Molecular phylogeography and population structure of a mid-elevation montane frog *Leptobrachium ailaonicum* in a fragmented habitat of southwest China

Mingwang Zhang a,b,1, Dingqi Rao a,1, Junxing Yang a,*, Guohua Yu a, Jeffery A. Wilkinson c,*

a Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, Yunnan 650223, China
b Graduate School of the Chinese Academy of Sciences, Beijing 100039, China
c Department of Herpetology, California Academy of Sciences, 55 Music Concourse Drive, San Francisco, CA 94118, USA

**A R T I C L E  I N F O**

Article history:
Received 18 February 2009
Revised 14 September 2009
Accepted 9 October 2009
Available online 19 October 2009

Keywords:
Conservation
Mitochondrial DNA
Orogeny
Pleistocene
Secondary contact
Vicariance

**A B S T R A C T**

*Leptobrachium ailaonicum* is a vulnerable anuran restricted to a patchy distribution associated with small mountain streams surrounded by forested slopes at mid-elevations (approximately 2000–2600 m) in the subtropical Mount Wuliang and Mount Ailao ranges in southwest China (Yunnan Province) and northern Vietnam. Given high habitat specificity and lack of suitable habitat in lower elevations between these ranges, we hypothesized limited gene flow between populations throughout its range. We used two mitochondrial genes to construct a phylogeographic pattern within this species in order to test our hypothesis. We also examined whether this phylogeographic pattern is a response to past geological events and/or climatic oscillations. A total of 1899 base pairs were obtained from 81 individuals of nine populations yielding 51 unique haplotypes. Both Bayesian and maximum parsimony phylogenetic analyses revealed four deeply divergent and reciprocally monophyletic mtDNA lineages that approximately correspond to four geographical regions separated by deep river valleys. These results suggest a long history of allopatric separation by vicariance. The distinct geographic distributions of four major clades and the estimated divergence time suggest spatial and temporal separations that coincide with climatic and paleogeographic changes following the orogeny and uplift of Mount Ailao during the late Miocene to mid Pliocene in southwest China. At the southern distribution, the presence of two sympatric yet differentiated clades in two areas are interpreted as a result of secondary contact between previously allopatric populations during cooler Pleistocene glacial cycles. Analysis of molecular variance indicates that most of the observed genetic variation occurs among the four regions implying long-term interruption of maternal gene flow, suggesting that *L. ailaonicum* may represent more than one distinct species and should at least be separated into four management units corresponding to these four geographic lineages for conservation.

© 2009 Elsevier Inc. All rights reserved.

1. Introduction

Past geologic events and climatic fluctuations have had profound impacts on the phylogeographic structure and genetic patterns of species in southwest China (Liu et al., 2007; Yuan et al., 2008). These influences have resulted in a high number of species in the temperate, subtropical and tropical mountains of this region, which are the major components of the South-Central China and Indo-Burman biodiversity hotspots (Myers et al., 2000; Wikramanayake et al., 2001; CEPF, 2005; CI, 2007). Two hypotheses have been proposed to explain the production of such high diversity in this region. One hypothesis is that extreme topographic variation of deep river valleys between recently uplifted mountains associated with the collision of the Indian subcontinent and the mainland of Asia created climatic conditions that are complex and diverse both altitudinally and latitudinally resulting in a high abundance of species (Yang, 1991; Zhao, 1999). The Hengduan Mountains, beginning at the southeast corner of the Tibetan Plateau, are north to south ranges of high mountains with alternating deep river valleys, which formed and now accommodate the main stems of some of the largest rivers in China and Southeast Asia (i.e., Irrawaddy, Mekong, Red, Salween, and Yangtze; Li et al., 1995; Brookfield, 1998). At the southern subsection of the Hengduan mountain chain are the Wuliang and Ailao mountains, which stretch across central-south Yunnan from northwest to the southeast (Fig. 1). The Ailao and Wuliang mountains are considered part of the Indo-Burman Hotspot, as they are located at the convergence of the Central and South Asian tropical zones (Myers et al., 2000; CI, 2007).
The other hypothesis put forth to explain the high diversity of this region is that the southern temperate and tropical regions in China were refugia for various animals and plants during Pleistocene glacial periods (Wu, 1980, 1987; Wang and Liu, 1994; Li et al., 2005; Long et al., 2006). Though most of China has never been covered by ice sheets, it, together with neighboring areas in eastern Asia, has experienced a development of cooler and drier climates within the last 15-million year period (Axelrod et al., 1996). The tremendous climatic changes during this period, particularly Quaternary glaciations, have led to many extinctions and influenced the distribution and evolution of many plants and animals in China and its neighboring areas (Wang and Ge, 2006). However, to date there is little to no information on the phylogeographic patterns of most of the species from subtropical and tropical mountain regions of Asia (Hewitt, 2000). The tremendous climatic changes during this period, particularly Quaternary glaciations, have led to many extinctions and influenced the distribution and evolution of many plants and animals in China and its neighboring areas (Wang and Ge, 2006). However, to date there is little to no information on the phylogeographic patterns of most of the species from subtropical and tropical mountain regions of Asia (Hewitt, 2000). Thus, we wanted to test whether historical changes in geography (such as orogenesis) or Quaternary climatic oscillations affected the phylogeographic patterns of the mustache toad (*Leptobrachium ailaonicum*) in Yunnan.

