



Short Communication

Multiple unrelated founding events for the long-distance Pleistocene dispersal of the Salangid, *Neosalanx taihuensis*: A general demographic model for inshore-orientated freshwater fish

Liang Zhao^{a,b}, Jie Zhang^{a,c}, Zhijin Liu^a, Ming Li^{a,*}^a Key Laboratory of Animal Ecology and Conservation Biology, Institute of Zoology, Chinese Academy of Sciences, Chaoyang, Beijing 100101, China^b Faculty of Biology, Suzhou University, Suzhou, Anhui 234000, China^c Key Laboratory of Zoological Evolution and Systematics, Institute of Zoology, Chinese Academy of Sciences, Chaoyang, Beijing 100101, China

ARTICLE INFO

Article history:

Received 18 April 2010

Revised 26 October 2010

Accepted 3 November 2010

Available online 21 November 2010

Keywords:

Neosalanx taihuensis

Coalescent simulations

Long-distance dispersal

Phylogeographic history

ABSTRACT

The Salangid icefish *Neosalanx taihuensis* (Salangidae) originated from inshore of the East China seas and underwent adaptive freshwater radiation from the mid-Miocene to the early Pleistocene. The distribution of its genetic diversity presents a random pattern inconsistent with contemporary hydrological structure. In the present study, coalescent simulations were used to analyze its Pleistocene dispersal history. Population history simulation supported the hypothesis of long-distance dispersal during the Pleistocene based on multiple unrelated founding events. This analogous genetic pattern has been described for other inshore-orientated freshwater fish, and may represent a general history dispersal model for the phylogeography of these species. From network analysis, three subclades (Clades 1–3) grouped consistently with three probable ancestral haplotypes (H36, H27, and H33). Demographic analysis also revealed that the ancestral haplotype group (Clade 1) dispersed into freshwater during an interglacial age about 0.35 Ma, while Clades 2 and 3 dispersed about 0.12 and 0.145 Ma, respectively. The *N. taihuensis* population remained relatively small for a considerable amount of time during the Pleistocene ages, with population expansion events mainly occurring after the last glacial maximum (LGM).

© 2010 Elsevier Inc. All rights reserved.

1. Introduction

According to key concepts in biogeography, geological and climatic changes profoundly affect biotic distributions. For example, the separation of biomes caused by vicariant events divides previously continuous species ranges, enabling population genetic differentiation or speciation. Alternatively, geological dynamics can reduce or even eliminate physical barriers to range expansion through the formation of temporary land corridors. Under these circumstances, species can extend their range via neighbor migration in a process termed “geodispersal” (Lieberman, 2000). Range expansion can also occur via long-distance dispersal (LDD) involving movement beyond neighboring demes. Understanding the relative contribution of vicariance, geodispersal, and LDD to the composition of contemporary biotas is one of the most challenging goals of biogeography. If congruence between geological connections and population genetic patterns can be demonstrated for a species at temporal and spatial levels, then vicariant and/or geodis-

persal explanations for current distributions can be accepted; if congruence at either level is lacking, however, LDD over existing barriers can be assumed (Mansion et al., 2008). Discerning the formations of biotic distributions among the above-mentioned processes is, however, still in its infancy.

Populations of freshwater fish species from different basins often show significant genetic differentiation resulting from isolation, while populations within a basin demonstrate no or low levels of genetic differentiation (Cortey et al., 2004), which can be explained, to some degree, using the isolation by distance (IBD) model hypothesis. However, our recent study on Chinese Salangid icefish (*Neosalanx taihuensis*) showed that significant levels of genetic subdivision existed among populations within basins rather than between basins, demonstrating a complex pattern that could not be interpreted simply by the IBD model (Zhao et al., 2008). Several researchers have shown the genetic structure in some freshwater fish populations to be analogous to our previous study (Cortey et al., 2004; McGlashan and Hughes, 2000). This provided evidence that the contemporary drainage structure did not coincide with genetic relationships among populations. Instead, the observed population relationships agreed with the population structure predicted by normal dispersal models of recently expanding populations due to multiple unrelated long-distance

* Corresponding author. Address: Institute of Zoology, Chinese Academy of Sciences, Datun Rd., Chaoyang District, Beijing 100101, China. Fax: +86 10 6480 7099.

E-mail address: lim@ioz.ac.cn (M. Li).

migrants founding pocket populations (Kamal et al., 1996). However, verifying the genetic differentiation formation of a species in a rigorous statistical framework remains difficult in the field of phylogeography. One way is to simulate genetic evolution within competing hypotheses of population history, with the simulated genealogies compared to phylogenies inferred from sequence data to assess which hypothesis represents a best-fit for the data (Spellman and Klicka, 2006).

