

RESEARCH ARTICLE

Open Access



# Phylogeographic structure, cryptic speciation and demographic history of the sharpbelly (*Hemiculter leucisculus*), a freshwater habitat generalist from southern China

Weitao Chen<sup>1,2</sup>, Zaixuan Zhong<sup>1,2</sup>, Wei Dai<sup>1,2</sup>, Qi Fan<sup>1</sup> and Shunping He<sup>1\*</sup> 

## Abstract

**Background:** Species with broad distributions frequently divide into multiple genetic forms and may therefore be viewed as “cryptic species”. Here, we used the mitochondrial cytochrome *b* (Cytb) and 12 nuclear DNA loci to investigate phylogeographic structures of the sharpbelly (*Hemiculter leucisculus*) in rivers in southern China and explored how the geological and climatic factors have shaped the genetic diversity and evolutionary history of this species.

**Results:** Our mitochondrial phylogenetic analysis identified three major lineages (lineages A, B, and C). Lineages B and C showed a relatively narrower geographic distribution, whereas lineage A was widely distributed in numerous drainages. Divergence dates suggested that *H. leucisculus* populations diverged between 1.61–2.38 Ma. Bayesian species delimitation analysis using 12 nuclear DNA loci indicated the three lineages probably represented three valid taxa. Isolation-with-migration (IM) analysis found substantial gene flow has occurred among the three lineages. Demographic analyses showed that lineages B and C have experienced rapid demographic expansion at 0.03 Ma and 0.08 Ma, respectively.

**Conclusions:** *Hemiculter leucisculus* populations in drainages in southern China comprise three mtDNA lineages, and each of which may represent a separate species. Intense uplift of the Qinghai–Tibetan Plateau, evolution of Asian monsoons, changes in paleo-drainages, and poor dispersal ability may have driven the divergence of the three putative species. However, gene flow occurs among the three lineages. Climatic fluctuations have a prominent impact on the populations from the lineages B and C, but exerted little influence on the lineage A.

**Keywords:** Phylogeographic structure, Cryptic speciation, *Hemiculter leucisculus*, Southern China, Multilocus

## Background

In recent years, numerous phylogeographical studies have shown that species with wide distribution ranges are frequently subdivided into genetically distinct groups [1–4]. The members of each group often cannot be discriminated by morphology and could represent hidden diversity. The detection and delimitation of such cases contribute to the understanding of speciation and its underlying processes; such studies increase our

knowledge of the biodiversity of numerous taxa and are important guides for the conservation and management of biodiversity [5]. In some cases (e.g., *Haloptilus longicornis* [6] and *Microtus agrestis* [1]), these main genetic groups can be classified as “cryptic species” and the process generating such groups can be called “cryptic speciation”.

Recent improvements in molecular and analytical tools have aided in investigations of cryptic speciation. Multilocus sequence data involving newly developed theoretical models can be particularly helpful to generate species trees of closely related taxa [7]. Several methods have been developed to address the incongruence

\* Correspondence: clad@ihb.ac.cn

<sup>1</sup>The Key Laboratory of Aquatic Biodiversity and Conservation of Chinese Academy of Sciences, Institute of Hydrobiology, Chinese Academy of Sciences, Wuhan, Hubei 430072, China

Full list of author information is available at the end of the article



between gene trees and species trees due to incomplete lineage sorting [7–9].

Geologic changes and climatic instabilities are important factors that influence the evolutionary history of many taxa. Geologic changes have frequently shape the genetic structure of many species in East Asia [10–14], including fish species [13, 15]. For example, the uplift of the Qinghai–Tibetan Plateau [16, 17] has been reported to drive speciation and population diversification in a number of East Asian fish species [3, 18, 19]. Substantial changes in paleodrainage patterns in southwestern China due to the orogenesis of the Qinghai–Tibetan Plateau [20] also exerted a more direct influence on genetic divergence among fish populations [21, 22].

Climatic instabilities during the Quaternary Period have resulted in periodic expansions and contractions of the population sizes and distribution ranges of species [23, 24]. For example, in regions of Europe and North America, cyclical glacial fluctuations shaped the distributions of species and their genetic attributes [24, 25]. However, major glaciations did not occur in East Asia [26, 27], and a relatively mild Pleistocene climate prevailed in southern China [28, 29]. Therefore, climatic cycling may have exerted little effect on population demographics in southern China [30, 31]. Nevertheless, exceptions can be seen in some plants and animals [32–35].

Cyprinids represent one of the most diverse freshwater fish groups and are a major component of the primary freshwater ichthyofauna of Eurasia [36]. The wide distribution of cyprinids makes these taxa of particular interest for studying evolutionary history and testing biogeographic hypotheses. Cyprinids achieve their highest diversity in Asiatic waters and represent more than 50% of freshwater fish diversity in the major Asian rivers, including the Yangtze River [37] and Mekong River [38].

Species with broad ranges are frequently found to show diverse population structure and distinct evolutionary histories [3, 10, 11, 31, 39, 40]. This is the case for some cyprinids in East Asia rivers, e.g., *Rhynchocypris oxycephalus* [3], *Opsariichthys bidens* [41] and *Zacco platypus* [42]. The sharpbelly, *Hemiculter leucisculus* (Cyprinidae: Cultrinae), a small cyprinid fish, reaches a size up to 23.0 cm long, and is native to fresh and brackish water habitats with a pH of 7.0, a hardness of 15 dH, and a temperature of 18 to 22 °C [43]. This species has a wide distribution in the drainage basins of China, North Korea, South Korea, Japan, Mongolia and Russia (Additional file 2: Figure S1). In China, *H. leucisculus* widely occupies various freshwater habitats, such as rivers, lakes, reservoirs, and even pools [43, 44]. Habitat generalist and wide distribution ranges of the *H. leucisculus* populations make this species an ideal candidate for the study of phylogeographic structure and for the test of biogeographic hypotheses. In addition, with the

increasing introduction of commercial fishes, some *H. leucisculus* populations have been imprudently introduced into many nonnative distributions [44].

Understanding the evolutionary history and speciation process of the sharpbelly provides an important case of the influence of biogeographic process on East Asia fish species and is important to manage this species in the future. In the current study, therefore, we use both mitochondrial DNA (mtDNA) and multiple nuclear DNA (nDNA) markers to investigate the phylogeographical structure and evolutionary history of the *H. leucisculus* populations in the rivers in southern China. In addition, we attempt to discuss the implications of our phylogeographic results in identifying cryptic diversity in *H. leucisculus*. Lastly, we evaluate the possible effects of geologic changes and climatic instabilities on the evolutionary history of *H. leucisculus* populations.

## Methods

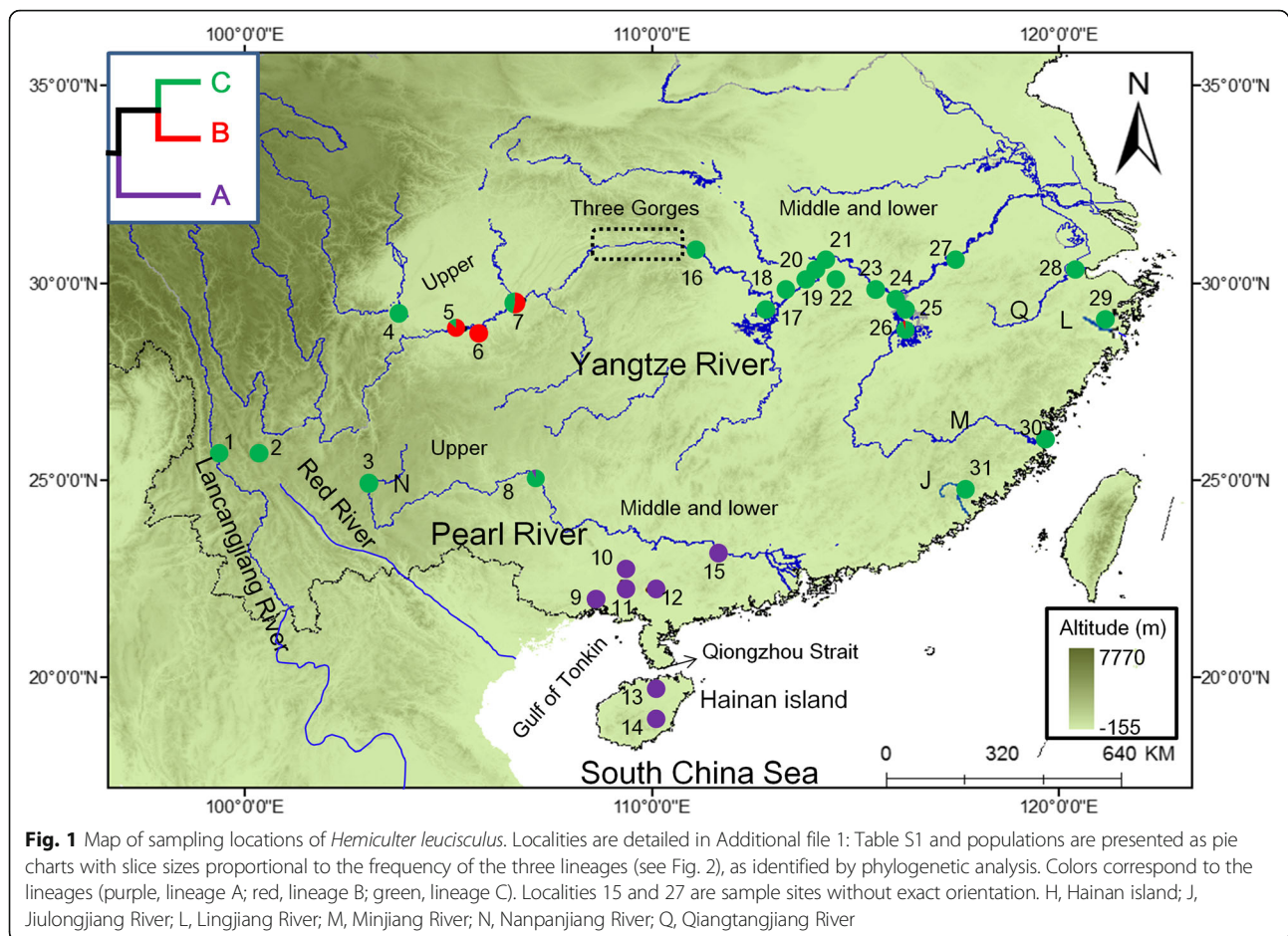
### Ethics statement

All experimental protocols were approved by the Ethics Committee of the Institute of Hydrobiology, Chinese Academy of Sciences. The policies were enacted according to Chinese Association for Laboratory Animal Sciences, and coordinated with the Institutional Animal Care and Use Committee (IACUC) protocols (<http://iacuc.usc.edu/>).

