Molecular Phylogenetics and Evolution 58 (2011) 142–147

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

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).

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 geodispersal 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).

1055-7903/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.ympev.2010.11.003

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 (*Ne*) and the lower and upper bounds of the 95% CI for *Ne* 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 *Ne* for the model parameters.

For all simulations, *Ne* 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 *Ne* using the formula $\theta = 2Ne\mu$. 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	*N*	*H*
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
Yangze 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.