Studying the phylogeographic pattern and its formations for *N. taihuensis* can provide important information regarding conservation options and the accurate assignation of appropriate fishery management. In the present study, complete mitochondrial *Cyt b* sequence data was used to examine its phylogeography and demographic history. Our research aimed to address the following questions: (i) Is the complex genetic pattern of *N. taihuensis* coherent with a scenario of multiple unrelated founding events via LDD as predicted by normal dispersal models? (ii) Is genetic divergence within the species due largely to the effects of LGM and/or the result of earlier events during the Pleistocene?

2. Materials and methods

2.1. Data description and phylogenetic analysis

A total of 354 *N. taihuensis* individuals collected from 13 populations in the Pearl River (P1–P2), Yangtze River (Y1–Y7), and Huai River (H1–H4) basins were analyzed (Table 1 and Fig. 1a). Of all samples collected, 36 *Cyt b* haplotypes were identified and deposited in the GenBank (EU376454–EU376489). A thorough description of the data set can be found in Zhao et al. (2008).

For phylogenetic analysis, we performed maximum-likelihood (ML) analyses using the program PAUP* 4.0 (Swofford, 2002). The TrN + I ($I = 0.8148$) model was selected as the best-fit model for analysis using MODELTEST 3.06 (Posada and Crandall, 1998), with *Protosalanx chinensis* (DQ191115), a close relative of *N. taihuensis* (Zhang et al., 2007), used as an outgroup. Statistical parsimony, as implemented by TCS version 1.18 (Clement et al., 2000), was then used as an unbiased criterion for assigning subclades.

2.2. Tests of LDD model based on coalescent simulations

A pair-wise relative rate test (pRRT) implemented in HyPhy (Pond et al., 2005) was used to estimate and verify molecular clock constancy at the intraspecific level for *N. taihuensis*. Analysis

revealed that one haplotype (H31) evolved at a significantly different pace to the others. After removing H31, no other violation of molecular clock constancy was found. Consequently, the other 35 identified haplotypes were used for further analysis.

Coalescent simulations of genealogies constrained within models of population divergence provide a powerful means for assessing the fit of observed genetic patterns to different phylogeographical hypotheses (Spellman and Klicka, 2006). Here, two alternative hypotheses were tested using Mesquite 2.5 (Maddison and Maddison, 2008). The first corresponded to a single founding event dispersal model, and the second corresponded to multiple unrelated founding events. Under the first hypothesis, speciation events for *N. taihuensis* populations occurred in the East China Seas (Zhang et al., 2007) and sequentially dispersed into freshwater by a single founding event. Under the second hypothesis, the genetic structure of *N. taihuensis* populations resulted from multiple unrelated founding events. Two types of coalescent simulations were performed in the present study (Fig. 2a and b). In the first set of simulations, 1000 coalescent genealogies were generated under each historical scenario and the distribution of S , the minimum number of sorting events required to explain population subdivision (Slatkin and Maddison, 1989), was recorded. During this simulation, overall effective population size (N_e) and the lower and upper bounds of the 95% CI for N_e were used as model parameters. In the second set of simulations, 100 gene matrices constrained within alternative models were simulated using DNA substitutions selected by Modeltest 3.06. The PAUP* 4.0 program was used to reconstruct trees from the simulated gene matrices, and S -values for these trees were recorded. The second set of simulations only used overall N_e for the model parameters.

For all simulations, N_e was estimated using θ -values calculated with FLUCTUATE 1.4 (Kuhner et al., 1998) and all runs employed the following strategy: 10 short chains of 4000 steps and five long chains of 400,000 steps, with sampling every 20th step; random starting trees. Runs were repeated five times to ensure consistency of estimates. The θ -values were converted to N_e using the formula $\theta = 2N_e\mu$. Since there is no fossil record to analyze for *N. taihuensis*, the genetic time clock is difficult to calibrate. However, a calibration rate of 0.8–1% per Myr per generation has been accepted for mitochondrial protein-coding genes in Salmoniformes and Osmeriformes in many studies (McCusker et al., 2000; Sheldon et al., 1996). Following Zhang et al. (2007) and Zhao et al. (2008), we used a mutation rate of $\mu = 1\%$ in this study.

Table 1
Localities, sample size, and number of haplotypes of *N. taihuensis* population.