### Sampling

A total of 381 specimens of *H. leucisculus* from 22 localities were collected from the Lancangjiang River, Pearl River and Yangtze River basins (Additional file 1: Table S1, Fig. 1). A small piece of white muscle tissue or fin was dissected from the right side of each specimen. All tissue for genomic DNA extraction was preserved in 95% ethanol. Voucher specimens were deposited in the collection of the Freshwater Fish Museum of the Institute of Hydrobiology, Chinese Academy of Sciences.

The total genomic DNA was extracted from muscle tissue or fin tissue using standard salt extraction [45]. The partial mitochondrial cytochrome *b* gene (Cytb) was amplified for all individuals using the universal primers L14724 and H15915 [22]. We sequenced 98 individuals de novo and retrieved 283 sequences from a previous study [46]. Nine published sequences were included in our analyses. Among the published sequences, six sequences were sampled from six small and coastal river basins, and the remaining sequences were sampled from the Pearl River and Yangtze River basins (Additional file 1: Table S1). A total of 390 Cytb sequences from 31 localities covering most of drainages in southern China were analysed (Additional file 1: Table S1). We sequenced up to 12 independent nuclear loci from a subset of individuals which covered the range of Cytb phylogenetic tree. These loci correspond to the protein-coding genes early growth



response 3 (EGR3), ectodermal-neural cortex 1 (ENC1), glycosyltransferase (Glyt), myosin heavy chain 6 (myh6), pleiomorphic adenoma gene-like 2 (plagl2), peptide transport (Ptr), recombination activating gene 2 (RAG2), Rhodopsin, ryanodine receptor 3 (RYR3), SH3 and PX domain containing 3 (SH3PX3), sterol regulatory element-binding protein 2 (sreb2) and zic family member 1 (zic1) (Additional file 2: Table S2). Primer information is listed in Additional file 2: Table S2. The sample size of 12 nuclear genes ranged from 26 to 55 individuals, which derived from 20 populations (Additional file 1: Table S1; Additional file 2: Table S3). These markers were amplified and sequenced following a previous protocol [47–49]. The amplification of genomic DNA was conducted in a 30- $\mu$ L reaction with an initial denaturation period of 3–5 min at 94 °C followed by 30–35 cycles of 94 °C for 0.5–1 min, primer-specific annealing temperature of 53–64 °C (Additional file 2: Table S2) for 0.5–1 min, 72 °C for 1–1.5 min and a single final extension at 72 °C for 5–10 min. The amplified fragments were purified using 1.0% low-melting agarose gel electrophoresis and sequenced with the identical primer pair using an ABI PRISM 3700 (Applied Biosystems) automatic DNA sequencer.

### Sequence analyses

The nucleotide sequences were initially edited using the DNASTAR multiple package (DNASTAR Inc., Madison, WI, USA), aligned using MUSCLE [50] and optimized by eye using MEGA version 6.0 [51]. For the nuclear genes, the polymorphic positions for individual sequences from nuclear loci were carefully inspected to ensure correct and consistent identification of double peaks in heterozygotes. Nuclear gene sequences containing more than one ambiguous site were resolved using PHASE 2.1.1 [52, 53], for which input files were prepared using SEQPHASE [54]. Three runs were performed for each locus with default settings. Recombination tests for detecting the longest non-recombining region for each locus were conducted using IMGIC [55].

### Phylogenetic analyses

Phylogenetic analyses of *H. leucisculus* of the Cytb data were reconstructed using Bayesian inference (BI) and maximum-likelihood (ML) methods. *Hemiculter bleekeri* (GenBank no: KF029693) and *Culter alburnus* (GenBank no: KM044500) downloaded from GenBank, were used as the outgroup. Nucleotide substitution models



(GTR + I + G) were selected using the Akaike information criterion in MRMODELTEST version 2.3 [56]. For BI analyses, three independent runs and four Markov chains (three heated chains and a single cold chain) using the best-fit models were performed in MRBAYES v 3.1.2 [57], starting from a random tree. Each run was conducted for 20 million generations and sampled every 1000 generations, with the first 25% discarded as burn-in. We checked for stationarity by ensuring that the potential scale reduction factor equaled 1 and that the average standard deviation of split frequencies between independent runs approached 0. We examined the MCMC samples in TRACER 1.5 [58] to ensure that all chains were sampled from the same target distribution. ML analyses were implemented in RAXML-VI-HPC [59] using a GTR + I + G model. Nodal support values were estimated from 1000 nonparametric bootstrap replicates. Nodes with 95% Bayesian posterior probabilities in the BI analysis [60] and nodes with 75% bootstrap support in the ML analysis were considered strongly supported. We also used NETWORK 4.6 [61] to construct a median-joining network for Cytb.

In addition, BI and ML analyses were employed for the nDNA loci using the identical parameter settings as well. The optimal nucleotide substitution model of each nDNA locus (Additional file 2: Table S4) was chosen using the Akaike information criterion in MRMODELTEST.

#### Bayesian species delimitation

To examine whether the three lineages inferred from Cytb trees represent different species, we used Bayesian Phylogenetics and Phylogeography (BPP) v3.0 [7, 62] with 12 nuclear genes. This method allowed for the absence of monophyly for DNA haplotypes and incomplete lineage sorting [7, 63]. The model of species delimitation rjMCMC algorithm0 and finetune(e) was used in this analyses. Because the prior distributions of the ancestral population size ( $\theta$ ) and root age ( $\tau_0$ ) could have affected the posterior probabilities of the models [63], three different combinations of  $\theta$  and  $\tau_0$  were implemented in BPP: ( $\theta = (1, 10)$   $\tau = (1, 10)$ ); ( $\theta = (2, 2000)$   $\tau = (2, 2000)$ ); and ( $\theta = (1, 10)$   $\tau = (2, 2000)$ ). We ran the rjMCMC analyses for 1,000,000 generations, sampled every five generations, and discarded samples from the first 50,000 generations as burn-in. Each analysis was repeated at least twice to confirm consistency between runs. Topology based on Cytb was used as a guide tree.

#### Divergence time estimation

Divergence times among the major lineages of *H. leucisculus* were incorporated in the program BEAST 1.6.1 [64]. For molecular time estimation, the molecular clock was assessed by relative-rate tests using PHYLST [65]. Lacking of fossil evidence, divergence times were

estimated according to the strict-clock model. Based on the mitochondrial Cytb data, a range of substitution rate of 1.00–2.00% per million years (Myr) is often adopted, as these rates have been widely employed for mitochondrial Cytb gene analysis in cyprinid fish [66–68]. Thus, we employed the average value (1.50% per Myr) of the above rate range for Cytb in our study. Two independent runs were run, and the output .log and .tre files were combined using LogCombiner (BEAST package). A Yule process was implemented for the tree prior model, and the GTR + I + G model of nucleotide substitution was implemented. Each run was conducted for 50 million generations and sampled every 1000 iterations. The trees were summarized with TreeAnnotator v.1.6.1 (BEAST package) into the maximum clade credibility tree and the first 25% of sampled trees were treated as burn-in. Burn-in and convergence of the chains were determined with TRACER 1.5. ESS were used to determine the Bayesian statistical significance of each parameter.

#### Gene flow

Potential gene flow among major lineages was estimated using the isolation-with-migration (IM) model in the IMA2 program [69]. We used Cytb and the longest non-recombining regions of the 12 nuclear loci for IM analyses. The Cytb tree was used as the guide tree. The method estimates the density functions and posterior-probability densities of the IM model parameters using a Markov chain (MCMC) method [70]. The functions of the model parameters were first estimated in M-mode with two million generations, and the first 10% were discarded as burn-in. The MCMC run was repeated three times to confirm convergence. Using these functions, the marginal posterior distribution and the maximum-likelihood estimates of the demographic parameters were estimated in the L-mode. The HKY model of the DNA substitution was employed for all markers, and 40 heated metropolis-coupled Markov chains were employed to ensure convergence. ESS were used to determine whether the convergence was satisfactory.

#### Genetic diversity and population structure based on Cytb sequences

The number of haplotypes ( $n$ ), private haplotypes ( $ph$ ), haplotype diversity ( $h$ ) and nucleotide diversity ( $\pi$ ) of each population and lineage were calculated using DnaSP v5 [71]. In addition, pairwise divergences between and within the lineages were estimated by Kimura's (1980) two-parameter [72] as implemented in MEGA version 6.0.

To investigate population structure, an analysis of molecular variance (AMOVA) [73] was performed in ARLEQUIN 3.5 [74] using Jukes and Cantor distances [75]. Populations were grouped into three groups (Lineages A, B, and C) suggested by the mtDNA

phylogenetic analysis and grouped into nine groups based on drainage system. Pairwise genetic differentiation ( $\phi_{ST}$ ) was calculated using Jukes and Cantor distances between populations having more than five samples (17 populations fulfilled this condition) in ARLEQUIN 3.5. A Mantel test was used to investigate isolation by distance using genetic differentiation ( $\phi_{ST}/(1-\phi_{ST})$ ) and the geographical distance (in kilometers) measured as the straight-line distance, in ARLEQUIN 3.5. Geographic distances were estimated using GOOGLE EARTH v.4.3. All calculations implemented in ARLEQUIN 3.5 used 1000 permutations to assess significance.

