Biogeographic history and high-elevation adaptations inferred from the mitochondrial genome of Glyptosternoid fishes (Sisoridae, Siluriformes) from the southeastern Tibetan Plateau

Xiuhui Ma1,2, Jingliang Kang2,3, Weitao Chen2,3, Chuanjiang Zhou2 and Shunping He2*

Abstract

Background: The distribution of the Chinese Glyptosternoid catfish is limited to the rivers of the Tibetan Plateau and peripheral regions, especially the drainage areas of southeastern Tibet. Therefore, Glyptosternoid fishes are ideal for reconstructing the geological history of the southeastern Tibetan drainage patterns and mitochondrial genetic adaptations to high elevations.

Results: Our phylogenetic results support the monophyly of the Sisoridae and the Glyptosternoid fishes. The reconstructed ancestral geographical distribution suggests that the ancestral Glyptosternoids was widely distributed throughout the Brahmaputra drainage in the eastern Himalayas and Tibetan area during the Late Miocene (c. 5.5 Ma). We found that the Glyptosternoid fishes lineage had a higher ratio of nonsynonymous to synonymous substitutions than those found in non-Glyptosternoids. In addition, $\omega_{\text{pss}}$ was estimated to be 10.73, which is significantly higher than 1 ($p$-value 0.0002), in COX1, which indicates positive selection in the common ancestral branch of Glyptosternoid fishes in China. We also found other signatures of positive selection in the branch of specialized species. These results imply mitochondrial genetic adaptation to high elevations in the Glyptosternoids.

Conclusions: We reconstructed a possible scenario for the southeastern Tibetan drainage patterns based on the adaptive geographical distribution of the Chinese Glyptosternoids in this drainage. The Glyptosternoids may have experienced accelerated evolutionary rates in mitochondrial genes that were driven by positive selection to better adapt to the high-elevation environment of the Tibetan Plateau.

Keywords: Glyptosternoids, High-elevation environment, Mitochondrial genome, PCR-based next-generation sequencing, Southeastern Tibetan drainage patterns

Background

The Tibetan Plateau (the “Roof of the World”) is the highest plateau on earth, with an average elevation of more than 4000 m. The plateau, which covers more than 2,500,000 km of plateaus and mountains in central Asia and is surrounded by towering mountain ranges, has been designated as a global hotspot of biodiversity [1].
role in ATP synthesis and heat generation, and intense selection pressures may preferentially affect mitochondria in high-elevation environments. Previous studies have detected signals of positive selection in the mitochondrial genomes of organisms living at high elevations, including goats [15], Tibetan antelope [16], Tibetan asses [17], Tibetan horses [18], pikas [19], Chinese snub-nosed monkeys [20], and bar-headed geese [21]. However, most of these studies have focused on mammals or birds. Among fish, only the high-elevation adaptations of the schizothoracine fishes (Cyprinidae) have been examined [22].

Glyptosternoids refer to catfishes in the family Sisoridae subfamily Glyptosterninae tribe Glyptosternina. Currently, there are around 10 genera and 71 species of glyptosternoids, which 9 genera and 31 species distributed in China (http://www.calacademy.org/scientists/projects/catalog-of-fishes). Chinese Glyptosternoids are found in the rivers around the Tibetan Plateau and eastern Himalayas, e.g., the Yalujiang (Brahmaputra River), Irrawaddy, Nujiang (Upper Salween), Lancangjiang (Upper Mekong River), Jinshajiang (Upper Yangtze), Yuanjiang (Red River), Nanpanjiang (Upper Pearl River) and the Brahmaputra basin [23]. The Glyptosternoids (Siluriformes) represent one of the three broad fish lineages (including the schizothoracines and Triptopterus) commonly found on the Tibetan Plateau. Habitat is thought to play a crucial role in diversification, and changes in habitat likely affect the distribution and diversification of biota in a particular region [24]. In turn, the historical biogeography of a lineage reflects aspects of the history of the region in which the species, or lineage, is distributed. The collision between India and Asia caused the uplift of the Tibetan Plateau in the Late Eocene [25, 26], which affected the fauna (e.g., fish [23, 27, 28], frogs [29] and pikas [30]), the climate [31], and the rivers in this region [32]. Chinese Glyptosternoids provide an excellent resource with which to infer the geological and environmental history of the region. Several studies have investigated the phylogeny, biogeography and evolution of the Glyptosternoids [23, 28, 33–35]. Due to the unique distribution and morphology of the fishes of this lineage, the relationships between the speciation, evolution and biogeography of these species and the Tibetan Plateau has become an area of intense research [28, 33, 34, 36–38].

Three different explanations have been suggested for the extant distribution patterns of the Chinese Glyptosternoids. (1) Hora and Silas suggested that the Glyptosternoids originated in the eastern Himalayan area of Yunnan province, southwestern China, but the exact origin and route of expansion were not clear [36]. (2) Based on the fossil records of Bagarius yarrelli, Chu inferred that the Glyptosternoids originated in southeastern Tibet during the late Pliocene [39]. According to Chu, the Glyptosternum-like species then expanded eastward to western Sichuan and northern Yunnan after the formation of the Jinshajiang River. Diversification at the genus level was proposed to have occurred during the Pleistocene, with species then expanding into the rivers of Yunnan and Sichuan during the most recent uplift of the Himalayan area. (3) Other authors have suggested that the ancestor of the Glyptosternoids was widely distributed throughout the Tibetan Plateau in the early Pleistocene [33, 40] and that the ancestor of the Glyptosternum-like species maintained this distribution during the initial uplift of the Tibetan Plateau. The ancestor of the Euchiloglanis-like fish subsequently originated in the eastern Himalayan area during the second uplift of the Tibetan Plateau. Euchiloglanis-like species were then isolated to the Jinsha, Lancang, Nujiang, Yuanjiang, Pearl and Irrawaddy Rivers by the third uplift of the Tibetan Plateau. The specialized Glyptosternoids achieved their present distribution pattern due to the isolation of the rivers.