*Leptobrachium ailaonicum* is a medium size toad belonging to an unusual group restricted to southern China and Vietnam in which the males possess external cornified spines on the maxillary region and are larger than the females (reverse sexual size dimorphism). Based on these characters, this small group of toads was separated into its own genus *Vibrissaphora* (Liu, 1945). *Leptobrachium ailaonicum* was first described by Yang et al. (1983) with recent taxonomic changes by Rao and Wilkinson (2008), who synonymized *L. echinatum* from Vietnam into *L. ailaonicum*. Until recently this species was known only from localities in the northern parts of the Mount Wuliang and Mount Ailao ranges in central Yunnan Province, China (Yang et al., 1983; Chen et al., 1984; Zhao, 1998; Fei, 1999; Fei et al., 2005; Yang and Rao, 2008) and the Fan Si Pan Mountains of northern Vietnam (Dubois and Ohler, 1998; Ho et al., 1999). However, our fieldwork has considerably extended the known range of this species by the discovery of several populations between these known localities, and two localities east of the Red River (Fig. 1). We failed to find new localities to the west of the Mekong River and along the Mount Ailao range between the northern and southern localities even though we searched over the past decade in suitable habitat.

This species is uncommon and highly adaptive to montane environments, inhabiting moist evergreen broad-leaf primary forests at elevations between 2000 and 2600 m above sea level but absent below 1500 m (Zhao, 1998; Fei, 1999; personal field observations). This species only breeds in moderately broad forested mid to high elevation mountain streams that have a relatively slow flow of cold clear water at an average depth of about 10 cm (Chen et al., 1984; Ho et al., 1999; personal field observations). We postulate that the extensive low-elevation “sea” surrounding the higher elevation montane “islands” acts as a strong dispersal barrier that isolates local populations resulting in strong genetic subdivision of the species due to its specialized ecological requirements.

![Fig. 1. The geographic distribution of *Leptobrachium ailaonicum* sampled in this study. WS, Wenshan; PB, Pingbian; MAD, Maandi; BMH, Baimahe; FSL, Fenshuiling; XLS, Xilongshan; LC, Luchun; JD, Jingdong; NJ, Nanjian. Solid black dots indicate the presence of haplotypes from Lineage A, white dots from Lineage B, dots with horizontal hatching from Lineage C, and dots with vertical hatching from Lineage D. The two dots at XLS and FSL indicate that haplotypes of Lineage A are present at those localities with individuals of Lineage B and C, respectively. Grey shading signifies elevations of 1500 m and above.](image)
Due to potential fragmentation and deterioration of its habitat, this species is listed as Near Threatened by the International Union for the Conservation of Nature and Natural Resources (IUCN, 2006). Also, recent global declines in amphibian communities have been especially prominent in montane taxa (Stuart et al., 2004), although a mechanism has not yet been pinpointed. Therefore, an understanding of the genetic structure of existing populations is necessary to design effective conservation plans for this species, which may be beneficial to other anuran species in this region.