### Demographic estimation

Sample size allowing observed lineages was screened for historical demographic events in each deme. Selective neutrality test of Fu's  $F_s$  statistics [76] based on Cytb was conducted for each lineage using ARLEQUIN 3.5 to find evidence of recent expansion. Significantly negative values are expected in populations that have undergone recent population expansion [76, 77]. We assessed significance by generating null distributions from 1000 permutations. To understand how population sizes changed through time, we examined the historical demographics of all lineages using coalescent-based extended Bayesian skyline plots (EBSPs). EBSPs were constructed using Cytb and three randomly selected nuDNA loci (EGR, Ptr and RAG2) in BEAST 1.6.1 to assess the variation in effective population sizes through time. The best model of nucleotide substitution was selected for each lineage and each loci (Additional file 2: Table S5), and time was scaled using a substitution rate of 1.50% substitutions/million years for Cytb data. The evolutionary rates for three nuDNA loci were estimated as a function of the Cytb evolutionary rate. A strict clock model was set as prior, 100 million generations were run for the three lineages. Each EBSP was sampled every 1000 iterations, with 10% of the initial samples discarded as burn-in. Stationarity was assessed by analyzing the effective sizes of all parameters in TRACER 1.5.

## Results

### Sequence information

A 1040-bp segment of Cytb was obtained from 390 individuals of *H. leucisculus* including 98 novel sequences, 283 sequences from the study by Fan & He (2014) [46] and 9 sequences from NCBI. The 390 sequences contained 213 variable sites and 150 parsimony-informative sites. A total of 171 unique haplotypes were defined from ingroup sequences. In addition, our nuclear dataset consisted of 12 independent loci with lengths ranging from 676 bp (Ptr) to 1227 bp (RAG2) (Additional file 2: Table S3). The number of individuals, the length of each

nDNA locus, the variable sites and the informative parsimony sites are showed in Additional file 2: Table S3. We translated all sequences into amino acids, and no stop codon was detected. All novel sequences analyzed in this study have been deposited in GenBank under Accession nos KY292565–KY293216 (Additional file 1: Table S1).

### Phylogenetic analyses of Cytb and nDNA genes

The ML and BI trees based on Cytb produced similar gene tree topologies and congruently revealed three well-supported monophyletic lineages (only BI tree shown, Fig. 2a). Lineage A, the sister group of the other two lineages, consisted of specimens from the Pearl River basin and Hainan Province (Fig. 1). Lineage B, the sister group of lineage C, mainly occurred in a small area (Localities 5–7) in the upper branch of the Yangtze River (Fig. 1). Lineage C included samples from most rivers in southern China, e.g., Yangtze River, the Nanpanjiang River, the Lancangjiang River (upper Mekong), and several eastern rivers (Fig. 1). Twelve individuals sampled from Locality 8, which is within the Pearl River basin, were also grouped into this lineage. In addition, lineages A and B displayed a high degree of geographic structure, whereas lineage C revealed a poor relationship with geographic distribution (Fig. 1). The network analyses for Cytb (Fig. 2b) showed similar patterns to the gene tree (Fig. 2a).

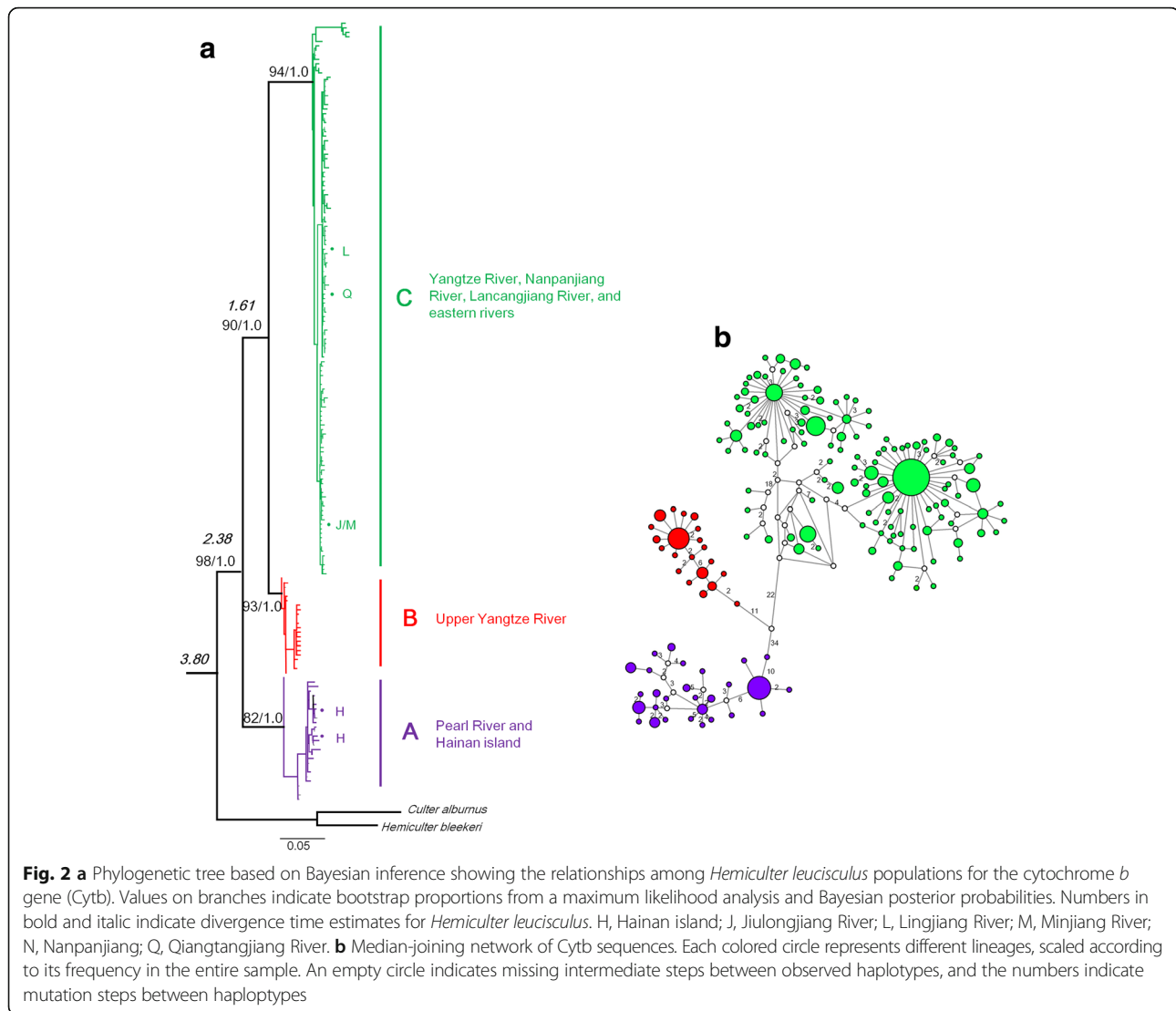
The nDNA phylogenetic trees showed that members of the three mtDNA lineages can be distinguished with three nDNA genes (i.e. ENC1, RAG2 and zic1), although the support is not very strong in some nodes (Fig. 3). The three lineages did not group into reciprocally independent clusters in the phylogenetic trees based on nine other nDNA genes (Fig. 3). Furthermore, the support value or Bayesian posterior probability of the nine nDNA trees was low in most nodes (Fig. 3).

### Bayesian species delimitation

The nuclear data strongly supported that the three lineages defined by the Cytb data represent three different species with perfect statistical support (PP = 1.00; Fig. 4). Different sets of prior distributions for  $\theta$  and  $\tau_0$  did showed consistent results.

### Divergence dating

The null hypothesis of clock-like evolution was not rejected at the 5% level. Therefore, we estimated the divergence times for the monophyletic Cytb lineages identified in the phylogenetic analyses (lineages A, B, and C as labeled in Fig. 2) using a strict molecular clock. The dating analyses (Fig. 2a) suggested that the most recent common ancestor of *H. leucisculus* diverged at 3.80 Ma (95% highest posterior density (HPD) = 2.92–4.65 Ma). Lineage A diverged at 2.38 Ma (95% HPD, 1.82–2.93 Ma), and lineages B and C split at 1.61 Ma (95% HPD, 1.19–1.98 Ma).