In this study, we aimed to reconstruct the ancestral distribution of the Glyptosternoids to test hypotheses concerning speciation with respect to southeastern Tibetan drainage patterns following the uplift of the Tibetan Plateau. Molecular clock approaches were used to infer divergence dates for this molecular phylogeny; to test whether the speciation, diversification and evolution of the Chinese Glyptosternoids are associated with the uplift of the Tibetan Plateau; and to examine high-elevation adaptive mitochondrial evolution in this lineage.

**Methods**

**Muscle samples and DNA extraction**

The experiments were performed in accordance with the Ethics Committee of the Institute of Hydrobiology, Chinese Academy of Sciences. The Ethics Committee has also given ethics approval for our study. The policies were enacted according to Chinese Association for Laboratory Animal Sciences, and coordinated with the Institutional Animal Care and Use Committee (IACUC) protocols [41, 42]. The field work sample collection has also been permission according the Key Fund and NSFC-Yunnan mutual funds of the National Natural Science Foundation of China (Grant Nos. 31130049 and U1036603). Samples, including seventeen glyptosternoid species (18 individuals, more than half of Chinese glyptosternoids species), four other sisorids and three non-sisorids, following the system of Chu et al. [43] and references [44, 45], were collected from a variety of locations in China (Fig. 1 and Additional file 1: Table S1). As outgroups, Liobagrus nigricauda (Siluriformes: Amblycipitidae), Cranoglanis boudierii (Siluriformes: Amblycipitidae),
Cranoglanididae) and *Ictalurus punctatus* (Ictaluridae) included, putative close relatives to Sisoridae according to a recent study [46]. Voucher specimens were deposited at the Institute of Hydrobiology at the Chinese Academy of Sciences. Total genomic DNA was extracted from the muscle of a specimen using the OMEGA Genomic DNA Extraction Kit.

Long PCR amplification
The complete mitochondrial genomes were amplified from the genomic DNA of the Sisoridae fishes using four overlapping amplification primers by long PCR methods: L9752 AGTACRAGTGACTTCCAATCACC, H2627 GT CCTGATCCAAACATCGAGG, L295 GTAAAATTCGTGCCAGCCACC, and H10174 TCTGAGCCGAAATCGAGGTC. The PCR reactions were prepared in total volumes of 50 μL as follows: 5x LongAmp buffer, 10 μL of each 2.5 mM dNTP, 6 μL of each 0.4 μM primer, 2.0 U LongAmp polymerase, and 20–50 ng of genomic DNA. The PCR conditions included an initial denaturation step at 94 °C for 30 s followed by 30–35 cycles at 94 °C for 30 s, 61–68 °C for 1 min, and 65 °C for 10 min, with a final extension of 10 min at 65 °C. Annealing temperatures were varied within these ranges in order to optimize the efficiency of different primers and samples.

Library preparation
Library preparation was conducted using the “With-Bead” Method [47] with a slight modification: The Long PCR product was subsequently sheared to approximately 500 bp using a Covaris S2 Focused-ultrasonicator (Covaris, Inc., Woburn, Massachusetts, USA) before library preparation. The fragment sizes between 250 bp and 500 bp were selected by the gel extract method. After shearing, overhanging 5’- and 3’-ends were repaired by T4 DNA polymerase, 5’-phosphates were attached using T4 polynucleotide kinase, and P5 and P7 adapters were ligated to the ends of the repaired molecules using T4 polynucleotide kinase, and P5 and P7 adapters were ligated to the ends of the repaired molecules using T4 DNA ligase. The resulting single-strand nicks are completed using Bst polymerase to allow amplification of the insert. The library was amplified using ‘off-bead amplification’ and tagged by different indexing primers.

Quantifying and pooling library using q-PCR
We used the LightCycler 480 SYBR Green I Master (Roche, Basel, Switzerland) and qPCR Primers 1.1 and 2.1 according to the Illumina protocol (Illumina, USA).
We pooled 40 indexed samples in equimolar ratios after they were quantified.

**Sequencing and assembly**
We used 600 μL of 16 pM samples for paired-end 300 bp sequencing on a MiSeq sequencer (Illumina, Inc., San Diego, CA, USA). Sequence reads were sorted into each sample by the indices. As a first step, quality control checks were run on raw sequence data via fastqc (http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc). The adapter sequences and sites with lower qualities were trimmed using Cutadapt [48] and the fastq_quality_filter tool (http://hannonlab.cshl.edu/fastx_toolkit/). Contigs were assembled de novo for each species using Trinity [49]. We then mapped all contigs to the mitochondrial genomes of their relative species by one by one using LASTZ (available at http://www.bx.psu.edu/miller_lab/). Additional file 1: Table S2 provides the detailed characteristics of these mitochondrial genomes. The resulting consensus sequences were compared with sequences of ND2, D-Loop, and several other mitochondrial fragments generated from the same sample using independent Sanger sequencing.

**Phylogenetic analyses based on mitochondrial genomes**
The original mitochondrial genome sequences of 10 species of Glyptosternoids were determined in this study, and the published mitochondrial genome sequences of 15 teleost species from GenBank were used to conduct phylogenetic analyses. Cranoglanis boudieri, Ictalurus punctatus and Liobagrus nigricauda were selected as outgroups. The accession numbers of all the sequences used in this study are summarized in Additional file 1: Table S1. Twelve protein-coding genes encoded in the heavy strand of DNA and two rRNA genes were used for the analyses. We excluded the ND6 gene because this gene is encoded on the light strand, and its nucleotide compositions are very different from other genes. Each gene sequence was automatically aligned using the MAFFT program [50] and carefully checked by eye. All ambiguous portions were excluded. After removing the start and stop codons, 12 protein-coding genes and 2 rRNA genes were concatenated.

We inferred the phylogenetic relationships via the maximum-likelihood (ML) method [51] and MrBayes software [52]. For the ML analyses, we used the RAxML program version 7.2.6 [53] with the general time reversible model with gamma distribution and a proportion of invariant sites (GTR + G + I) as estimated by Python programs (i.e., run mraic.py) [54]. Taking into account the different tempo and mode of the nucleotide substitutions, the parameters of the nucleotide substitution model and the branch lengths of the first, second, and third codon positions and ribosomal RNAs were separately estimated.