Population	Location	Latitude	Longitude	<i>N</i>	<i>H</i>
Pearl River basin				44	5
P1	Luofu River	22°30'33"	114°03'25"	31	5
P2	Pearl River	22°34'59"	113°29'24"	13	3
Yangtze River basin				197	25
Y1	Taihu lake	31°30'12"	120°30'06"	31	8
Y2	Chaohu lake	31°30'12"	117°12'30"	25	8
Y3	Bohu lake	30°06'06"	116°24'06"	29	5
Y4	Poyanghu lake	29°0'30"	115°48'48"	22	5
Y5	Dongtinghu lake	28°48'30"	111°54'06"	27	5
Y6	Tianrezhou	29°18'48"	113°24'0"	34	5
Y7	Xujiahe Reservoir	31°49'05"	113°50'40"	29	2
Huai River basin				113	15
H1	Hongzehu lake	33°16'20"	118°42'12"	26	8
H2	Weishanhu lake	35°14'42"	116°38'08"	31	4
H3	Wabuhu lake	32°22'35"	116°54'22"	30	2
H4	Chengdonghu lake	32°18'41"	116°22'57"	26	6
Entire region (all samples)				354	36

Note: *N*, number of individuals; *H*, number of haplotypes.

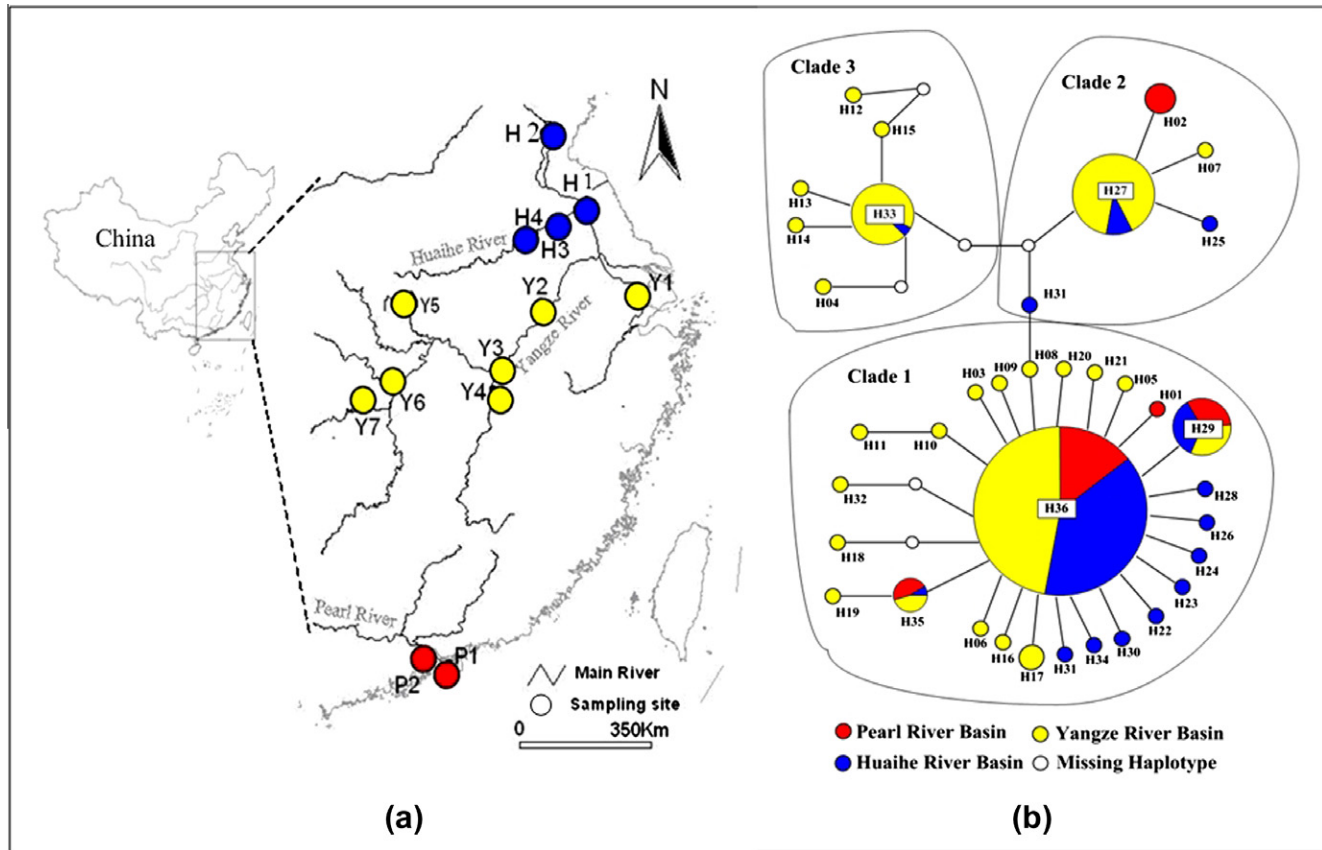


Fig. 1. Map of sampling locations of *N. taihuensis* and network redrawn from Zhao et al. (2008). (a) Sampling locations and (b) minimum spanning network: nodes contain the haplotype number and are proportional to the haplotype frequency, white nodes indicate undetected intermediate haplotype states separated by one mutational step, and boxes indicate one-step to two-step nesting levels for the nested clades.