The objectives of our study are (1) to test the prediction that *L. ailaonicum* is made up of highly differentiated populations, (2) to identify the phylogenetic and geographic relationships among these populations if they are highly differentiated, (3) to investigate if past geologic events and/or climatic oscillations are responsible for observed genetic phylogeographic patterns, and (4) to assess whether the current management strategies for this species are sufficient to maintain its long-term survival. We approached these questions by examining the genetic variation in the mitochondrial cytochrome *b* (cyt *b*) and NADH dehydrogenase 4 (ND4) genes, and sequence for transfer RNA for leucine (tRNA^Leu^) within and among populations of *L. ailaonicum*.

2. Materials and methods

2.1. Sampling of specimens

We analyzed a total of 81 individuals from nine localities that span the entire known range of *L. ailaonicum* in China (Table 1). Sample sites included one site (Nanjian, NJ) in the Mount Wuliang range, one site (Jingdong, JD) from the north, and five sites (Luchun, LC, Xilongshan, XLS, Fenshuiling, FSL, Baimahe, BMH, and Maandi, MAD) from the south of the Mount Ailaiao range, and two sites (Pingbian, PB and Wenshan, WS) from east of the Red River (Fig. 1). Two individuals of *L. boringiae* and one individual of *L. promustache* were used as outgroups, with *L. promustache* as the most distant outgroup based on the results of Rao and Wilkinson (2008), and Zheng et al. (2008). Samples were obtained from either tadpole fin clips preserved in ethanol or toe clips from live specimens subsequently sequenced at the point of capture. Tissue samples were preserved in the field in 95% ethanol and transferred later to a −20 °C freezer.

2.2. DNA extraction, polymerase chain reaction, and sequencing

We extracted DNA using a standard phenol/chloroform method (Sambrook et al., 1989). The entire cyt *b* gene was amplified with the primers GLU-5' and CB6THR-3' of Palumbi et al. (1991). Polymerase chain reactions (PCR) were run in a total volume of 50 μL in a GeneAmp PCR System 9700 thermocycler (Applied Biosystems) as follows: pre-denaturing at 94 °C for 5 min, cycling at 94 °C for 1 min, 48 °C for 0.5 min, and 72 °C for 1 min for 35 cycles, and a final extension at 72 °C for 5 min. Additionally, primers ND4 and ND4-tRNA^Leu^ of Arévalo et al. (1994) were used to amplify an approximately 900 bp ND4-tRNA^Leu^ fragment. PCR conditions were the same as for cyt *b* except that the annealing temperature was 50 °C.

We conducted all sequencing on ABI PRISM 3730 automated sequencers. We sequenced all individuals in both directions, and used fully confirmed sequences in most cases with occasionally unconfirmed, unambiguous ends of sequences determined by visual inspection. All sequences were deposited in GenBank (Accession Nos. GQ503897–GQ504004). Sequences were examined for signal quality and confirmed for complementarity using DNASTAR (DNASTAR Inc.) and aligned by SEQUENCE NAVIGATOR (Applied Biosystems).

2.3. Analyses of genetic diversity and phylogeny

We computed numbers of haplotypes (*N*) and values of haplotype diversity (*h*; Nei, 1987) and nucleotide diversity (*π*; Nei and Tajima, 1981) using DNASP 4.0 (Rozas et al., 2003). We performed a partition-homogeneity test using 1000 replicates as implemented in paup 4.10b (Swofford, 2002) in order to examine whether the two analyzed regions (cyt *b* and ND4-tRNA^Leu^) could be combined into a larger data matrix (Farris et al., 1995). Because the result of the partition-homogeneity test was not significant, further analyses were performed on the combined data. Phylogenetic trees were estimated using Bayesian inference and maximum parsimony (MP) based on these combined data. Because many haplotypes were identical, we first merged all redundant OTUs using MACCLADE 4.0 ( Maddison and Maddison, 2000).

Prior to the Bayesian phylogenetic analysis, we selected the settings for the DNA substitution model that best fit the data by a hierarchical likelihood-ratio test using the programs MODELTEST.