To evaluate the confidence of the internal nodes, the rapid bootstrap method [53] was applied with 1000 replications. Using the MrBayes software, four independent chains were run for 10,000,000 generations with a burn-in length of 2500 generations and a sampling frequency of 1000 generations. Three of the four chains were heated, and the analysis was run twice.

**Reconstruction of ancestral geographical distribution**
To trace the historical biogeography during the evolution of the Glyptosternoids, the ancestral distribution of the internal nodes were reconstructed with the Dispersal-Extinction-Cladogenesis (DEC) model [55] using the RASP program v. 3.02. [56] based on the ML tree inferred from the data from 12 mitochondrial protein-coding genes (Fig. 2) and the distribution pattern (Additional file 1: Table S3) of Sisoridae.

**Divergence time estimation**
The divergence times of the Chinese Glyptosternoids were determined using the protein-coding genes with the Bayesian molecular dating program Beast [57], according to the manual for the program. Estimates were calibrated using two age constraints (C1 and C2; Fig. 3). The C1 calibration point is based on the fossil record of Bagarius yarrelli from the Pliocene (5.3-1.8 Ma ago) of the Siwalik Hills in India [37]. C2 represents an upper bound of 4 Ma, derived from the capture of the Tsangpo by the Brahmaputra River, which occurred prior to about this time [32]. These time estimates were conducted using the GTR + G + I model. Following a burn-in of the initial 25 % cycles, divergence times were sampled once every 1000 generations from 10^6 Markov Chain Monte Carlo (MCMC) iterations. Convergence of the chains to a stationary distribution was checked by visual inspection using TRACER v1.4 [58]. We repeated this analysis twice with different MCMC conditions and confirmed the stability of our estimates.

**Substitution rate estimation and selection analyses**
To estimate lineage-specific evolutionary rates for each branch of the Glyptosternoids and their closest relatives, the CODEML program in the PAML package [59] with the free-ratio model (model = 1) was run on each of the 12 protein-coding mitochondrial genes. The parameters dN, dS, dN/dS, N*dN, and S*dS values were obtained for each branch of the tree (Fig. 2), and genes were discarded if N*dN or S*dS were less than 1 or dS was greater than 1, according to a previous study [60].

In molecular evolutionary biology, the natural selection acting on protein-coding genes is often characterized by comparison of synonymous and nonsynonymous substitution rates [38]. The nonsynonymous over synonymous substitution rates, ω = dN/dS, is a widely used
indicator for measuring the direction and magnitude of selective pressure on amino acid replacements. The ω values <1, =1, and >1 represent purifying selection, neutral evolution, and positive selection, respectively [61, 62]. The ω ratios of the mitochondrial genomes were estimated with the CODEML program of PAML [59] using the concatenated sequence of the 12 protein-coding genes (excluding ND6). To detect positive selection in the limited codon sites in particular lineages across the phylogenetic tree, we applied the branch-site model [61]. To confirm that the ω ratios of the positively selected sites (ω_{p ss}) were significantly higher than 1, we performed the likelihood ratio test (LRT) with the null hypothesis that the ω_{p ss} value was 1. The candidates of the positively selected sites were predicted by the Empirical Bayes method [62].

Hydrological and geological events pertinent to this study

It has long been recognized that paleo-drainages of major continental East Asian Rivers, draining the southeastern Tibet plateau margin, differed markedly from...
their current drainage patterns [32, 63–69]. Clark et al. [32] suggested that these rivers were once tributaries to a single southward flowing system, which drained into the South China Sea. Subsequent reorganization into modern major river drainages was primarily caused by river capture and reversal events associated with the initiation of Miocene uplifts in eastern Tibet [32]. Although large-magnitude tectonic shear, prompted by the In-dian–Asian collision around the eastern Himalayan syntaxis, river capture and reversal events cannot be ruled out as an additional factor influencing these large-scale changes in drainage patterns [32, 66]. As reviewed by Ruber et al. [27], the evolution of drainage systems in Asia can be summarized in four stages [32]. (a) Upper Yangtze, Middle Yangtze, Upper Me-kong, and Upper Salween rivers drained into the South China Sea through the paleo Red River. (b) Capture/reversal of the Middle Yangtze river redirected drainage away from the Red River and into the East China Sea through the Lower Yangtze river. (c) Capture of the Upper Yangtze River into the Lower Yangtze River, and of the Upper Mekong and Upper Salween rivers into their modern drainage position. The Tsangpo River was also captured to the south through the Irrawaddy River. (d) Capture of the Tsangpo river through the Brahmaputra river into its modern drainage position.

Results

Phylogenetic reconstruction

Saturation tests [70] that included all taxa found no evidence for saturation in the twelve protein gene tests (except ND6) and the RNA gene tests. In each case, the index of substitution saturation (Iss) was significantly less than the critical value (Iss.c; see [70]).

The results of the Bayesian and ML nucleotide analyses produced by MrBayes and RAxML for the 12 protein-coding gene sequence datasets showed a marked consistency in topological congruence, differing only in the support values for certain nodes (Fig. 2). Just like the results based on the concatenation datasets from the 12 mitochondrial protein-coding and 2 rRNA genes (see Additional file 1: Figs. S1 and S2), the phylogenies indicate monophyly of the Chinese Sisoridae (including Pseudecheneis, Bagarius, Glyptothorax and Glyptosternoids) with very high support values (PP = 1.00 and BP = 95, Fig. 2). Likewise, monophyly of glyptosternoids (including Glyptosternum, Glaridoglanis, Exostoma, Euchiloglanis, Pseudexostoma, Oreoglanis, Pareuchiloglanis, and Creteuchiloglanis) are recovered with great posterior probability (PP = 1.00 and BP = 100, Fig. 2) The Glyptosternoids and non-Glyptosternoids (Bagarius, Glyptothorax) form a sister group with high support values (Fig. 2). Exostoma labiatum was placed with other Glyptosternoids to form a sister group to Glyptosternum maculatum in the phylogenetic trees. However, G. maculatum was placed with other Glyptosternoid fishes to form a sister group to E. labiatum in the phylogenetic trees (Fig. 3). Specialized Glyptosternoids diverged into three main lineages: The first lineage included Euchiloglanis kishinouyei and Pareuchiloglanis from the Upper Yangtze (P. anteanalis and P. sinensis); the second included the Creteuchiloglanis from Nujiang, Pareuchiloglanis gracilicaudata and Pseudexostoma; and the third lineage included the Oreoglanis, Pareuchiloglanis longicauda and P. macrostrem. The latter two lineages form sister groups separate from the Upper Yangtze lineage. Pareuchiloglanis was not resolved as monophyletic.