**Fig. 2 a** Phylogenetic tree based on Bayesian inference showing the relationships among *Hemiculter leucisculus* populations for the cytochrome *b* gene (Cytb). Values on branches indicate bootstrap proportions from a maximum likelihood analysis and Bayesian posterior probabilities. Numbers in bold and italic indicate divergence time estimates for *Hemiculter leucisculus*. H, Hainan island; J, Jiulongjiang River; L, Lingjiang River; M, Minjiang River; N, Nanpanjiang; Q, Qiangtangjiang River. **b** Median-joining network of Cytb sequences. Each colored circle represents different lineages, scaled according to its frequency in the entire sample. An empty circle indicates missing intermediate steps between observed haplotypes, and the numbers indicate mutation steps between haplotypes

### Gene flow among the three lineages

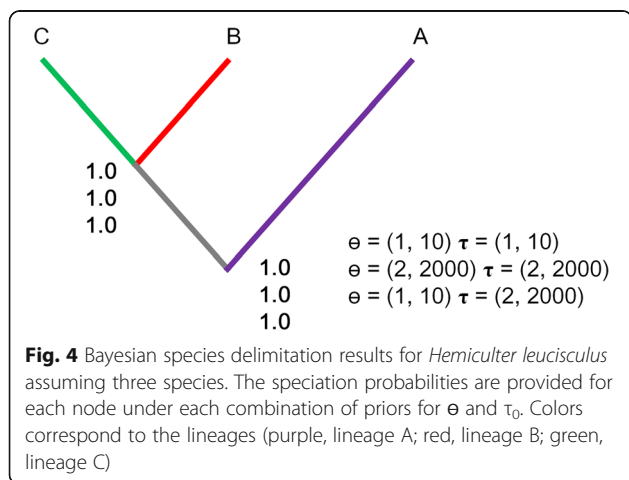
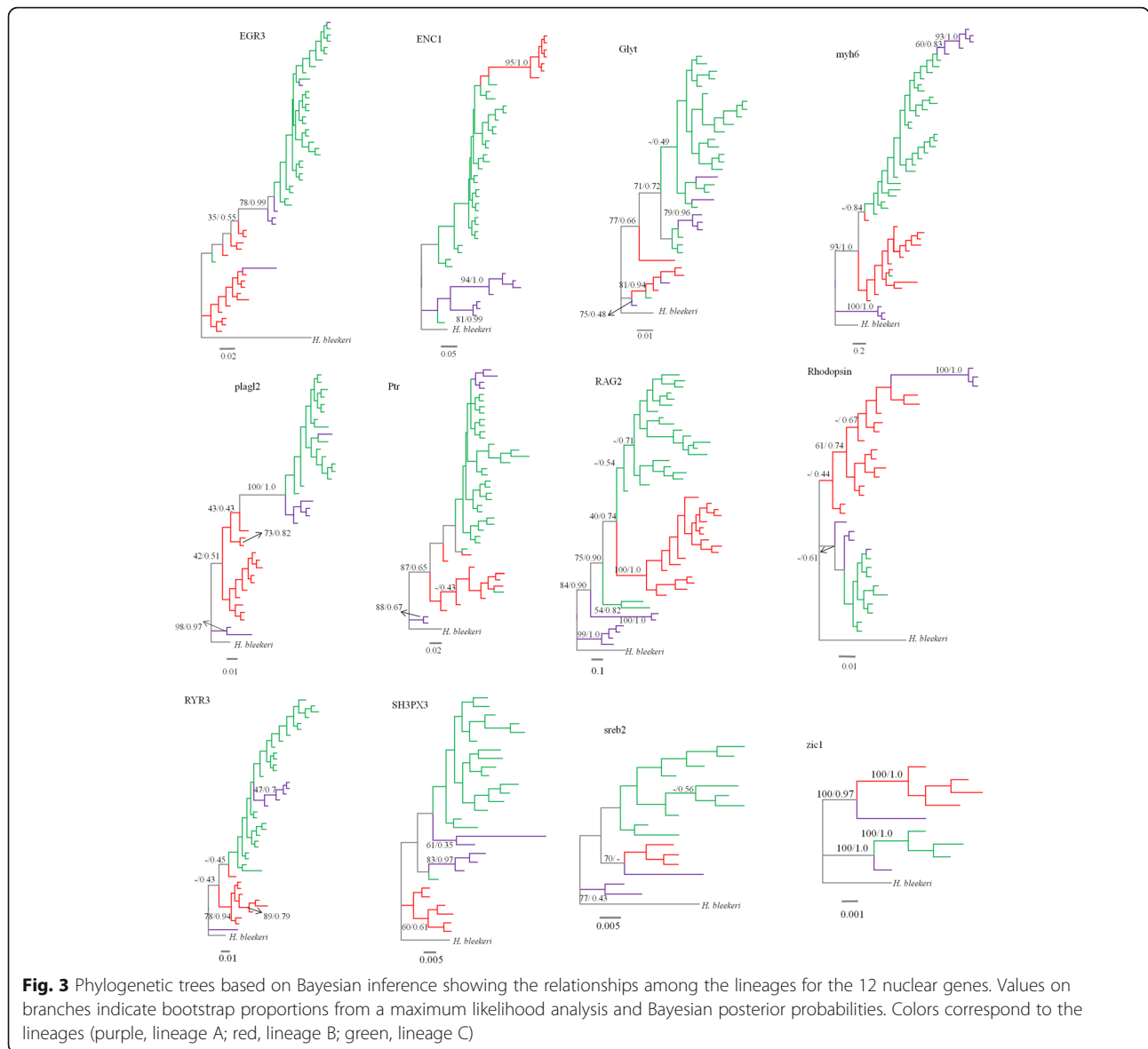
We detected three statistically significant ( $P < 0.05$ ) and unidirectional migration events among the three lineages (Fig. 5). The three migration events were from lineages C to B with a population migration rate  $2NM = 0.071$ , from lineages C to A with  $2NM = 0.29$ , and from lineages B to A with  $2NM = 0.051$ .

### Genetic diversity and population structure

Overall  $h$  was high ( $0.9737 \pm 0.0043$ ), with the highest within-lineage  $h$  occurring in lineage C ( $h = 0.9573 \pm 0.0082$ ), and was similar in lineages A ( $h = 0.868 \pm 0.0400$ ) and B ( $h = 0.891 \pm 0.0320$ ) (Table 1). Overall  $\pi$  was  $0.0290 \pm 0.0013$ , with the highest within-lineage  $\pi$  in lineage A and lowest in lineage B (Table 1). The number of haplotypes ( $n$ ), private haplotypes ( $ph$ ),  $h$ , and  $\pi$  varied among populations (Additional file 2: Table S6). A total of 134

private haplotypes were defined in all populations. The uncorrected K2P distance within each lineage is listed in Table 1. The largest distance within lineages was found in the lineage A (0.82%), the distance was intermediate in the lineage C (0.73%), and smallest in the lineage B (0.45%). The interlineage genetic distances (Additional file 2: Table S7) ranged from 4.39% (between lineages A and B) to 6.87% (between lineages A and C).

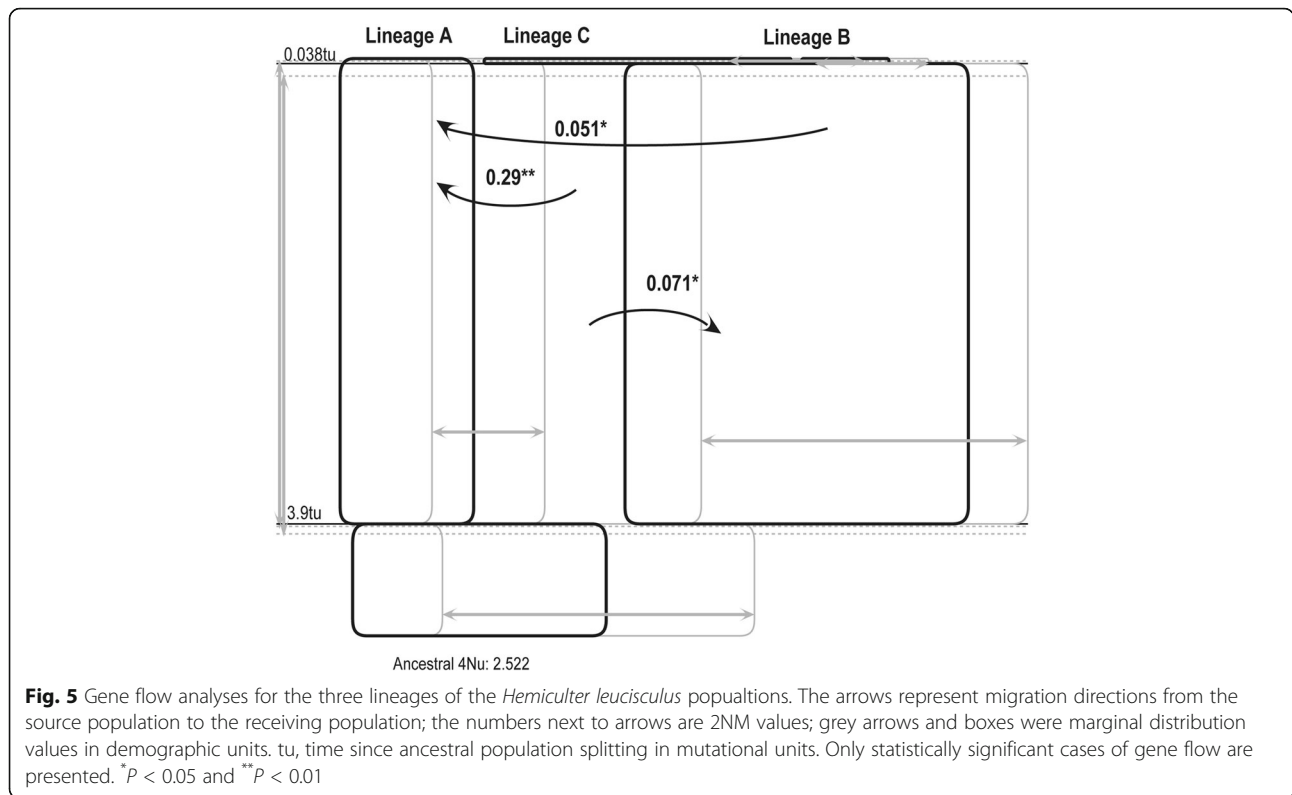
Results from the two-level hierarchical AMOVA suggested that *H. leucisculus* was highly structured geographically: 72.88% of the genetic variation was attributed to differentiation among populations, and this was highly significant ( $P < 0.001$ ) (Table 2). Furthermore, the AMOVA based on the three lineages revealed that most of the variation (89.37%,  $P < 0.001$ ) occurred among groups (Table 2). Lastly, the AMOVA based on drainage systems accounted for 59.21% of the overall variation



among groups and accounted for 22.32% among populations within groups (Table 2). One hundred and one out of 136 pairwise  $\phi_{ST}$  values were statistically significant ranging from 0.048 to 0.942 ( $P < 0.001$  to  $P = 0.043$ ). A significant correlation between the pairwise genetic differentiations ( $\phi_{ST}/(1-\phi_{ST})$ ) and the geographical distance was found over the entire dataset (Mantel test;  $r = 0.27$ ,  $P = 0.015$ ), indicative of isolation by distance (IBD).