<table>
<thead>
<tr>
<th>Regional groups</th>
<th>Sampling localities</th>
<th>Latitude (N)</th>
<th>Longitude (E)</th>
<th>Phylogroup</th>
<th>Haplotypes present(number of specimens)</th>
<th>N</th>
<th>Haplotype diversity</th>
<th>Nucleotide diversity</th>
</tr>
</thead>
<tbody>
<tr>
<td>East Range</td>
<td>Pingbian (PB)</td>
<td>22°54’</td>
<td>103°41’</td>
<td>Lineage D</td>
<td>H42(1), H43(5), H44(2), H45(1), H46(1)</td>
<td>10</td>
<td>0.911</td>
<td>0.00451</td>
</tr>
<tr>
<td></td>
<td>Wenshan (WS)</td>
<td>23°21’</td>
<td>103°55’</td>
<td>Lineage D</td>
<td>H47(3), H48(2), H49(2), H50(2), H51(1)</td>
<td>10</td>
<td>0.975</td>
<td>0.00979</td>
</tr>
<tr>
<td>Middle Range</td>
<td>Maandi (MAD)</td>
<td>22°46’</td>
<td>103°30’</td>
<td>Lineage C</td>
<td>H1(1), H2(2), H3(4), H4(1), H5(1), H6(1), H7(1)</td>
<td>11</td>
<td>0.933</td>
<td>0.02740</td>
</tr>
<tr>
<td></td>
<td>Baimahe (BMH)</td>
<td>22°48’</td>
<td>103°19’</td>
<td>Lineage C</td>
<td>H16(1), H17(1), H18(1), H19(1), H20(1)</td>
<td>5</td>
<td>0.825</td>
<td>0.00089</td>
</tr>
<tr>
<td></td>
<td>Fenshuiling (FSL)</td>
<td>22°52’</td>
<td>103°14’</td>
<td>Lineage C</td>
<td>H8(2), H9(2), H10(1), H11(1), H12(1), H13(1), H14(1), H15(2), H21(1), H22(1)</td>
<td>2</td>
<td>0.911</td>
<td>0.00451</td>
</tr>
<tr>
<td>South Range</td>
<td>Xilongshan (XLS)</td>
<td>22°41’</td>
<td>102°47’</td>
<td>Lineage B</td>
<td>H24(1), H25(1), H27(1)</td>
<td>3</td>
<td>0.825</td>
<td>0.00089</td>
</tr>
<tr>
<td></td>
<td>Lineage A</td>
<td></td>
<td></td>
<td></td>
<td>H23(2), H26(1)</td>
<td>3</td>
<td>0.974</td>
<td>0.03684</td>
</tr>
<tr>
<td>West Range</td>
<td>Luchun (LC)</td>
<td>22°53’</td>
<td>102°17’</td>
<td>Lineage A</td>
<td>H28(1), H29(1)</td>
<td>2</td>
<td>0.825</td>
<td>0.00089</td>
</tr>
<tr>
<td></td>
<td>Jingdong (JD)</td>
<td>24°32’</td>
<td>101°01’</td>
<td>Lineage A</td>
<td>H31(1), H33(1), H34(7), H35(1), H36(1)</td>
<td>11</td>
<td>0.974</td>
<td>0.03684</td>
</tr>
<tr>
<td></td>
<td>Nanjian (NJ)</td>
<td>24°48’</td>
<td>100°25’</td>
<td>Lineage A</td>
<td>H30(1), H32(1), H34(4), H37(2), H38(1), H39(1), H40(2), H41(1)</td>
<td>13</td>
<td>0.974</td>
<td>0.03684</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>81</td>
<td>0.974</td>
<td>0.03684</td>
</tr>
</tbody>
</table>
3.6 (Posada and Crandall, 1998) and PAUP* 4.0b10 (Swofford, 2002). Our phylogenetic reconstruction for Bayesian inference was based on this model. We ran Bayesian analyses with MRBAYES version 3.0b4 (Hueslenbeck and Ronquist, 2001). We performed Metropolis-coupled Markov chain Monte Carlo sampling with four chains that were run for $3 \times 10^6$ generations, sampling every 100th tree. We plotted $-\ln L$ scores against generations, and discarded $-\ln L$ scores obtained prior to the chains reaching stationarity as burn-in.