Consistent topologies were found among the MrBayes phylogenies using only protein-coding genes (Fig. 2) as well as using concatenating rRNA and the 12 protein-coding genes by gene region partitioning. We therefore used the 12 protein gene sequences in our phylogenetic reconstructions.

Divergence times among the Glyptosternoid fishes

Our newly estimated divergence times based on the mitochondrial genomes are shown in Fig. 3. Chinese Sisoridae were found to originate in the Late Miocene (c. 7.7 Ma), the Glyptosternoids later in the Late Miocene (c. 5.5 Mya), and the specialized Glyptosternoids, Pareuchiloglanis, Oreoglanis, Creteuchiloglanis and Pseudexostoma, between the Pleistocene and Holocene. These results also show that explosive speciation of the specialized Glyptosternoids occurred between the late Pliocene and the Quaternary (c. 2.8 Ma).

Ancestral reconstruction of the geographical distribution

The optimal distributions at each ancestral node are given in Fig. 4. The analysis suggests basal lineages for the Glyptosternoid in the Tsangpo drainages (node 44). All the ancestors of basal Glyptosternoids lineages (nodes 42, 43, and 44) were distributed in the Tsangpo basin and then spread into drainages of the Tibetan Plateau. The reconstruction of the ancestral distribution ranges in certain deeper nodes is expected to be less robust due to the higher number of ranges among daughter lineages, especially for node 41 (Additional file 1: Table S4).

Accelerated evolution in the lineage of Glyptosternoid fish

Averaged across all 12 protein-coding genes, the ratio of nonsynonymous to synonymous substitutions is significantly higher in most of the Glyptosternoids lineages than observed in other non-Glyptosternoid lineages (Additional file 1: Table S5), suggesting accelerated function evolution in the Glyptosternoids lineages (Fig. 5). In addition, the basal species of Glyptosternoids had lower
ratios of nonsynonymous to synonymous substitutions than the specialized species (Fig. 5a).

**Positively selected genes in the mitochondrial genome**

Using the branch-site model, positively selected signals were detected in fifteen branches. Among them, the signals of eight branches (branches 7, 8, 12, 13, 23, 24, 26, and 29) were weak, and AIC preferred the null hypothesis (\(\omega_{pss} = 1\)). Accordingly, the likelihood of positive selection in these eight branches was negligible. In contrast, AIC preferred the alternative hypothesis (\(\omega_{pss} = \text{maximum likelihood estimation}\)) for the other seven branches (branches 1, 10, 16, 21, 22, 25, and 31; Fig. 6 and Additional file 1: Table S3). Branch 1 represents the common ancestral branch of the Chinese Glyptosternoids. According to our estimates, \(\omega_{pss} = 10.73\); thus, the \(\omega_{pss}\) value was significantly higher than 1 (the \(p\)-value was 0.0002). Positively selected sites occurred in the COX1 gene. The candidates of the positively selected sites are shown in Additional file 1: Table S6. In the other six branches, which represent specialized Glyptosternoids, positively selected sites occurred in most of the mitochondrial protein-coding genes (Additional file 1: Table S6).

**Discussion**

**Origin and expansion of Chinese Glyptosternoid fishes and the southeastern Tibetan drainage patterns**

The evolutionary time scale of Glyptosternoids diversification is still a controversial issue. Our estimates are much younger than those of Peng [28] and Guo [23], both of which were based on several gene sequences (such as cytb) and limited sample sizes, especially with respect to the strict clock method used by Guo [23]. This work incorporated far more Glyptosternoid samples, especially *Pareuchiloglanis*, *Oreoglanis* and *Creteuchiloglanis*. Thus, the Glyptosternoids most likely originated in the Late Miocene and radiated during the Pliocene and Quaternary.

Based on our biogeographic results, we believe that the third explanation (see section *introduction*) agrees best with our phylogeny (Fig. 2). An ancestral Glyptosternoids species was widely distributed throughout the Brahmaputra drainage in the eastern Himalayas and Tibetan area during the Late Miocene (c. 7.7 Ma). The eastern Himalayas areas included what are now Nepal, India, Bhutan and China (Tibet and Yunnan). Then, the Glyptosternoids dispersed into the eastern Tibetan drainages and species evolved specialized adaptive traits suited to rapidly flowing water habitats, such as depressed bodies and heads, smaller gill openings and
Fig. 5 Evolutionary rates of Glyptosternoids. The species of each branch are listed in Additional file 1: Table S5: branches of the Glyptosternoids in Fig. 5.

Fig. 6 Detection of positively selected codon sites in the complete mitochondrial protein genes (excluding ND6). The MrBayes tree topology (Fig. 2) was used for this analysis. *ω_pla* represents the ω value of positively selected sites. *P*-value (LRT) represents the p-value of the likelihood ratio test of the alternative hypothesis (*ω_pla* = MLE) and the null hypothesis (*ω_pla* = 1). The branches in which positively selected codon sites were detected are indicated by bold black lines.
pinniform rugae on the rims of paired fins (such as pectoral fins). The teeth of these species also diversified, and their feeding habits changed (pointed-form teeth: fish and arthropods; shovel-form teeth: algae; *Exostoma*-like teeth: algae and arthropods). A rapid uplift of the Tibetan Plateau occurred at approximately 3.6 Ma, and the Yangtze, Nujiang, Lancang and Yuanjiang Rivers formed at this time. The specialized species, *Pareuchiloglanis, Creteuchiloglanis, Euchiloglanis, Oreoglanis* and *Pseudosexostoma*, originated during this phase. These species then spread to the downstream portions of the Nujiang, Lancangjiang, Yaluzangbujiang (Brahmaputra) and Irrawaddy Rivers to form the current day distribution pattern, yielding a great species diversity. The process of speciation among Chinese Glyptosternoids resulted from dispersal and vicariance events associated with the uplift of the Tibetan area and the newly formed river systems.