2.3. Demographic analysis

Several approaches were used to analyze recent and historic population demography of *N. taihuensis*. First, changes over time in effective population size were evaluated with Bayesian skyline plots using BEAST 1.4.1 (Drummond and Rambaut, 2007). We used the HKY model, uncorrelated log-normal relaxed clock, linear skyline with twelve groups for 10,000,000 iterations, after a burn-in phase of 100,000 iterations. Tracer 1.4 (Rambaut and Drummond, 2007) was used to plot the Bayesian skyline. We then examined the observed distribution of pair-wise differences between sequences (mismatched distribution) using Arlequin 3.11 (Excoffier et al., 2007), with significance assessed by parametric bootstraps (10,000 replicates). Second, for all populations and for each subclade detected by TCS, D^* (Fu and Li, 1993), F_S (Fu, 1997) and R_2 (Ramos-Onsins and Rozas, 2002) tests for mutation/drift equilibrium were performed in DnaSP 4.10 (Rozas et al., 2003) with 10,000 simulations. The D^* test is designed to detect an excess of old mutations, characteristic of a population that has experienced a historical reduction in effective population size (Fu, 1996; Fu and Li, 1993). In contrast, F_S and R_2 are sensitive to an excess of recent mutations (Fu, 1997; Ramos-Onsins and Rozas, 2002).

Divergence time for each subclade was estimated by the Bayesian phylogenetic method implemented in BEAST 1.4.1. To gain a more detailed picture of the mutational pattern from population expansion, GENETREE (Bahlo and Griffiths, 2000) was also used to construct a gene tree to determine the distribution of each *N. taihuensis* mutation. GENETREE is a coalescent method that simulates gene trees conditional on their topology and can be used to determine the distribution of the most recent common ancestor

(TMRCA). In the present study, the most probable ancestral haplotype was reconstructed by the Mesquite 2.5 program using parsimony methods. The TMRCA for a constant population was then calculated in GENETREE by finding the maximum-likelihood estimate of θ (Joy et al., 2003). Each run was repeated three times with a different starting seed number and 1000,000 coalescent simulations.

3. Results

3.1. Phylogenetic analysis and tests of the LDD model

Phylogenetic inference revealed that *N. taihuensis* formed a monophyletic lineage with respect to outgroup *P. chinensis* (ML tree not shown). Only a few missing haplotypes were detected in network, which indicated that sufficient samples had been collected for phylogeographical analysis. The H36 haplotype yielded the highest outgroup weight (0.2941) followed by H27 (0.0588), H33 (0.0588), H29 (0.0441), and H35 (0.0441), indicating that H36, H27, and H33 were the most probable ancestral haplotypes. Three subclades (Clades 1–3) grouped consistently with these ancestral haplotypes. Clade 1 was the dominant group and contained 25 different haplotypes, which produced a pattern of population relationships associated with haplotype ancestry within each population rather than with the hydrological network (Fig. 1b).

Using FLUCTUATE 1.4, an overall θ_{total} of 0.0094 (95% CI 0.0052–0.0186) was estimated from the data set. Using 1% substitutions per site per million years and a generation time of 1 year, a point