#### Demographic history

A neutrality test conducted on the *Cytb* suggested that lineages B and C experienced a demographic expansion, as indicated by significant negative values of Fu's  $F_S$ -statistics (Table 1). Similar results were demonstrated by the EBSP analyses (Fig. 6). EBSPs suggested that demographic expansion occurred approximately 0.025 Ma for



lineage B and approximately 0.08 Ma for lineage C. Neither Fu's  $F_S$  nor EBSP detected significant expansion for lineage A.

**Discussion**

**Distribution patterns and three lineages**

In this study, we clarify the phylogeographical structure and hidden diversity of *H. leucisculus* in the drainages of southern China in detail. We confirmed the existence of three highly divergent mtDNA lineages (A, B, and C) based on Cytb sequences. Lineages A and B display a highly geographical structure. Members of lineage A originate from the middle and lower Pearl River drainage and Hainan Island, and most representatives of lineage B are limited to a small area (Localities 5–7) in the Upper Yangtze River. Regarding lineage A, the populations from middle and lower Pearl River drainages and Hainan Island were unexpectedly found to show a close

relationship, though they are separated by the Qiongzhou Strait. This may be a result of the fact that the Gulf of Tonkin and the Qiongzhou Strait were once part of the coastal plain of the Asian continent during the Pleistocene glaciations [78, 79]. Populations from lineage B are restricted to the Upper Yangtze River, suggesting that they originated and developed from these regions.

In contrast, lineage C is poorly correlated with its geographic distribution and is extensively distributed in many isolated drainage basins, including the Lancangjiang River

**Table 1** Genetic diversity and neutrality test statistics for each lineage based on Cytb

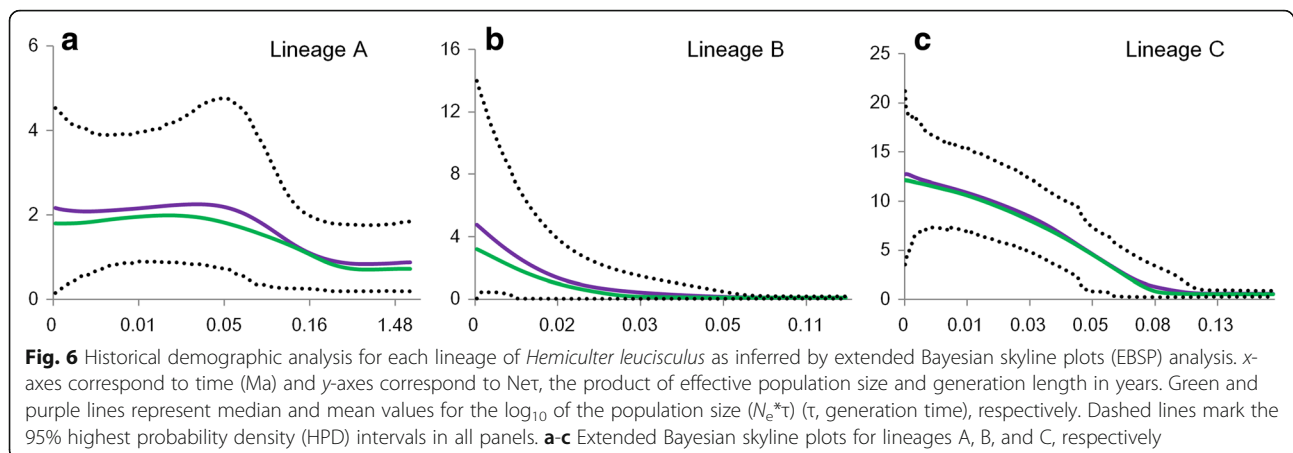
	n/N	h	$\pi$	Fu's	K2P
A	28/65	0.891 ± 0.0320	0.0081 ± 0.0004	-5.36	0.82%
B	21/50	0.868 ± 0.0400	0.0044 ± 0.0005	-6.72*	0.45%
C	122/275	0.957 ± 0.0082	0.0072 ± 0.0003	-24.40***	0.73%

N, individual numbers; n, haplotype numbers; h, haplotype diversity;  $\pi$ , nucleotide diversity; K2P, Kimura's two-parameter (K2P) distance within each lineage. \* $P < 0.05$ , \*\*\* $P < 0.001$

**Table 2** Results of an analysis of molecular variance (AMOVA) for three grouping options of *Hemiculter leucisculus* populations estimated using  $\phi$ -statistics based on Cytb

Source of variation	Percentage of variation	$\phi_{CT}$	P
Grouped by clades			
Among clades	89.37	0.89	< 0.001
Within clades	10.63	0.11	
Grouped by drainages			
Among drainages	59.21	0.59	< 0.001
Among populations, within drainage	22.32	0.22	< 0.001
Within populations	18.47	0.18	< 0.001
Grouped by sample sites			
Among populations	72.88	0.73	< 0.001
Within populations	27.12	0.27	





(Localities 1 and 2), the Upper Pearl River (Localities 3 and 8), the Yangtze River, the Minjiang River, the Jiulongjiang River, the Lingjiang River and the Qiangtangjiang River (Fig. 1). The close relationship of *H. leucisculus* populations between the Yangtze River basin and the four coastal rivers in southeastern China (i.e., the Jiulongjiang River, the Lingjiang River, the Minjiang River, and the Qiangtangjiang River) may due to their spatial proximity. The spatial proximity among these rivers enables *H. leucisculus* populations from the Yangtze River basin to expand to nearby rivers easily due to floods in summer. The hypothesis of active river capture in southwestern China [80] may explain the close relationship of populations between the Yangtze River and Upper Pearl River (Localities 3 and 8). Active river capture and reverse events may have caused the Yangtze River and Upper Pearl River to connect in some regions. However, no Cytb haplotype shared between Nanpanjiang River populations (Locality 3) and Yangtze River populations (Additional file 2: Table S6) indicated that there may have been rare gene exchange between the two populations, at least in mtDNA level. Lower genetic diversity and high levels of Cytb haplotype sharing with the Yangtze River populations (Additional file 2: Table S6) indicated that the Lancangjiang River populations (Localities 1 and 2) likely originated from the Yangtze River populations due to imprudent introduction by human in company with the large-scale introduction of economically valuable fish half a century ago [44, 81, 82].

In addition to the three strongly supported mtDNA lineages, a high level of mtDNA genetic difference observed between the lineages (4.39–6.87%) supported a deep divergence among *H. leucisculus* populations (Additional file 2: Table S7). This finding is common in small freshwater fish species in China, e.g., *Zacco platypus* [42, 83], *Opsariichthys bidens* [41], *Rhynchocypris oxycephalus* [3] and *Garra orientalis* [79].

We speculate that *H. leucisculus* originated in the drainages of southern China at 3.80 Ma (95% HPD, 2.92–4.65 Ma), suggesting that this species originated near the beginning of the Late Pliocene [84, 85]. This period is broadly consistent with the intense uplift of the Qinghai–Tibetan Plateau and evolution of Asian monsoons during 3.6–2.6 Ma [86, 87]. The divergence time estimation between lineage A and lineages B and C is at ~2.38 Ma (95% HPD, 1.82–2.93 Ma), which fits well with this timescale as well. The drastic uplift of the plateau and evolution of Asian monsoons largely re-shaped the landscape features of eastern Asia, including drainage systems [88, 89], which have been hypothesized to be important driving forces of vicariant speciation or intra-specific divergence in many fish species [3, 18, 90, 91]. Similar effects are reported for many taxa, including plants [92, 93] and other animals [12, 31, 94].

Lineages B and C, which comprised most individuals from the Yangtze River basin, separated at ~1.61 Ma (95% HPD, 1.19–1.98 Ma), which corresponds to the hypothetical formation of the current Yangtze River [95]. Historically, the Upper Yangtze River flowed southward into the Paleo-Red River [20, 96] and the middle lower Yangtze River flowed into the East China Sea [96]. Although the exact time for the connection at the Three Gorges is controversial, this appears to have occurred sometime between the Late Pliocene and Early Pleistocene, when the Yangtze River shifted its drainage network [95]. The time of splitting between the two lineages estimated in our study fits this timescale. This historical break is recorded in the level of genetic divergence between the populations of many taxa [21, 22]. Considering only Cytb was used to estimate divergence times among the lineages, our presented hypotheses should be treated cautiously.

In addition to the influence of the geologic movements, poor dispersal ability may play prominent roles in the divergence of *H. leucisculus* populations. Firstly,

the lower dispersal potential inferred from Mantel analyses indicates that geographic distance acts as a barrier for gene exchange between populations and may lead to genetic divergence. Furthermore, Cytb analyses suggest that *H. leucisculus* is geographically restricted to local areas, which may be due to the poor dispersal potential. For example, approximately 78.4% of Cytb haplotypes were only detected in single locality, which mainly distributed in Localities 3–12 and 17–26 (Additional file 2: Table S6). The AMOVA and  $\Phi_{ST}$  calculations suggest a high level of geographic structuring. This pattern may result from the isolation of populations due to habitat restriction.

### Demographic history

Cyclical Pleistocene glaciations in Europe and North America have been demonstrated to influence the population demographic history of many extant species [23, 24]. Although it is widely acknowledged that East Asia did not undergo major glaciations during the Quaternary [26, 27], a number of studies have suggested that climatic oscillations were responsible for the genetic diversity and current distribution for some taxa [35, 97–100]. A similar scenario has been detected in some extant taxa in southern China [32–34, 100].