The historical biogeography of a lineage reflects aspects of the history of the region in which the species, or lineage, is distributed. We reconstructed a possible scenario for biogeographic history of the southeastern Tibet inferred from the geographical distribution of the Chinese Glyptosternoids across this drainage area. An exact search with DEC suggested basal lineages for the Glyptosternoids in the Tsangpo drainages (Fig. 4). The inferred ancestral areas (based on DEC results) are shown in Fig. 4. To distinguish between these scenarios, geological evidence and the likelihood of widespread ancestors must be considered. The use of a wide range of calibrated rates allows us to compare the molecular divergence times with the available data on geological events in this area and to test various vicariance/dispersal hypotheses. Vicariance theorists assume that common distributional patterns result from shared vicariance events. The hypothesis of a vicariant event between the formerly connected Tsangpo and Upper Irrawaddy is also supported by data from Badidae species [27].

Our results seem to agree with the geological evidence for the separation of these drainages due to tectonic uplift in eastern Tibet. The geological events appear to have played a primary role in the diversification of Chinese Glyptosternoids. In southeastern Tibet, the order of the appearance of the specialized species represents the order in which the rivers became isolated [26]. The rivers were once tributaries of a single southward-flowing system that drained into the South China Sea [31, 32]. In the early Pliocene, the Tsangpo and Brahmaputra Rivers became isolated from the ancient Red River. Until the middle Pliocene, the Jinsha River was isolated and fostered specialized diversification at the genus level; these fish then expanded into the rivers of Yunnan and Sichuan during the most recent uplift episode in the Himalayan area. The modern Red River and Pearl River became isolated in the late Pliocene. At this time, the Glyptosternoids spread into the Nujiang and Lancang Rivers. In the Pleistocene, the Lancangjiang River was isolated following the vicariance between the modern Red River and Pearl River. In the Holocene, the Nujiang River separated from the Irrawaddy River. These results agree with [32]. River capture and reversal occurred during the rapid uplift of the Tibetan Plateau [32]. The origin and expansion of Chinese Glyptosternoid fishes were affected not only by the three cycles of uplift and two large-scale peneplanation events of the Qinghai-Tibetan Plateau but also the river capture and reversal events in eastern Tibet in the Miocene due to the uplift of the Qinghai-Tibetan Plateau. Both factors have influenced the modern drainage patterns in eastern Tibet.

**Mitochondrial adaptive genetic basis for high-elevation living**

Adaptive evolution may preferentially occur at the molecular level and may be expressed as an increased ratio of nonsynonymous substitutions to synonymous substitutions [71]. Our study adds to the growing body of evidence for adaptive evolution in the mitochondrial genome of high-elevation species. Similar to previous studies of major adaptations to high-elevation habitats of different endothermic animals based on genomic data [8, 9, 11], the Glyptosternoid fish lineage exhibits accelerated evolution in the mitochondrial genome relative to other non-Glyptosternoid fish lineages. A consequence of the fact that species living in similar ecological environments can be shaped by convergent evolution to form physiological or morphological similarities [72]. In particular, the specialized Glyptosternoid fishes have higher nonsynonymous to synonymous substitutions than the basal species, suggesting the specialized species developed accelerated evolutionary rates in order to adapt to the high-elevation environment. Thus, the mitochondrial genes of Glyptosternoid fishes may have experienced adaptively accelerated evolutionary rates to better adapt to the extreme environments of the Tibetan Plateau because accelerated evolution is usually driven by positive selection.

We identified signatures of positive selection in the branch of the Chinese Glyptosternoid fish, and these signatures may indicate adaptation to physiological hypoxia and cold stress. Branch 1 (Fig. 6) represents the common ancestral branch of the Glyptosternoid fishes in China. According to our estimates, \( \omega_{pss} \) was 10.73, i.e., significantly higher than 1 (\( p \)-value 0.0002), for the gene COX1 under positive selection. These findings are similar to previous studies on native high-elevation animals that found that the COX1 gene experienced positive selection in Tibetan antelope [16] and plateau pikas [19]. In cold environments, a less efficient OXPHOS is preferred because it results in maximum heat generation.
and minimum ATP and ROS production [73]. Mitochondria produce increased quantities of NO under hypoxia. Cytochrome c oxidase has been identified as the mitochondrial enzyme that reduces NO$_2^-$ to NO, which induces expression of nuclear hypoxic genes, possibly via a pathway that involves protein nitration [74]. Therefore, the modification of the structure and/or activity of cytochrome c oxidase or complex IV of the respiratory chain may contribute to hypoxia adaptation. Furthermore, because oxygen is the ultimate electron acceptor, which results in the production of H$_2$O in a process catalyzed by cytochrome c oxidase, modifications of the cytochrome c oxidase activities are expected to facilitate coping with a reduced oxygen supply. These modifications would affect mitochondrial NO production and, consequently, hypoxic signaling. Considering the reconstructed ancestral geographical distribution areas, this positive selection likely occurred during the process of high-elevation adaptation.

The other branches’ candidates for positively selected sites were predicted by the Empirical Bayes method [62] and are shown in Additional file 1: Table S6. Branch 21, branches 16 and 22, branch 25, branch 10 and branch 31 appear to have been distributed in the Brahmaputra/Tsangpo, Upper Yangtze, Pearl, Irrawaddy and Salween Rivers, respectively. Branch 21, which includes *Pareuchiloglanis kamengensis*, is found in the Brahmaputra/Tsangpo drainages and experienced positive selection. In this branch, the proportion of positively selected sites was 0.97 %, which corresponds to 35 codon sites in different genes. These codon sites are composed of Complex I, Complex III, Complex IV and Complex V. Branches 10, 16, 22, and 25, which were distributed in Irrawaddy, Salween, Upper Yangtze, and Pearl Rivers, were also under positive selection and correspond with many codon sites. However, these codon sites do not contain the genes associated with Complex V. Polypeptides are all subunits of the oxidative phosphorylation (OXPHOS) enzyme complexes. In aerobic organisms, OXPHOS supplies most of the ATP needed for cell metabolism. During this process, electrons from NADH or FADH$_2$ are transferred to O$_2$, generating a proton, via a series of electron carriers, which pump protons through the inner mitochondrial membrane, generating a proton gradient that drives ATP synthesis via ATP synthase (Complex V). Complex V consists of two main structural domains: an intrinsic membrane domain (F0) and an extrinsic globular domain (F1), linked by a central and a peripheral stalk [75]. The mammalian mitochondrial ATP synthase comprises at least 16 subunit types [76], among which the mitochondrial-encoded ATP6 and ATP8 are essential subunits [77]. In branch 21 (*Pareuchiloglanis kamengensis*), there are one and three positively selected candidate sites corresponding to ATP 6 and ATP 8, respectively (Additional file 1: Table S6).