We also performed a MP analysis using PAUP* 4.0b10 (Swofford, 2002) to provide an alternative to the Bayesian phylogeny. In the MP analysis, data were treated with equal weight. We performed 100 heuristic searches with the starting trees obtained by random stepwise addition, followed by tree-bisection-reconnection branch-swapping. Gaps in tRNA LEU were treated as a fifth character. We also performed a nonparametric bootstrap analysis (Felsenstein, 1985) with 100 replicates, each executed as a heuristic search as above, to evaluate support for relationships, as implemented in PAUP* 4.0b10.

We calculated estimates of sequence divergence among the lineages of *L. ailaonicum*, correcting for variation among haplotypes within each of the lineages (e.g., Edwards, 1997), as $p$-distances (pairwise genetic distance) in mega 4 (Tamura et al., 2007). To test for constancy among lineages, we calculated the likelihood scores with and without enforcing a molecular clock using PAUP* 4.0b10 (Swofford, 2002) and performed a likelihood-ratio test (LRT) to compare both likelihoods. Because the data did not conform to expectations of clock-like evolution ($X^2 = 496$, d.f. = 87, $P < 0.01$), we used Bayesian molecular dating (Thorne et al., 1998) to date splits using the MULTIDIVTIME package of Thorne and Kishino (2002).

We were unable to calibrate the timing of clade divergences using this dataset because of the absence of appropriate fossils and/or geologic events as independent sources for dating within *Leptobrachium* or the family Megophryidae. We therefore expanded our inclusion of taxa to other pelobatoid species, such as species within Pelobatidae, which is considered closely related to Megophryidae (García-París et al., 2003).

We used default settings of MULTIDIVTIME as recommended by Rutschmann (2005) except for the following parameters: $rtrate = 0.07$, $rtratesd = 0.07$, $rttmsd = 2$, $bigtime = 25.0$; with $rtrate$ and $rtratesd$ in millions of years. We defined the vicariance event of Europe separating from Africa following the end of the Messinian salinity crisis at $5.33 \pm 0.02$ million years ago (Ma) (Krijgsman et al., 1999) as the only reliable calibration point for our expanded data set. This geological event is thought to be responsible for a number of speciation events in amphibians (e.g., Busack et al., 1985; Fromhage et al., 2004), among them *Pelobates varaldii* and *P. cultripes* (García-París et al., 2003; Veith et al., 2006; Crottini et al., 2007). To estimate the time of divergence among clades, we used only *cyt b* sequences (702 bp) because sequences of ND4 of *P. varaldii* are not available. We used homologous *cyt b* sequences of three other species of *Pelobates* (*P. cultripes, P. syriacus* and *P. varalidi*) and of *Spea bombifrons* and *S. multiplicata* (GenBank Accession Nos. DQ333373, DQ333372, EF191042, EF191043, and EF191044) for hierarchical outgroup rooting according to the phylogeny in García-París et al. (2003) to construct a Bayesian phylogeny, using the above mentioned protocols, to be used in the dating analysis.

The *cyt b* gene of frogs and toads have been demonstrated to have a rate of divergence of $0.7–1.6\%$ per Myr (Macey et al., 1998; Jaeger et al., 2005; Eggert et al., 2006; Hofman et al., 2007), though some ranids appear to have faster rates (e.g., Babik et al., 2004; Palo et al., 2004). We therefore applied a $0.7–1.6\%$ range for estimating the divergences between any major clades in our study as a comparison with the results obtained using MULTIDIVTIME.

2.4. Analyses of geographic structuring

We implemented a hierarchical analysis of molecular variance (AMOVA; Excoffier et al., 1992) using Arlequin version 3.0 (Excoffier et al., 2005) to assess the most probable population configuration and geographic subdivision. Populations were grouped into 2, 3, 4, 5, and 6 groups, respectively, according to different geographic hierarchies that matched the mtDNA lineages recovered in the phylogenetic analyses or according to geographic proximity. We also evaluated other possible groupings of populations regardless of geographic or phylogenetic proximity. The groupings that maximized values of among group variation ($F_{CT}$) and were statistically significant indicated the most parsimonious geographic subdivisions.