In our study, lineages B and C of *H. leucisculus* fit this scenario. Population demographic analyses, including a neutrality test (Fu's) and EBSPs, consistently indicated that rapid population expansion has occurred in the two lineages. The absence of sublineages probably further reflects an expansion event for the two lineages. Lineage B mainly occurred in narrow regions (Localities 5–7) and exhibited low genetic diversity, which was easily influenced by fluctuant climate. In contrast, the expansion time of lineage C (~0.08 Ma) was much earlier than that of lineage B, occurring in the Late Pleistocene with the Riss-Würm interglaciation (0.073–0.125 Ma), when the climate in China was warm and wet [101, 102]. During this period, the warm climate was suitable for *H. leucisculus* survival and dispersal. Furthermore, the floods frequently occurred in summer may be a potential factor that can promote the population expansion of this species. During annual rainfall periods, the floods are common, allowing fluvial connectivity restoration and providing a contemporary condition for *H. leucisculus* population expansion. Certainly, the decreasing of competition and predation may also contribute to the population expansion of these two lineages.

However, we found no evidence of long-distance dispersal and population expansion in lineage A, which is principally distributed in southernmost China. In these temperate regions and the tropics, the relatively mild Pleistocene climate in southern China [28, 29, 103] and the varied topography provided many suitable habitats

for *H. leucisculus* [23]. The suitable habitats may still have been present in these regions during the glacial periods. Without pressure from a lack of habitat, it is possible that *H. leucisculus* inhabiting these regions might have remained in stable niches and did not experience drastic demographic fluctuations. The high level of genetic diversity reflects this scenario. Similar examples were reported in other taxa in southern China [30, 31, 104].

### Implications of cryptic speciation

Many hidden species (cryptic species) with the same morphological characteristics have been recognized using molecular tools and integrative taxonomy, which enrich species diversity but challenge traditional morphological taxonomy [105]. In our study, the BPP analysis based on nuclear genes suggests that the three lineages (A, B and C) tested are valid species. The three-lineage divergence is also supported by AMOVA results (Table 2). In addition, the three mtDNA lineages could be clearly distinguished in the three nDNA gene (ENC1, RAG2 and zic1) phylogenetic trees. Consequently, the three lineages may correspond to three separate species in *H. leucisculus* under the phylogenetic species concept [106]. Thus, given that no significant morphological differentiation among these populations was revealed in a previous study [107], *H. leucisculus* may have undergone cryptic speciation. However, the formal taxonomies of the DNA-based cryptic lineages are problematic [108]. The cryptic lineages in *H. leucisculus* detected by unlinked molecular markers merely conform to the primary species hypotheses, and more morphological analysis and phylogenetic evidence are needed to clarify formal taxonomic assignments of each lineage under the criteria of secondary species hypotheses [109].

### Gene flow between lineages

IM analysis provides unambiguous evidence of gene flow from lineages B to A, from C to A, and from C to B (Fig. 5). In the three cases, the exchange is unidirectional. It is widely acknowledged that speciation with ongoing gene flow is common [110]. The detected levels of gene flow ( $2NM = 0.051–0.29$ ) represent substantial gene flow but not high enough to prevent divergence; a  $2NM$  greater than one would limit the divergence process in the absence of selection [69, 111, 112].

### Conclusions

Our study documents the phylogeographic structure and cryptic speciation of *H. leucisculus* in the drainages of southern China. Three mtDNA lineages were observed, and each lineage represents a valid species. Intense uplift of the Qinghai–Tibetan Plateau, evolution of Asian monsoons, changes in paleo-drainages, local adaptation and poor dispersal ability may have driven the

divergence of the three putative species. However, gene flow occurs among the three lineages. Demographic analyses suggest that lineages B and C have experienced rapid demographic expansion, whereas lineage A remained stable long term during the cyclical Pleistocene glacial periods. Our results reveal the phylogeographical structure, cryptic diversity and evolutionary history of *H. leucisculus* populations and reinforce the notion that the number of widespread Chinese cyprinid species has been underestimated.

## Additional files

**Additional file 1: Table S1.** Summary of sample localities for *Hemiculter leucisculus* and outgroups. Locality numbers correspond to Fig. 1. Locality, coordinates (latitude/longitude), voucher number, and GenBank accession number for Cytb and nuclear loci are presented. Genbank numbers in bold and italic are sequences from Fan & He (2014) [1], and in bold are sequences from NCBI dataset. (DOCX 108 kb)

**Additional file 2: Table S2.** Information of primer pairs used in our study. **Table S3** Summary statistics for the 13 loci used in this study, bp, length of locus in bases; n, number of sequences; ps, polymorphic sites; pis, parsimony informative sites; LNR, length of the longest non-recombining regions, h, number of haplotypes. **Table S4** Nucleotide substitution models used in tree reconstruction. Pinvar, proportion of invariable sites; Gamma, gamma shape parameter. **Table S5** Nucleotide substitution models for Cytb and three nuDNA loci used in extend Bayesian skyline plots. **Table S6** Genetic diversity statistics for each population based on Cytb, n, individual numbers; Nh, haplotype numbers; ph, private haplotype within each population; Hap, Cytb haplotype; h, haplotype diversity;  $\pi$ , nucleotide diversity. **Table S7** Genetic distance based on Cytb among the three lineages estimated by K2P distance (%). **Figure S1** The entire distribution of *Hemiculter leucisculus* in global scale. The map derived from <http://www.discoverlife.org>. (DOCX 267 kb)

## Acknowledgements

This work was supported by the Key Fund and NSFC-Yunnan mutual funds of the National Natural Science Foundation of China (Grant Nos. 31130049 and U1036603) and the Pilot projects (Grant Nos. XDB13020100).

## Availability of data and materials

DNA sequences have been deposited in GenBank under Accession numbers KY292565–KY293216. Details regarding individual samples are available in Additional file 1: Table S1.

## Authors' contributions

WTC and SPH designed the research and wrote the manuscript. WTC, ZXZ, WD and QF performed the sampling and molecular experiments. All authors read and approved the final manuscript.

## Consent for publication

Not applicable

## Competing interests

The authors declare that they have no competing interests.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Author details

<sup>1</sup>The Key Laboratory of Aquatic Biodiversity and Conservation of Chinese Academy of Sciences, Institute of Hydrobiology, Chinese Academy of Sciences, Wuhan, Hubei 430072, China. <sup>2</sup>University of Chinese Academy of Sciences, Beijing 100049, People's Republic of China.

Received: 20 February 2017 Accepted: 30 August 2017

Published online: 12 September 2017

## References

1. Pauperio J, Herman JS, Melo-Ferreira J, Jaarola M, Alves PC, Searle JB. Cryptic speciation in the field vole: a multilocus approach confirms three highly divergent lineages in Eurasia. *Mol Ecol.* 2012;21(24):6015–32.
2. Rastorgueff PA, Chevaldonne P, Arslan D, Verna C, Lejeune C. Cryptic habitats and cryptic diversity: unexpected patterns of connectivity and phylogeographical breaks in a Mediterranean endemic marine cave mysid. *Mol Ecol.* 2014;23(11):2825–43.
3. Yu D, Chen M, Tang QY, Li XJ, Liu HZ. Geological events and Pliocene climate fluctuations explain the phylogeographical pattern of the cold water fish *Rhynchocypris oxycephalus* (Cypriniformes: Cyprinidae) in China. *BMC Evol Biol.* 2014;14:225.
4. Zhou WW, Wen Y, Fu J, Xu YB, Jin JQ, Ding L, Min MS, Che J, Zhang YP. Speciation in the *Rana chensinensis* species complex and its relationship to the uplift of the Qinghai-Tibetan plateau. *Mol Ecol.* 2012;21(4):960–73.
5. Moritz C. Defining 'evolutionarily significant units' for conservation. *Trends Ecol Evol.* 1994;9(10):373–5.
6. Andrews KR, Norton EL, Fernandez-Silva I, Portner E, Goetze E. Multilocus evidence for globally distributed cryptic species and distinct populations across ocean gyres in a mesopelagic copepod. *Mol Ecol.* 2014;23(22):5462–79.
7. Yang Z, Rannala B. Bayesian species delimitation using multilocus sequence data. *Proc Natl Acad Sci U S A.* 2010;107(20):9264–9.
8. Edwards SV, Liu L, Pearl DK. High-resolution species trees without concatenation. *Proc Natl Acad Sci U S A.* 2007;104(14):5936–41.
9. Heled J, Drummond AJ. Bayesian inference of species trees from multilocus data. *Mol Biol Evol.* 2010;27(3):570–80.
10. Ding L, Gan XN, He SP, Zhao EM. A phylogeographic, demographic and historical analysis of the short-tailed pit viper (*Gloydius brevicaudus*): evidence for early divergence and late expansion during the Pleistocene. *Mol Ecol.* 2011;20(9):1905–22.
11. Qu Y, Luo X, Zhang R, Song G, Zou F, Lei F. Lineage diversification and historical demography of a montane bird *Garrulax Elliotii*—implications for the Pleistocene evolutionary history of the eastern Himalayas. *BMC Evol Biol.* 2011;11:174.
12. Lu B, Zheng Y, Murphy RW, Zeng X. Coalescence patterns of endemic Tibetan species of stream salamanders (Hynobiidae: Batrachuperus). *Mol Ecol.* 2012;21(13):3308–24.
13. Wu X, Luo J, Huang S, Chen Z, Xiao H, Zhang Y. Molecular Phylogeography and evolutionary history of *Poropuntius Huangchuchieni* (Cyprinidae) in Southwest China. *PLoS One.* 2013;8(11):e79975.
14. Zhang DR, Chen MY, Murphy RW, Che J, Pang JF, Hu JS, Luo J, Wu SJ, Ye H, Zhang YP. Genealogy and palaeodrainage basins in Yunnan Province: phylogeography of the Yunnan spiny frog, *Nanorana yunnanensis* (Microglossidae). *Mol Ecol.* 2010;19(16):3406–20.
15. He DK, Chen YF. Phylogeography of *Schizothorax o'connori* (Cyprinidae: Schizothoracinae) in the Yarlung Tsangpo River, Tibet. *Hydrobiologia.* 2009; 635(1):251–62.
16. Shi Y, Li J, Bea L. Uplift and environmental evolution of Qinghai Xizang (Tibet) plateau. In: Sun H, Zheng D, editors. Formation, evolution and development of Qinghai-Xizang. Guangdong: Guangdong Science and Technology Press; 1998.
17. Li JJ, Fang XM. Uplift of the Tibetan plateau and environmental changes. *Chinese Sci Bull.* 1999;44(23):2117–24.
18. He SP, Cao WX, Chen YY. The uplift of Qinghai–Xizang (Tibet) plateau and the vicariance speciation of glyptosternoid fishes (Siluriformes: Sisoridae). *Sci China Ser C.* 2001;44:644–51.
19. He DK, Chen YF. Biogeography and molecular phylogeny of the genus *Schizothorax* (Teleostei: Cyprinidae) in China inferred from cytochrome b sequences. *J Biogeogr.* 2006;33(8):1448–60.
20. Clark MK, Schoenbohm LM, Royden LH, Whipple KX, Burchfiel BC, Zhang X, Tang W, Wang E, Chen L. Surface uplift, tectonics, and erosion of eastern Tibet from largescale drainage patterns. *Tectonics.* 2004;23:TC1006.
21. Yang L, Mayden RL, He S. Population genetic structure and geographical differentiation of the Chinese catfish *Hemibagrus macropterus* (Siluriformes, Bagridae): evidence for altered drainage patterns. *Mol Phylogenet Evol.* 2009;51(2):405–11.
22. Xiao W, Zhang Y, Liu H. Molecular systematics of *Xenocyprinae* (teleostei: cyprinidae): taxonomy, biogeography, and coevolution of a special group restricted in East Asia. *Mol Phylogenet Evol.* 2001;18(2):163–73.

