatics 24, 637–644.
Tarailo-Graovac, M., and Chen, N. (2009). Using RepeatMasker to identify repetitive elements in genomic sequences. Curr Protoc Bioinformatics 25.
Thisse, C., and Thisse, B. (2008). High-resolution *in situ* hybridization to whole-mount zebrafish embryos. Nat Protoc 3, 59–69.
Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L., Rinn, J.L., and Pachter, L. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc 7, 562–578.
Vaser, R., Sović, I., Nagarajan, N., and Šikić, M. (2017). Fast and accurate *de novo* genome assembly from long uncorrected reads. Genome Res 27, 737–746.
Walker, B.J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., Cuomo, C.A., Zeng, Q., Wortman, J., Young, S.K., et al. (2014). Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. PLoS ONE 9, e112963.
Wang, Y., Shen, Y., Feng, C., Zhao, K., Song, Z., Zhang, Y., Yang, L., and He, S. (2016). Mitogenomic perspectives on the origin of Tibetan loaches and their adaptation to high altitude. Sci Rep 6, 29690.
Wang, K., Shen, Y., Yang, Y., Gan, X., Liu, G., Hu, K., Li, Y., Gao, Z., Zhu, L., Yan, G., et al. (2019). Morphology and genome of a snailfish from the Mariana Trench provide insights into deep-sea adaptation. Nat Ecol Evol 3, 823–833.
Wu, Y.F., and Wu, C.Z. (1992). The Fishes of the Qinghai-Xizang Plateau

(in Chinese). Chengdu: Sichuan Science and Technology Press.
Xiao, S.J., Mou, Z.B., Yang, R.B., Fan, D.D., Liu, J.Q., Zou, Y., Zhu, S.L., Zou, M., Zhou, C.W., and Liu, H.P. (2021). Genome and population evolution and environmental adaptation of *Glyptosternon maculatum* on the Qinghai-Tibet Plateau. Zoological Res 42, 502–513.
Xie, C., Mao, X., Huang, J., Ding, Y., Wu, J., Dong, S., Kong, L., Gao, G., Li, C.Y., and Wei, L. (2011). KOBAS 2.0: a web server for annotation and identification of enriched pathways and diseases. Nucleic Acids Res 39, W316–W322.
Xiong, Y., Fan, L., Hao, Y., Cheng, Y., Chang, Y., Wang, J., Lin, H., Song, G., Qu, Y., and Lei, F. (2020). Physiological and genetic convergence supports hypoxia resistance in high-altitude songbirds. PloS Genet 16, e1009270.
Yang, L., Wang, Y., Sun, N., Chen, J., and He, S. (2021). Genomic and functional evidence reveals convergent evolution in fishes on the Tibetan Plateau. Mol Ecol 30, 5752–5764.
Yang, Z. (2007). PAML 4: phylogenetic analysis by maximum likelihood. Mol Biol Evol 24, 1586–1591.
Yu, L., Wang, G.D., Ruan, J., Chen, Y.B., Yang, C.P., Cao, X., Wu, H., Liu, Y.H., Du, Z.L., Wang, X.P., et al. (2016). Genomic analysis of snub-nosed monkeys (*Rhinopithecus*) identifies genes and processes related to high-altitude adaptation. Nat Genet 48, 947–952.
Yue, P.Q. (2000). Fauna Sinica: Osteichthyes Cypriniformes III (in Chinese). Beijing: Science Press.
Zhang, Z., Xu, D., Wang, L., Hao, J., Wang, J., Zhou, X., Wang, W., Qiu, Q., Huang, X., Zhou, J., et al. (2016). Convergent evolution of rumen microbiomes in high-altitude mammals. Curr Biol 26, 1873–1879.
Zhu, X., Guan, Y., Signore, A.V., Natarajan, C., DuBay, S.G., Cheng, Y., Han, N., Song, G., Qu, Y., Moriyama, H., et al. (2018). Divergent and parallel routes of biochemical adaptation in high-altitude passerine birds from the Qinghai-Tibet Plateau. Proc Natl Acad Sci USA 115, 1865–1870.

SUPPORTING INFORMATION

The supporting information is available online at https://doi.org/10.1007/s11427-022-2253-7. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.