We calculated unrooted statistical parsimony haplotype networks with the program TCS 1.21 (Clement et al., 2000). The network was nested according to the rules outlined in Templeton et al. (1995) and Templeton (2004). We subsequently used GEODIS 2.5 (Posada et al., 2000) to perform a nested clade analysis (NCA) to test the association between haplotypes and geography. Geographic associations were estimated from the distance parameters $D_x (x)$, $D_y (x)$, $I-T$, $D_y$ and $T$. The clade distance $D_x (x)$ is the mean distance of all individuals of clade $x$ from the geographic center of clade $x$. The nested clade distance $D_y (x)$ measures how far bearers of clade $x$ haplotypes are from all bearers of the higher nested clade $y$ (which includes clade $x$), $I-T$, $D_y$ and $T$. The distance differences between interior (ancestral) and tip clades within clades, $D_y$, or nested within nested clades, $D_y$ (Templeton et al., 1995; Templeton, 1998). To test whether these distances were significantly small or large at the $5\%$ level, we used 1000 permutations with the Monte Carlo technique (Roff and Bentzen, 1989). We interpreted the potential historical and biological associations (i.e., allopatric fragmentation, continuous range expansion, isolation by distance, long distance colonization, long distance dispersal, past fragmentation, restricted gene dispersal, restricted gene flow) using the inference keys in Templeton (1998, 2004) for those clades in which the null hypothesis of random geographic distribution was rejected.

2.5. Analyses of population structuring

We calculated mismatch distributions for each of the groups in the haplotype networks to assess graphically the changes in population size inferred by NCA. Assuming the infinite sites model, the mismatch distribution is smooth and often unimodal as a result of population expansion whereas for stationary populations the distribution is ragged and often multimodal (Harpending et al., 1998). We also applied Fu’s Fs neutrality test (Fu, 1997) as an additional assessment of possible population expansion. Under the assumption of neutrality, a population expansion produces a large negative value of $Fs$ (Fu, 1997).

We tested for evidence of population expansion within lineages by examining the mismatch distributions using Arlequin version 3.0 (Excoffier et al., 2005). We used the equation $\tau = 2ut$ (Nei and Tajima, 1981; Rogers and Harpending, 1992) to estimate the approximate expansion time in generations (t) for *L. ailaonicum* populations, where $\tau$ is the mode of mismatch distribution expressed in units of evolutionary time, $t$ is the date of growth or decline measured in units of mutational time, and $u$ is the mutation rate per sequence and per generation. Finally, we calculated the approximate time of expansion in years by multiplying $t$ by the generation time (3 years; Chen et al., 1984) of *L. ailaonicum*. We
assumed a mitochondrial divergence rate of 0.7–1.6% per Myr for *L. ailaonicum* as explained above.

3. Results

3.1. DNA sequence variation and genetic diversity

In total, we identified 51 haplotypes among the 81 *L. ailaonicum* individuals based on the combined data (1141 bp of the mitochondrial *cyt b* and 848 bp ND4-*tRNA LEU* sequences). These 51 haplotypes are listed in Table 1 to show their distributions in subpopulations. No premature stop codons, ambiguous nucleotides in translation, deletions, or insertions were observed in either the *cyt b* or ND4 gene sequences, suggesting that these sequences represent functional genes. There were 250 variable sites, of which 212 were parsimony informative for just the ingroup individuals, and 485 variable sites with 330 parsimony informative when the outgroup taxa were included.

3.2. Phylogenetic analysis and divergence time estimation

A Bayesian tree on the *cyt b* and ND4-*tRNA LEU* sequences was constructed with the TrN + G model following the parameter settings: Nst = 6; base frequencies A = 0.2551, C = 0.2677, G = 0.1431, and T = 0.3342; shape parameter of the gamma distribution = 0.1796, and Pinvar = 0. In the Bayesian analysis, two independent runs were conducted to avoid entrapment in local optima. In L scores reached stationarity at or prior to 5000 generations, the 50% majority rule consensus of trees was computed after discarding 50 trees of each run as burn-in.