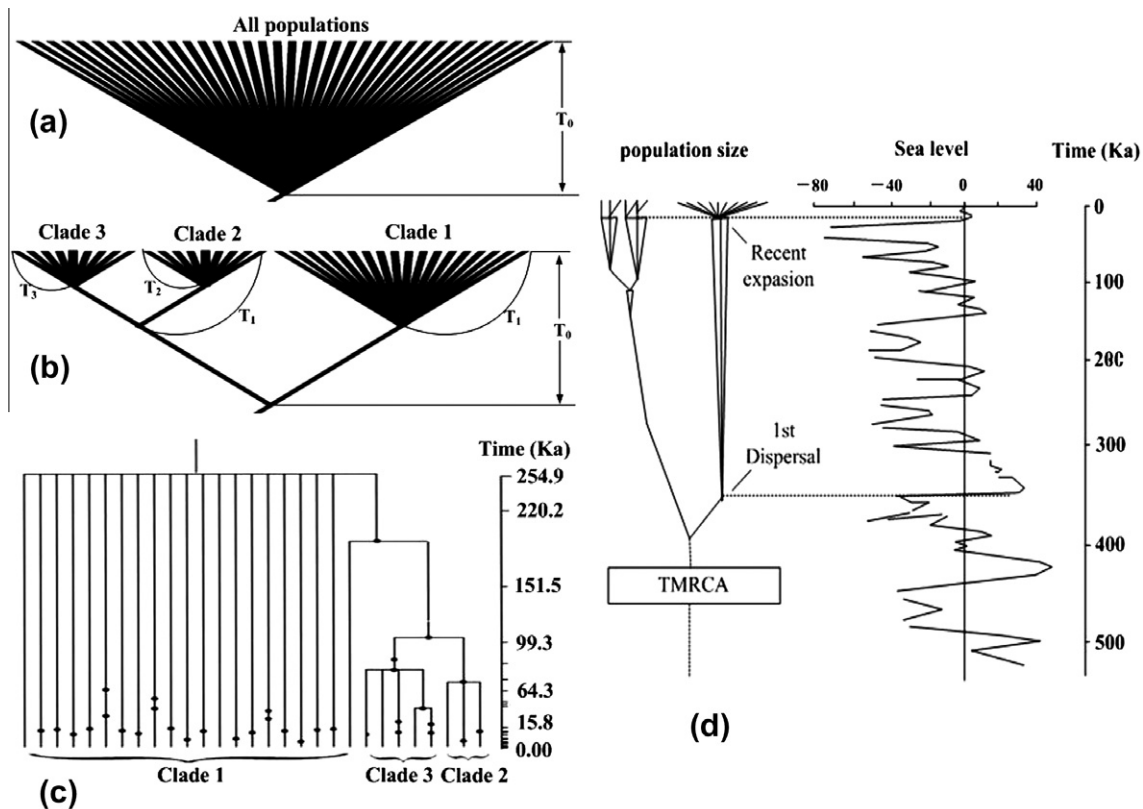


Fig. 2. Models used to test LDD population founding hypotheses, gene tree and proposed evolutionary history of *N. taihuensis*. (a) Dispersal model hypothesis based on a single founding event, which assume that current populations were derived from a single founding event of dispersal, (b) dispersal model hypothesis based on two founding events, which assume that ancestral populations separated into mainland freshwater by multiple unrelated founding events, (c) gene tree of the *N. taihuensis* mitochondrial *Cyt b* gene. The tree is based on 1000,000 coalescent simulations. The absolute time scale on the right shows the TMRCA (kyr) in generations using a mutation rate of 1×10^{-8} per year. The tree shows the ancestral distribution of mutations and events (TMRCA, recent rapid expansion) for population history and (d) proposed evolutionary history of *N. taihuensis* highlighting fluctuations in population size and dispersal events. The fluctuation of the sea levels is redrawn from Butzer (1964) and Chappell (1974). Ka is kilo years ago.

estimate of $N_{\text{total}} = 470,000$ was calculated for this species, with a 95% CI ranging from 260,000 to 930,000.

For the ML genealogy, a value of $S = 5$ was calculated. From the first set of simulations, it was not possible to reject either hypothesis model. For the second set of simulations (100 gene simulated gene matrices constrained within each model), the $\text{TrN} + I$ ($I = 0.8148$) model was used to generate the gene matrices. Under these simulations, the single founding event of dispersal model was rejected and the multiple unrelated founding events model was accepted ($P = 0.10$, $T_2 = 0.013$ Ma, $T_1 = 0.36$ Ma). A dispersal hypothesis based on multiple unrelated founding events was therefore suitable for this species.

3.2. Demographic analysis

For all populations, the Bayesian skyline plots (Fig. 1S) showed recent population expansion. The mtDNA mismatch analyses for all populations showed a bimodal profile, which might result from a

constant population size among an old population or an admixture population (Fig. 2S). There was no significant mismatch in the distribution goodness-of-fit test (Table 2), however, which suggested no severe departure from the estimated demographic expansion model. Additionally, for all populations and each subclade (Clades 1–3), the F_S and R_2 tests showed a statistically significant departure from the neutrality hypothesis. Although a bottleneck could produce an analogous pattern in the neutrality test, the D^* tests were not statistically significant for these populations, suggesting there was no historical reduction in effective population size for this species (Table 2).

The Bayesian phylogenetic method was used to estimate divergence times for each subclade, which were determined to be 0.350 (95% HPD 0.2827–0.5446), 0.120 (HPD 0.0551–0.2191), and 0.145 (HPD 0.0652–0.2550) Ma for Clades 1–3, respectively. GENETREE was used to construct a gene tree to determine mutation time distribution and gain a more detailed picture of the mutational pattern for *N. taihuensis* population expansion (Fig. 2c). GENETREE is

Table 2

Statistical test for neutrality, mismatch analysis, and the estimate of demographic parameters for *N. taihuensis* based on mitochondrial *cyt b* sequence data.