Taking into account the reconstructed ancestral states of the geographical distribution (Fig. 4) and positively selected codon sites (Fig. 6), we conclude that the common ancestral branch of the Glyptosternoid fishes was distributed across the Qinghai-Tibet Plateau in China and adapted to physiological hypoxia and cold stress through transforming the codon sites in COX1. These findings are similar to those suggested for the Tibetan antelope [16] and plateau pika [19] but differ from findings reported for artiodactyls, perissodactyls, snub-nosed monkeys and humans living in high-elevation environments [2, 3, 18, 20]. Different adaptive strategies have likely been developed by different lineages. Following the dispersion of the Glyptosternoid fish into the drainages surrounding southeastern Tibet, the changes to additional codon sites corresponded to other mitochondrial genes associated with Complex I, Complex III, Complex IV and Complex V.

**Conclusions**

We reconstructed a possible scenario for the southeastern Tibetan drainage patterns based on the adaptive geographical distribution of the Chinese Glyptosternoids in this drainage. In addition, the Glyptosternoid fishes lineage had a higher ratio of nonsynonymous to synonymous substitutions than those found in non-Glyptosternoids. They may have experienced accelerated evolutionary rates in mitochondrial genes that were driven by positive selection to better adapt to the high-elevation environment of the Tibetan Plateau.

**Availability of supporting data**

The data set supporting the results of this article is available in the GenBank under KP872682, KP872683, KP872690-KP872697 and provided as supplementary data.

**Additional file**

Additional file 1: Figure S1. Phylogenetic tree estimated using the MrBayes algorithm with 12 protein-coding genes and 2 rRNA genes. Numbers below nodes represent Bayesian posterior probability. Figure S2. Phylogenetic tree estimated using the maximum-likelihood method with 12 protein-coding genes and 2 rRNA genes. Numbers below nodes represent bootstrap support from the maximum-likelihood tree. Table S1. Location of samples and accession numbers for all sequences used in this study. Table S2. Characteristics of the original mitochondrial genome sequences of 10 species of Glyptosternoids fishes. Table S3. Known distribution of Chinese Siuoraeidae fishes in the study. A: Brahmaputra B: Yalujiang (Tsangpo), C: Irrawaddy, D: Nuijiang (Salween), E: Lancangjiang (Mekong), F: Jinghaijiang (Upper Yangtze), G: Ganges, H: Red River, I: Pearl River. Table S4. Probabilities in nodes for the ancestral states of the Glyptosternoids distribution areas with Dispersal-Extinction-Cladogenesis (DEC) analysis (Fig. 4). Table S5. Branches of the Glyptosternoids in figure 5. Table S6. Candidate of the positively selected genes. (DOCX 533 kb)
Competing interests
The authors declare that they have no competing interests.

Authors' contributions
XMa's interests include comparative genomics and molecular phylogenies that address questions of species evolution, speciation processes and historical biogeography. XM contributed to the sampling, molecular experiments, data analysis and writing the manuscript. JK, WC and CJ contributed to the molecular experiments and sampling. SH contributed to the research design and writing the manuscript. All authors have read and approved the final version of the manuscript.

Acknowledgements
This work was supported by grants from the Pilot projects (No. XDB13020100).

Author details
1School of Life Science, Southwest University, Beibei, Chongqing 400715, China. 2The Key Laboratory of Aquatic Biodiversity and Conservation of Chinese Academy of Sciences, Institute of Hydrobiology, Chinese Academy of Sciences, Wuhan, Hubei 430072, China. 3University of Chinese Academy of Sciences, Beijing 100081, China.

Received: 2 April 2015 Accepted: 20 October 2015
Published online: 28 October 2015