23. Hewitt G. The genetic legacy of the quaternary ice ages. *Nature*. 2000; 405(6789):907–13.
24. Hewitt GM. Genetic consequences of climatic oscillations in the quaternary. *Philos T Roy Soc B*. 2004;359(1442):183–95.
25. Soltis DE, Morris AB, McLachlan JS, Manos PS, Soltis PS. Comparative phylogeography of unglaciated eastern North America. *Mol Ecol*. 2006; 15(14):4261–93.
26. Shi YF, Ren BH, Wang JT, Derbyshire E. Quaternary Glaciation in China. *Quaternary Sci Rev*. 1986;5:503–7.
27. Liu KB. Quaternary history of the temperate forests of China. *Quaternary Sci Rev*. 1988;7(1):1–20.
28. Ju LX, Wang HK, Jiang DB. Simulation of the last glacial maximum climate over East Asia with a regional climate model nested in a general circulation model. *Palaeogeogr Palaeoclimatol*. 2007;248(3–4):376–90.
29. Weaver AJ, Eby M, Augustus FF, Wiebe EC. Simulated influence of carbon dioxide, orbital forcing and ice sheets on the climate of the last glacial maximum. *Nature*. 1998;394(6696):847–53.
30. Gao B, Yu LJ, Qu YH, Song G, Dai CY, Zhang RY, Yin ZH, Wang KF, Gao XB, Li SH, et al. An unstructured Phylogeographic pattern with extensive gene flow in an endemic bird of South China: collared Finchbill (*Spizixos Semitorques*). *Int J Mol Sci*. 2011;12(6):3635–47.
31. Yan F, Zhou W, Zhao H, Yuan Z, Wang Y, Jiang K, Jin J, Murphy RW, Che J, Zhang Y. Geological events play a larger role than Pleistocene climatic fluctuations in driving the genetic structure of *Quasipaa Boulengeri* (Anura: Dicroglossidae). *Mol Ecol*. 2013;22(4):1120–33.
32. Qiu YX, Guan BC, Fu CX, Comes HP. Did glacials and/or interglacials promote allopatric incipient speciation in east Asian temperate plants? Phylogeographic and coalescent analyses on refugial isolation and divergence in *Diosma versipellis*. *Mol Phylogenet Evol*. 2009;51(2):281–93.
33. Zhan AB, Fu JZ. Past and present: Phylogeography of the *Bufo Gargarizans* species complex inferred from multi-loci allele sequence and frequency data. *Mol Phylogenet Evol*. 2011;61(1):136–48.
34. Huang ZH, Liu NF, Liang W, Zhang YY, Liao XJ, Ruan LZ, Yang ZS. Phylogeography of Chinese bamboo partridge, *Bambusicola Thoracica Thoracica* (aves: Galliformes) in south China: inference from mitochondrial DNA control-region sequences. *Mol Phylogenet Evol*. 2010;56(1):273–80.
35. Gao Y, Wang SY, Luo J, Murphy RW, Du R, Wu SF, Zhu CL, Li Y, Poyarkov AD, Nguyen SN, et al. Quaternary palaeoenvironmental oscillations drove the evolution of the Eurasian *Carassius Auratus* Complex (Cypriniformes, Cyprinidae). *J Biogeogr*. 2012;39(12):2264–78.
36. Bănărescu P, Coad BW. Cyprinids of Eurasia. In: *Cyprinid Fishes*. Netherlands: Springer; 1991.
37. Fu C, Wu J, Chen J, Wu Q, Lei G. Freshwater fish biodiversity in the Yangtze River basin of China: patterns, threats and conservation. *Biodivers Conserv*. 2003;12(8):1649–85.
38. Dudgeon D. Asian river fishes in the Anthropocene: threats and conservation challenges in an era of rapid environmental change. *J Fish Biol*. 2011;79(6):1487–524.
39. Fan ZX, Liu SY, Liu Y, Liao LH, Zhang XY, Yue BS. Phylogeography of the South China field mouse (*Apodemus Draco*) on the Southeastern Tibetan plateau reveals high genetic diversity and glacial refugia. *PLoS One*. 2012; 7(5):e38184.
40. Reilly SB, Wake DB. Cryptic diversity and biogeographical patterns within the black salamander (*Aneides Flavipunctatus*) complex. *J Biogeogr*. 2015; 42(2):280–91.
41. Perdices A, Sayanda D, Coelho MM. Mitochondrial diversity of *Opsariichthys Bidens* (Teleostei, Cyprinidae) in three Chinese drainages. *Mol Phylogenet Evol*. 2005;37(3):920–7.
42. Perdices A, Cunha C, Coelho MM. Phylogenetic structure of *Zacco Platypus* (Teleostei, Cyprinidae) populations on the upper and middle Chang Jiang (=Yangtze) drainage inferred from cytochrome b sequences. *Mol Phylogenet Evol*. 2004;31(1):192–203.
43. Search Fishbase [<http://www.fishbase.org>].
44. Chu XL, Chen YR. The fishes of Yunnan, China. Cyprinidae. Beijing: Sciences Press; 1989.
45. Aljanabi SM, Martinez I. Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. *Nucleic Acids Res*. 1997;25(22):4692–3.
46. Fan Q, He S. The pattern of upper and middle yangtze drainages shapes the genetic structure and diversity of hemiculter *leucisculus* revealed by mitochondrial dna locus. *Acta Hydrobiologica Sinica*. 2014;38(4):627–35.
47. Chen WJ, Miya M, Saitoh K, Mayden RL. Phylogenetic utility of two existing and four novel nuclear gene loci in reconstructing tree of life of ray-finned fishes: the order Cypriniformes (Ostariophysi) as a case study. *Gene*. 2008; 423(2):125–34.
48. Li CH, Orti G, Zhang G, Lu GQ. A practical approach to phylogenomics: the phylogeny of ray-finned fish (Actinopterygii) as a case study. *BMC Evol Biol*. 2007;7:44.
49. Lopez JA, Chen WJ, Orti G. Esociform phylogeny. *Copeia*. 2004;3:449–64.
50. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res*. 2004;32(5):1792–7.
51. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: Molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol*. 2013;30(12):2725–9.
52. Smith SA, Stephens PR, Wiens JJ. Replicate patterns of species richness, historical biogeography, and phylogeny in Holarctic treefrogs. *Evolution*. 2005;59(11):2433–50.
53. Stephens M, Smith NJ, Donnelly P. A new statistical method for haplotype reconstruction from population data. *Am J Hum Genet*. 2001;68(4):978–89.
54. Flot JF. SeqPhase: a web tool for interconverting phase input/output files and fasta sequence alignments. *Mol Ecol Resour*. 2010;10(1):162–6.
55. Woerner AE, Cox MP, Hammer MF. Recombination-filtered genomic datasets by information maximization. *Bioinformatics*. 2007;23(14):1851–3.
56. Nylander J. In: JAA N, editor. MrModeltest v2. Program distributed by the author. Uppsala: Evolutionary Biology Centre, Uppsala University; 2004.
57. Ronquist F, Huelsenbeck JP. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*. 2003;19(12):1572–4.
58. Tracer v1.4 [<http://beast.bio.ed.ac.uk/Tracer>].
59. Stamatakis A. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*. 2006; 22(21):2688–90.
60. Felsenstein J. *Inferring phylogenies*, vol. 2. Sunderland: Sinauer Associates; 2004.
61. Bandelt HJ, Forster P, Rohl A. Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol*. 1999;16(1):37–48.
62. Rannala B, Yang Z. Bayes estimation of species divergence times and ancestral population sizes using DNA sequences from multiple loci. *Genetics*. 2003;164(4):1645–56.
63. Leache AD, Fujita MK. Bayesian species delimitation in west African forest geckos (*Hemidactylus Fasciatus*). *P Roy Soc B-Biol Sci*. 2010;277(1697):3071–7.
64. Drummond AJ, Rambaut A. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol*. 2007;7:214.
65. Kumar S. Phyltest: Phylogenetic hypothesis testing software (version 2.0). Pennsylvania: The Pennsylvania State University; 1996.
66. Durand JD, Tsigonopoulos CS, Unlu E, Berrebi P. Phylogeny and biogeography of the family Cyprinidae in the Middle East inferred from cytochrome b DNA - evolutionary significance of this region. *Mol Phylogenet Evol*. 2002;25(1):218.
67. Ketmaier V, Bianco PG, Cobolli M, Krivokapic M, Caniglia R, De Matthaes E. Molecular phylogeny of two lineages of Leuciscinae cyprinids (*Telestes* and *Scardinus*) from the peri-Mediterranean area based on cytochrome b data. *Mol Phylogenet Evol*. 2004;32(3):1061–71.
68. Meyer A. *Evolution of mitochondrial DNA in fishes*, vol. 2. The Hague: Elsevier; 1993.
69. Hey J. Isolation with migration models for more than two populations. *Mol Biol Evol*. 2010;27(4):905–20.
70. Hey J, Nielsen R. Integration within the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics. *Proc Natl Acad Sci U S A*. 2007;104(8):2785–90.
71. Librado P, Rozas J. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*. 2009;25(11):1451–2.
72. Kimura M. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J Mol Evol*. 1980;16(2):111–20.
73. Excoffier L, Smouse PE, Quattro JM. Analysis of molecular variance inferred from metric distances among DNA Haplotypes - application to human mitochondrial-DNA restriction data. *Genetics*. 1992;131(2):479–91.
74. Excoffier L, Lischer HEL. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and windows. *Mol Ecol Resour*. 2010;10(3):564–7.
75. Jukes TH, Cantor CR. Evolution of protein molecules. In: Munro NH, editor. *Mammalian Protein Metabolism*. New York: Academic Press; 1969.
76. Fu YX. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*. 1997;147(2):915–25.