	Hri (P value)	SSD (P value)	D^* (P value)	R_2 (P value)	F_S (P value)
All populations	0.0675 (0.8590)	0.0388 (0.3002)	−0.0591 (0.4570)	0.0726 (0.0160)*	−18.9472 (0.0002)**
Clade 1			−0.0499 (0.4590)	0.0881 (0.0000)**	−32.054 (0.0000)**
Clade 2			−0.0345 (0.4540)	0.1198 (0.0000)**	−3.6060 (0.0057)**
Clade 3			−0.0977 (0.4950)	0.0859 (0.0020)**	−9.2412 (0.0019)**

* $P < 0.05$.

** $P < 0.01$.

implemented based on DNA sequence data under a strict infinitely many sites model. In analysis, GENETREE tested the second site among all variable nucleotide sites violated this model. To ensure the sequence data was suitable for the infinite-sites model, four haplotypes (H06, H07, H10, and H13) were automatically excluded after this second site was removed. It is worth noting that the possible removal of data can affect inference and lead to unusual conclusions, but the results from this procedure can provide a rough profile for the distribution of mutations. The results revealed that the vast majority of mutations occurred near the tips of the gene tree, with Clade 1 forming the most ancient founding populations and Clades 2 and 3 representing subsequent founding events.

4. Discussion

4.1. Phylogeographic pattern and its formations

Parsimony network analysis revealed a lack in geographical structure, which was consistent with the AMOVA and the Mantel test in our previous study (Zhao et al., 2008). Phylogeographic studies often reveal disparate levels of sequence divergence between lineages spanning a common geographic barrier, leading to the conclusion that isolation is nonsynchronous. Only rarely, however, do researchers account for the expected variance associated with ancestral coalescence and among-population variation in demographic history (Joy et al., 2003; Zhao et al., 2008). In the present study, three subclades (Clades 1–3) grouped consistently with the probable ancestral haplotypes (H36, H27, and H33) found within each population rather than with the hydrological network. This agrees with the population structure predicted by normal dispersal models of recently expanding populations due to rare long-distance migrant founding events (Kamal et al., 1996).

The coalescent simulations of genealogies constrained within the alternative models of population divergence found in this study provided evidence to reject the single founding event dispersal model and accept the multiple unrelated founding events model ($P = 0.10$, $T_2 = 0.013$ Ma, $T_1 = 0.36$ Ma). The mtDNA mismatch analyses also showed a bimodal profile, which might result from a constant population size among an old population or an admixture of divergent lineages (Fig. 2S). Demographic analysis did not support a relatively constant population size in *N. taihuensis* (Table 2). In terms of population demographic scenarios, therefore, multimodal distributions in mismatch analysis might be expected for multiple unrelated dispersal events for this species. Thus, we suggest that a dispersal hypothesis based on multiple unrelated founding events is suitable for this species. Additionally, speciation events for the Salangidae family occurred in the East China Seas (Zhang et al., 2007), indicating that sequential multiple unrelated long-distance founding events are a reasonable scenario for the current genetic structure of *N. taihuensis*. Research on other inshore-orientated freshwater fish species, specifically *Coilia nasus* (Yang et al., 2008) and *Neosalanx jordani* (Zhao, 2008), revealed very similar phylogeographic patterns. As with *N. taihuensis*, current populations of these species exhibited closely related maternal haplotypes throughout the Chinese mainland, with population relationship patterns associated with ancestral haplotypes rather than within the hydrological network. Although further studies should be conducted on other inshore-orientated freshwater fish, the phylogeographic dispersal model of *N. taihuensis* could be considered a general demographic model for this kind of species in this area.

4.2. Demographic dispersal procedure

The haplotype network clades revealed star-like patterns around each of the widespread haplotypes, a pattern indicative