References
1. Qi L, Chao Y, Guo S, Zhao L, Li T, Wei F, et al. Convergent, parallel and Correlated Evolution of Trophic Morphologies in the Subfamily Schizothoracinae from the Qinghai-Tibetan Plateau. Plos One. 2012, 7(3):e40740.
2. Wang Z, Yonezawa T, Liu B, Ma T, Shen X, Su J, et al. Domestication relaxed selective constraints on the yak mitochondrial genome. Mol Biol Evol. 2011;28(5):1553–6.
3. Simonson TS, Yang Y, Huff CD, Yun H, Qin G, Witherspoon DJ, et al. Genetic evidence for high-altitude adaptation in Tibet. Science. 2010;329(5987):72–5.
4. Yi X, Liang Y, Huerta-Sanchez E, Jin X, Cuo ZX, Pool JE, et al. Sequencing of 50 human exomes reveals adaptation to high altitude. Science. 2010;329(5987):75–8.
5. Brall CM, Cavallini GL, Deng L, Elston RC, Gao Y, Knight J, et al. Natural selection on EPAS1 (HIF2 alpha) associated with low hemoglobin concentration in Tibetan highlanders. Proc Natl Acad Sci U S A. 2010;107(25):11459–64.
6. Bigham A, Bauchet M, Pinto D, Mao X, Akey JM, Mei R, et al. Identifying signatures of natural selection in Tibetan and Andean populations using dense genome scan data. PLoS Genetics. 2010, 6(9):e1001116.
7. Peng Y, Yang Z, Zhang H, Cui C, Qi X, Luo X, et al. Genetic variations in Tibetan Populations and high-altitude adaptation at the Himalayas. Mol Biol Evol. 2011;28(2):1075–81.
8. Qiu Q, Zhang G, Ma T, Qian W, Wang J, Ye Z, et al. The yak genome and adaptation to life at high altitude. Nat Genet. 2012;44(8):946–9.
9. Ge R-L, Cai Q, Shen Y-Y, San A, Ma L, Zhang Y, et al. Draft genome sequence of the Tibetan antelope. Nat Commun. 2013;4:11.
10. Li M, Tian S, Jin L, Zhou G, Li Y, Zhang W, et al. Genome analyses identify distinct patterns of selection in domesticated pigs and Tibetan wild boars. Nat Genet. 2013;45(2):1431–1438.
11. Qu Y, Zhao H, Han N, Zhou G, Song G, Gao B, et al. Ground tib genome reveals avian adaptation to living at high altitudes in the Tibetan plateau. Nat Commun. 2013;4:2071.
12. Li Y, Wu D-O, Boyko AR, Wang G-D, Wu S-F, Jinw DM, et al. Population variation revealed high-altitude adaptation of Tibetan Mustangs. Mol Biol Evol. 2014;31(5):1200–5.
13. Gou X, Wang Z, Li N, Qiu F, Xu Z, Yan D, et al. Whole-genome sequencing of six dog breeds from continuous altitudes reveals adaptation to high-altitude hypoxia. Genome Res. 2014;24(8):1308–15.
14. Yang L, Wang Y, Zhang Z, He S. Comprehensive transcriptome analysis reveals accelerated genic evolution in a Tibet fish, gymnodytichys pachychelis. Genome Biol Evol. 2015;7(1):251–61.
15. Hassanin A, Ropiquet A, Coulong A, Cracaud C. Evolution of the mitochondrial genome in mammals living at high altitude: new insights from a study of the tribe Caprini (Bovidae, Antilopinae). J Mol Evol. 2009;68(4):295–310.
16. Xu S-Q, Yang Y-Z, Zhou J, Jing G-E, Chen Y-T, Wang J, et al. A mitochondrial genome sequence of the Tibetan antelope (Pantholops hodgsonii). Genomics Proteomics Bioinformatics. 2005;3(1):15–17.
17. Luo Y, Chen Y, Liu F, Gao Y. Mitochondrial genome of Tibetan wild ass (Equus kiang) reveals substitutions in NDH1 which may reflect evolutionary adaptation to cold and hypoxic conditions. Asia Life Sci. 2012;21(1):1–11.
18. Xu S, Luosang J, Hua S, He J, Ciren A, Wang W, et al. High-altitude adaptation and phylogenetic analysis of Tibetan horse based on the mitochondrial genome. J Genet Genomics. 2007;34(8):720–9.
19. Luo Y, Gao W, Gao Y, Tang S, Huang Q, Tan X, et al. Mitochondrial genome analysis of Ochotona curzoniae and implication of cytochrome c oxidase in hypoxic adaptation. Mitochondrion. 2008;8(6–7):352–7.
20. Yu L, Wang X, Ting N, Zhang Y. Mitogenomic analysis of Chinese snub-nosed monkeys: Evidence of positive selection in NADH dehydrogenase genes in high-altitude adaptation. Mitochondrion. 2011;11(3):497–503.
21. Scott GR, Schulte PM, Egginton S, Scott ALM, Richards JG, Milsom WK. Molecular evolution of cytochrome c oxidase underlies high-altitude adaptation in the bar-headed goose. Mol Biol Evol. 2011;28(1):351–63.
22. Li Y, Ren Z, Shedlock AM, Wu J, Sang L, Tensing T, et al. High altitude adaptation of the schizothoracine fishes (Cyprinidae) revealed by the mitochondrial genome analyses. Gene. 2013;517(2):169–78.
23. Guo X, He S, Zhang Y. Phylogeny and biogeography of Chinese sisorid catfishes re-examined using mitochondrial cytochrome b and 16S rRNA gene sequences. Mol Phylogenet Evol. 2005;35(2):344–62.
24. Hou Z, Sket B, Frizer C, Li S. Eocene habitat shift from saline to freshwater promoted Tethyan amphipod diversification. Proc Natl Acad Sci U S A. 2011;108(35):14533–8.
25. Spencer RA, Harris NBW, Widdowson M, Herman AB, Guo SX, Valdes PJ, et al. Constant elevation of southern Tibet over the past 15 million years. Nature. 2003;421(6923):622–4.
26. Royden LH, Burchfield BC, van der Hilst RD. The geological evolution of the Tibetan plateau. Science. 2008;321(5892):1544–8.
27. Ruber L, Britz R, Kullander SO, Zardoya R. Evolutionary and biogeographic patterns of the Badidae (Teleostei : Perciformes) inferred from mitochondrial and nuclear DNA sequence data. Mol Phylogenet Evol. 2004;32(3):1010–22.
28. Peng Z, Ho SY, Zhang Y, He S. Uplift of the Tibetan plateau: evidence from divergence times of glyptosternoid catfishes. Mol Phylogenet Evol. 2006;39(2):568–72.
29. Che J, Zhou WW, Hu JS, Yan F, Papenfuss TJ, Wake DB, et al. Spiny frogs (Paini) illuminate the history of the Himalayan region and Southeast Asia. Proc Natl Acad Sci U S A. 2010;107(31):13765–70.
30. Yu N, Zheng CL, Zhang YP, Li WH. Molecular systematics of pikas (genus Ochotona) inferred from mitochondrial DNA sequences. Mol Phylogenet Evol. 2000;16(1):85–95.