77. Tajima F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*. 1989;123(3):585–95.
78. Yang L, He S. Phylogeography of the freshwater catfish *Hemibarbus guttatus* (Siluriformes, Bagridae): implications for South China biogeography and influence of sea-level changes. *Mol Phylogenet Evol*. 2008;49(1):393–8.
79. Yang JQ, Hsu KC, Liu ZZ, Su LW, Kuo PH, Tang WQ, Zhou ZC, Liu D, Bao BL, Lin HD. The population history of *Garra orientalis* (Teleostei: Cyprinidae) using mitochondrial DNA and microsatellite data with approximate Bayesian computation. *BMC Evol Biol*. 2016;16:73.
80. Ren M, Bao H, Han T. The Jinsha River valley landforms and river-capture in northwestern Yunnan. *Acta Geograph Sin*. 1959;25(2):135–55.
81. Chen XY. Checklist of fishes of Yunnan. *Zool Res*. 2013;34(4):281–343.
82. Liu M, Chen D, Duan X, Ke W. Ichthyofauna composition and distribution of fishes in Yunnan section of Lancang River. *J Fishery Sci China*. 2011;18(1):156–70.
83. Zheng X, Zhou TQ, Wan T, Perdices A, Yang JQ, Tang XS, Wang ZP. Huangshan population of Chinese *Zacco platypus* (Teleostei, Cyprinidae) harbors diverse matrilineal and high genetic diversity. *Zool Res*. 2016;37(02):103–9.
84. Cui ZJ, Gao QZ, Liu GN, Pan BT, Chen HL. Planation surfaces, palaeokarst and uplift of Xizang (Tibet) plateau. *Sci China Ser D*. 1996;39:391–400.
85. Li JJ, Fang XM, Pan BT, Zhao ZJ, Song YG. Late Cenozoic intensive uplift of Qinghai-Xizang plateau and its impacts on environments in surrounding area. *Quaternary Sci*. 2001;21(5):381–91.
86. 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.
87. Zheng BX, Xu QQ, Shen YP. The relationship between climate change and quaternary glacial cycles on the Qinghai-Tibetan plateau: review and speculation. *Quaternary Int*. 2002;97:93–101.
88. Sun XJ, Wang PX. How old is the Asian monsoon system? Palaeobotanical records from China. *Palaeogeogr Palaeoclimatol*. 2005;222(3):181–222.
89. Jia GD, Peng PA, Zhao QH, Jian ZM. Changes in terrestrial ecosystem since 30 ma in East Asia: stable isotope evidence from black carbon in the South China Sea. *Geology*. 2003;31(12):1093–6.
90. Guo XG, He SP, Zhang YG. Phylogeny and biogeography of Chinese sisorid catfishes re-examined using mitochondrial cytochrome b and 16S rRNA gene sequences. *Mol Phylogenet Evol*. 2006;38(1):291.
91. Yang J, Yang JX, Chen XY. A re-examination of the molecular phylogeny and biogeography of the genus *Schizothorax* (Teleostei: Cyprinidae) through enhanced sampling, with emphasis on the species in the Yunnan-Guizhou plateau, China. *J Zool Syst Evol Res*. 2012;50(3):184–91.
92. Yuan QJ, Zhang ZY, Peng H, Ge S. Chloroplast phylogeography of *Dipentodon* (Dipentodontaceae) in southwest China and northern Vietnam. *Mol Ecol*. 2008;17(4):1054–65.
93. Liu JQ, Wang YJ, Wang AL, Hideaki O, Abbott RJ. Radiation and diversification within the *Ligularia-Cremnathodium-Parasenecio* Complex (Asteraceae) triggered by uplift of the Qinghai-Tibetan plateau. *Mol Phylogenet Evol*. 2006;38(1):31–49.
94. Zhang MW, Rao DQ, Yang JX, Yu GH, Wilkinson JA. Molecular phylogeography and population structure of a mid-elevation montane frog *Leptobrachium ailaonicum* in a fragmented habitat of southwest China. *Mol Phylogenet Evol*. 2010;54(1):47–58.
95. Li J, Xie S, Kuang M. Geomorphic evolution of the Yangtze gorges and the time of their formation. *Geomorphology*. 2001;41(2):125–35.
96. Yang DY. The formation of Yangtze River geomorphology. Beijing: The Geological Publishing House; 2006.
97. Qu YH, Ericson PGP, Lei FM, Li SH. Postglacial colonization of the Tibetan plateau inferred from the matrilineal genetic structure of the endemic red-necked snow finch, *Pyrgilauda ruficollis*. *Mol Ecol*. 2005;14(6):1767–81.
98. Huang S, He SP, Peng ZG, Zhao K, Zhao EM. Molecular phylogeography of endangered sharp-snouted pitviper (*Deinagkistrodon acutus*; Reptilia, Viperidae) in mainland China. *Mol Phylogenet Evol*. 2007;44(3):942–52.
99. Zhou W, Yan F, Fu J, Wu S, Murphy RW, Che J, Zhang Y. River islands, refugia and genetic structuring in the endemic brown frog *Rana kukunoris* (Anura, Ranidae) of the Qinghai-Tibetan plateau. *Mol Ecol*. 2013;22(1):130–42.
100. Li SH, Yeung CKL, Feinstein J, Han LX, Manh HL, Wang CX, Ding P. Sailing through the late Pleistocene: unusual historical demography of an east Asian endemic, the Chinese Hwamei (*Leucodioptron canorum canorum*), during the last glacial period. *Mol Ecol*. 2009;18(4):622–33.
101. Yuan DX, Cheng H, Edwards RL, Dykoski CA, Kelly MJ, Zhang ML, Qing JM, Lin YS, Wang YJ, Wu JY, et al. Timing, duration, and transitions of the last interglacial Asian monsoon. *Science*. 2004;304(5670):575–8.
102. Kelly MJ, Edwards RL, Cheng H, Yuan DX, Cai YJ, Zhang ML, Lin YS, An ZS. High resolution characterization of the Asian monsoon between 146,000 and 99,000 years BP from Dongge cave, China and global correlation of events surrounding termination II. *Palaeogeogr Palaeoclimatol*. 2006;236(1):20–38.
103. Pinot S, Ramstein G, Harrison SP, Prentice IC, Guiot J, Stute M, Joussaume S. Tropical paleoclimates at the last glacial maximum: comparison of Paleoclimate Modeling Intercomparison project (PMIP) simulations and paleodata. *Clim Dynam*. 1999;15(11):857–74.
104. Wang HW, Ge S. Phylogeography of the endangered *Cathaya argyrophylla* (Pinaceae) inferred from sequence variation of mitochondrial and nuclear DNA. *Mol Ecol*. 2006;15(13):4109–22.
105. Bickford D, Lohman DJ, Sodhi NS, Ng PKL, Meier R, Winker K, Ingram KK, Das I. Cryptic species as a window on diversity and conservation. *Trends Ecol Evol*. 2007;22(3):148–55.
106. De Queiroz K. Species concepts and species delimitation. *Syst Biol*. 2007;56(6):879–86.
107. Luo Y, Chen Y. *Culterinae*. Fauna Sinica, Osteichthyes, Cypriniformes II. Beijing: Science Press; 1998.
108. Satler JD, Carstens BC, Hedin M. Multilocus species delimitation in a complex of morphologically conserved trapdoor spiders (Mygalomorphae, Antrodiaetidae, Aliatypus). *Syst Biol*. 2013;62(6):805–23.
109. Pante E, Puillandre N, Viricel AE, Arnaud-Haond S, Aurelle D, Castelin M, Chenuil A, Destombe C, Forcioli D, Valero M, et al. Species are hypotheses: avoid connectivity assessments based on pillars of sand. *Mol Ecol*. 2015;24(3):525–44.
110. Nosil P. Speciation with gene flow could be common. *Mol Ecol*. 2008;17(9):2103–6.
111. Wright S. Evolution in Mendelian populations. *Genetics*. 1931;16(2):97–159.
112. Pinho C, Hey J. Divergence with gene flow: models and data. *Annu Rev Ecol Evol S*. 2010;41:215–30.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