of rapid ancestral haplotype expansion over a large geographic area (Joy et al., 2003). This was supported by Bayesian skyline plots and the F_S and R_2 tests (Fig. 1S and Table 2). The estimated divergence time for Clades 1–3 were 0.350 (95% HPD 0.2827–0.5446), 0.120 (HPD 0.0551–0.2191), and 0.145 (HPD 0.0652–0.2550) Ma, respectively. This indicates that colonization of *N. taihuensis* in mainland China pre-dated the LGM. It is worth noting that the timing of Clade 1 expansion was consistent with an interglacial period occurring 0.35 Ma, while Clades 2 and 3 were consistent with an interglacial period occurring 0.1–0.15 Ma (Butzer, 1964; Chappell, 1974). It is likely that appropriate large-scale habitats formed with the warming climate and rising sea levels present during the interglacial period, and that *N. taihuensis* founders expanded randomly into different locations in mainland China by water flow. Our gene tree results provide a detailed profile on how the Pleistocene ice ages affected the demographic history of this species. The vast majority of mutations occurred near the tips of the gene tree, which suggests that the *N. taihuensis* population remained relatively small for a considerable amount of time due to periodic climatic oscillations in the Pleistocene ages, with the majority of population expansion events occurring after the LGM. Interestingly, most recent mutations were from Clade 1, indicating this population was growing faster than the others. This also suggests that the rapid growth of the oldest founding populations and the subsequent resource competition may have prevented the other founding populations from simultaneous expansion (Hewitt, 2000).

4.3. Conclusion and some caveats

A summary of the evolutionary history model for *N. taihuensis* is shown in Fig. 2d. In the present study, coalescent simulations were used to test alternative dispersal history models for *N. taihuensis*, with results supporting a dispersal hypothesis based on multiple unrelated founding events. Considering the Salangidae family originated from the East China Seas, multiple unrelated long-distance founding events reasonably explain the formation of this species' current genetic structure. Furthermore, the phylogeographic dispersal model of *N. taihuensis* might represent a general demographic model for inshore-orientated freshwater fish in this area.

However, it is appropriate that statistical phylogeography method, which could consider complex histories with many processes, does not exist yet. Genetic histories are so strongly influenced by chance events that inferences should be made cautiously, particularly as information from any one genealogy may be misleading and conclusions are difficult to determine if lineage sorting is incomplete. Furthermore, the coalescent simulations used in this study rely heavily upon model parameters (θ) estimated directly from the data, so the chance of erroneous inferences may be high if coalescent simulations are conducted using poor and/or incorrect estimates of θ . Estimates of θ from single locus data can be unreliable and always have a high degree of error associated with them, so uncertainty in the gene tree should be considered (Spellman and Klicka, 2006). Beyond the uncertainty of the statistical phylogeography methodology mentioned above, the *Cyt b* gene used in the present study is only a single matrilineal locus; therefore it also represents only one observation of the *N. taihuensis* evolutionary history. It would be interesting to test whether multi-locus genes exhibit the same demographic history for *N. taihuensis*, and most importantly, whether the demographic dispersal process for *N. taihuensis* can be confirmed in other species.

Acknowledgments

This project was supported by the National Basic Research Program of China (973 Program: 2007CB411600), the National Natural

Science Foundation of China (No. 30570256), and the Natural Science Key Program of Anhui Province Office of Education (KJ2009A052Z). The authors would like to thank Dr. L.F. Zhu, and Dr. B.W. Zhang for their technical guidance and data analysis, and other staff for their help in collecting samples. Special thanks are given to Dr. T. Meng and Dr. D.W. Qi for providing geographic information and maps.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at [doi:10.1016/j.ympev.2010.11.003](https://doi.org/10.1016/j.ympev.2010.11.003).