31. An ZS, Kutzbach JE, Prell WL, Porter SC. Evolution of Asian monsoons and phased uplift of the Himalayan Tibetan plateau since Late Miocene times. Nature. 2001;411(6833):62–6.
32. Clark MK, Schoenbohm LM, Royden LH, Whipple KV, Burchfield BC, Zhang X, et al. Surface uplift, tectonics, and erosion of eastern Tibet from large-scale drainage patterns. Tectonics. 2004, 23(1), TC1006.
33. He S. The analysis of historical biogeography for the glyptosternoid fishes (Teleostei: Siluriform, Sisoridae). Biogeographica (Paris). 1995;71(4):145–60.
34. He S. The phylogeny of the glyptosternoid fishes (Teleostei: Siluriformes, Sisoridae). Cybium. 1996;20(2):115–59.
35. Jiang W, Ng HH, Yang J, Chen X. Monophyly and phylogenetic relationships of the catfish genus Glyptothorax (Teleostei: Sisoridae) inferred from nuclear and mitochondrial gene sequences. Mol Phylogenet Evol. 2011;62(2):278–89.
36. Hora SL, Silas EG. Evolution and distribution of Glyptosternoid fishes of the family Sisoridae (Cyprinidae). Palaeontologia Indica. 1886, (10)(iii):241-258.
37. Lydekker R. Indian tertiary and post-tertiary vertebrata. Tertiary fishes. Palaeontologia Indica. 1886, (10)(iii):241-258.
38. Nei M. Selectionism and neutralism in molecular evolution. Mol Biol Evol. 1987;4(4):485–92.
39. Chu XL. Systematics and evolutionary pedigree of the glyptosternoid fishes (Family Sisoridae). Acta Zootaxa Sinica. 1979;4:72.
41. Chinese association for laboratory animal sciences [http://www.calas.org.cn/]
   html/ypx/zcfg/]. Accessed 20 Mar 2015.
42. Institutional animal care and use committee [http://iacuc.usc.edu/].
   Accessed 20 Mar 2015.
43. Chu XL, Zheng BS, Dai DY. Fauna Sinica, Class Teleostei, Siluriformes. Beijing:
   Science Press; 1999.
44. Zhou W, Li X, Thomson AW. A new genus of glyptosternine catfish
   (Siluriformes: Sisoridae) with descriptions of two new species from Yunnan,
   China. Copeia. 2011;2:422–41.
45. Chen X-Y. Checklist of fishes of Yunnan. Zoological Res. 2013;34(4):281–343.
46. Sullivan JP, Lundberg JG, Hardman M. A phylogenetic analysis of the major
   groups of catfishes (Teleostei: Siluriformes) using rag1 and rag2 nuclear
   gene sequences. Mol Phylogenet Evol. 2006;41(3):636–62.
47. Fisher S, Barry A, Abreu J, Minie B, Nolan J, Delory TM, et al. A scalable, fully
   automated process for construction of sequence-ready human exome
   targeted capture libraries. Genome Biol. 2011;12(1):R1.
48. Martin M. Cutadapt removes adapter sequences from high-throughput
   sequencing reads. EMBO J. 2011;7(1):10–12.
49. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, et al.
   Full-length transcriptome assembly from RNA-Seq data without a reference
   genome. Nat Biotechnol. 2011;29(7):644–643.
50. Katoh K, Misawa K, Kuma K, Miyata T. MAFFT: a novel method for rapid
   multiple sequence alignment based on fast Fourier transform. Nucleic Acids
   Res. 2002;30(14):3059–66.
51. Felsenstein J. Evolutionary trees from dna-sequences - a maximum-
   likelihood approach. J Mol Evol. 1981;17(6):368–76.
52. Huelsenbeck JP, Ronquist F. MRBAYES: Bayesian inference of phylogenetic
   trees. Bioinformatics. 2001;17(8):754–5.
53. Stamatakis A, Hoover P, Rougemont J. A rapid bootstrap algorithm for the
   RAxML web servers. Syst Biol. 2008;57(3):758–71.
54. Faircloth BC, McCormack JE, Crawford NG, Harvey MG, Brumfield RT, Glenn
   TC. Ultraconserved elements anchor thousands of genetic markers spanning
   multiple evolutionary timescales. Syst Biol. 2012;61(5):717–726.
55. Ree RH, Smith SA. Maximum likelihood inference of geographic range
   evolution by dispersal, local extinction, and cladogenesis. Syst Biol.
   2008;57(1):14–14.
56. Yu Y, Harris AJ, Blair C, He X. RASP (Reconstruct Ancestral State in
   Phylogenies): a tool for historical biogeography. Mol Phylogenet Evol.
   2015;87:46–69.
57. Drummond AJ, Suchard MA, Xie D, Rambaut A. Bayesian Phylogenetics with
   BEAUti and the BEAST 1.7. Mol Biol Evol. 2012;29(8):1969–73.
58. Drummond AJ, Rambaut A. BEAST: Bayesian evolutionary analysis by
   sampling trees. BMC Evol Biol. 2007;7.
59. Yang Z, PAML 4: Phylogenetic analysis by maximum likelihood. Mol Biol
   Evol. 2007;24(8):1586–91.
60. Goodman M, Sterner KN, Islam M, Uddin M, Sherwood CC, Hof PR, et al.
   Phylogenomic analyses reveal convergent patterns of adaptive evolution in
   elephant and human ancestries. Proc Natl Acad Sci U S A.
   2009;106(49):20824–9.
61. Yang ZH, Nielsen R. Codon-substitution models for detecting molecular
   adaptation at individual sites along specific lineages. Mol Biol Evol.
   2002;19(6):908–17.
62. Yang ZH, Wong WSW, Nielsen R. Bayes empirical Bayes inference of amino
   acid sites under positive selection. Mol Biol Evol. 2005;22(4):1107–18.
63. Brookfield M. The evolution of the great river systems of southern Asia
   during the Cenozoic India-Asia collision: rivers draining southwards.
   Geomorphology. 1998;22(3):285–312.
64. Gregory J. The evolution of the river system of southeastern Asia. Scott
   Geogr Mag. 1925;41:129–41.
65. Gregory J, Gregory C. The Alps of Chinese Tibet and their geographical
   relations. Geogr J. 1923;61(3):153–174.
66. Hallet B, Molnar P. Distorted drainage basins as markers of crustal strain east of
   the Himalaya. J Geophys Res: Solid Earth (1978–2012).
   2001;106(7):13697–709.
67. Meteyer F, Gaudemer Y, Tapponnier P, Klein M. Mass accumulation rates in
   Asia during the Cenozoic. Geophys J Int. 1999;137(2):280–318.
68. Seeger L, Gomitz V. River profiles along the Himalayan arc as indicators of
   active tectonics. Tectonophysics. 1983;92(4):335–67.
69. Zeitler PK, Meltzer AS, Koons PO, Craw D, Hallet B, Chamberlain CP, et al.
   Erosion, Himalayan geodynamics, and the geomorphology of metamorphism.
   GSA Today. 2001;11(1):4–9.