References

- Bahlo, M., Griffiths, R.C., 2000. Inference from gene trees in a subdivided population. *Theoretical Population Biology* 57, 79–95.
- Butzer, K.W., 1964. *Environment and Archeology*. Aldine Pub. Co., Chicago.
- Chappell, J., 1974. Geology of coral terraces, Huon Peninsula, New Guinea: a study of Quaternary tectonic movements and sea-level changes. *Geological Society of America Bulletin* 85, 553–570.
- Clement, M., Posada, D., Crandall, K.A., 2000. TCS: a computer program to estimate gene genealogies. *Molecular Ecology* 9, 1657–1659.
- Cortey, M., Pla, C., Garcia-Marín, J.L., 2004. Historical biogeography of Mediterranean trout. *Molecular Phylogenetics and Evolution* 33, 831–844.
- Drummond, A.J., Rambaut, A., 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology* 7, 214.
- Excoffier, L., Laval, G., Schneider, S., 2007. Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics* 2005, 47–50.
- Fu, Y.X., 1996. New statistical tests of neutrality for DNA samples from a population. *Genetics* 143, 557–570.
- Fu, Y.X., 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* 147, 915–925.
- Fu, Y.X., Li, W.H., 1993. Statistical tests of neutrality of mutations. *Genetics* 133, 693–709.
- Hewitt, G., 2000. The genetic legacy of the Quaternary ice ages. *Nature* 405, 907–913.
- Joy, D.A., Feng, X., Mu, J., Furuya, T., Chotivanich, K., Krettli, A.U., Ho, M., Wang, A., White, N.J., Suh, E., Beerli, P., Su, X.-Z., 2003. Early origin and recent expansion of *Plasmodium falciparum*. *Science* 300, 318–321.
- Kamal, M.I., Richard, A.N., Godfrey, M.H., 1996. Spatial patterns of genetic variation generated by different forms of dispersal during range expansion. *Heredity* 77, 282–291.
- Kuhner, M.K., Yamato, J., Felsenstein, J., 1998. Maximum likelihood estimation of population growth rates based on the coalescent. *Genetics* 149, 429–434.
- Lieberman, B.S., 2000. *Paleobiogeography: Using Fossils to Study Global Change, Plate Tectonics, and Evolution*. Kluwer Academic Publishers, New York.
- Maddison, W.P., Maddison, D.R., 2008. MESQUITE: A Modular System for Evolutionary Analysis. Version 2.5. <<http://www.mesquiteproject.org>>
- Mansion, G., Rosenbaum, G., Schoenenberger, N., Bacchetta, G., Rossello, J.A., Conti, E., 2008. Phylogenetic analysis informed by geological history supports multiple, sequential invasions of the Mediterranean Basin by the Angiosperm family Araceae. *Systematic Biology* 57, 269–285.
- McCusker, M.R., Parkinson, E., Taylor, E.B., 2000. Mitochondrial DNA variation in rainbow trout (*Oncorhynchus mykiss*) across its native range: testing biogeographical hypotheses and their relevance to conservation. *Molecular Ecology* 9, 2089–2108.
- McGlashan, D.J., Hughes, J.M., 2000. Reconciling patterns of genetic variation with stream structure, earth history and biology in the Australian freshwater fish *Craterocephalus stercusmuscarum* (Atherinidae). *Molecular Ecology* 9, 1737–1751.
- Pond, S.L.K., Frost, S.D.W., Muse, S.V., 2005. HyPhy: hypothesis testing using phylogenies. *Bioinformatics* 21, 676–679.
- Posada, D., Crandall, K.A., 1998. MODELTEST: testing the model of DNA substitution. *Bioinformatics* 14, 817–818.
- Rambaut, A., Drummond, A.J., 2007. Tracer v1.4. <<http://www.beast.bio.ed.ac.uk/Tracer>>.
- Ramos-Onsins, S.E., Rozas, J., 2002. Statistical properties of new neutrality tests against population growth. *Molecular Biology and Evolution* 19, 2092–2100.
- Rozas, J., Sanchez-DelBarrio, J.C., Messeguer, X., Rozas, R., 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* 19, 2496–2497.
- Sheldon, J.M., Robert, H.D., Michael, J.S., 1996. Phylogeny of Pacific salmon and trout based on growth hormone type-2 and mitochondrial NADH dehydrogenase subunit 3 DNA sequences. *Canadian Journal of Fisheries and Aquatic Sciences* 53, 1165–1176.
- Slatkin, M., Maddison, W.P., 1989. A cladistic measure of gene flow inferred from the phylogenies of alleles. *Genetics* 123, 603–613.
- Spellman, G., Klicka, J., 2006. Testing hypotheses of Pleistocene population history using coalescent simulations: phylogeography of the pygmy nuthatch (*Sitta pygmaea*). *Proceedings of the Royal Society B: Biological Sciences* 273, 3057–3063.
- Swofford, D., 2002. *Paup*: Phylogenetic Analysis Using Parsimony (* and Other Methods)*, Version 4.0b10. Sinauer Associates, Sunderland, Massachusetts.
- Yang, J.Q., Hu, X.L., Tang, W.Q., 2008. Genetic structure and population history of *Coilia nasus* in Yangtze River and its south adjacent waters. *Journal of Shanghai Fisheries University* 17, 513–519.
- Zhang, J., Li, M., Xu, M.Q., Takita, T., Wei, F.W., 2007. Molecular phylogeny of icefish Salangidae based on complete mtDNA cytochrome *b* sequences, with comments on estuarine fish evolution. *Biological Journal of the Linnean Society* 91, 325–340.
- Zhao, L., 2008. Population Genetic Structure and Demographic History of the *Neosalanx taihuensis* and *Neosalanx jordani* Based on Cytochrome *b* Sequences. Doctoral Dissertation Submitted to the Institute of Zoology Chinese Academy of Sciences, pp. 77–92.
- Zhao, L., Zhang, J., Liu, Z., Funk, S., Wei, F., Xu, M., Li, M., 2008. Complex population genetic and demographic history of the Salangid, *Neosalanx taihuensis*, based on cytochrome *b* sequences. *BMC Evolutionary Biology* 8, 201.