In both Bayesian and MP phylogenetic analyses, the 51 haplotypes of *L. ailaonicum* observed in the combined dataset formed four distinct clades (Fig. 2). In all cases, the four *L. ailaonicum* mtDNA clades were supported with high bootstrap (BS = 100) and Bayesian posterior probability (BPP = 1.00) values; Lineage D, which includes all individuals collected from PB and WS; Lineage C including all individuals collected from MAD, BMH, and most specimens of FSL; Lineage B, including most specimens from XLS; and Lineage A, including all individuals collected from NJ, JD, and LC and some specimens from XLS and FSL (Fig. 2). In general, haplotypes from the same locality clustered together. However, it is notable that the individuals from the FSL locality did not cluster together within one of the four clades; eight haplotypes (H8, H9, H10, H11, H12, H13, H14, H15) from FSL are part of Lineage C, while two haplotypes (H21 and H22) are part of Lineage A, indicating that members of two clades (Lineages A and C) coexist at FSL. Likewise, Lineage A and B haplotypes occurred together within XLS. Furthermore, both analyses recovered Lineage A as the sister group of Lineage B with strong support (BPP = 1.00, BS = 97; Fig 2). However, placement of Lineages C and D differed between the Bayesian and MP trees. The Bayesian analysis recovered the A + B clade as the sister group to Lineage C with weak support (BPP = 0.51, not shown) and a more basal branch leading to Lineage D (Fig 2A). There was moderate support

---

Fig. 2. Bayesian phylogram (A) and the strict consensus tree from the parsimony analysis (B) of the observed haplotypes of *L. ailaonicum*, with *L. boringiae*, and *L. promustache* as outgroups. Bayesian posterior probabilities are given only for those nodes which obtained values > 0.95 with this method. For the parsimony tree, numbers beside nodes are bootstrap proportions greater than 75.
from the MP analysis for the grouping of the A + B clade and Lineage D (BS = 88) with a more basal branch leading to Lineage C (Fig. 2B).

Average $p$-distances between clades were used as a heuristic estimate of divergence time between regions. Sequence divergences among these clades (adjusted for within clade variation) were all similar and approximately half that observed between each clade and the taxa $L. boringiae$ and $L. promustache$ included in this study (Table 2). Sequence divergences within clades were about an order of magnitude lower than between clades. A second Bayesian analysis was constructed on just the cyt $b$ sequence for dating with MULTIDIVTIME. The TrN + I + G model was selected by MODELTEST using the Akaike information criterion (AIC) based on 702 bp of the cyt $b$ gene. The Bayesian analysis was run with Markov chain Monte Carlo (MCMC) ngen = $1 \times 10^7$, nchains = 4, and sumt burn-in = 25,000 to obtain one consensus tree. The Bayesian tree with $Spea$ bombifrons and $Spea$ multiplicata as outgroups is shown in Fig. 3. According to the estimations obtained with MULTIDIVTIME, the first splits within $Leptobrachium ailaonicum$ occurred during the late Miocene to early Pliocene $8.41 \pm 1.77$ Ma (lower and upper 95% credibility interval = 5.21–12.09 Ma), separating Lineage A + B + D and Lineage C (Fig. 3). The divergence time between Lineage A + B and Lineage D was calculated to have taken place in the mid Pliocene $7.97 \pm 1.72$ Ma (lower and upper 95% credibility interval = 4.89–11.51 Ma), whereas the age of the divergence time between Lineage A and

<table>
<thead>
<tr>
<th>Clade A</th>
<th>Clade B</th>
<th>Clade C</th>
<th>Clade D</th>
<th>$L. boringiae$</th>
<th>$L. promustache$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clade A</td>
<td>0.003</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clade B</td>
<td>0.043</td>
<td>0.004</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clade C</td>
<td>0.054</td>
<td>0.064</td>
<td>0.003</td>
<td>0.005</td>
<td></td>
</tr>
<tr>
<td>Clade D</td>
<td>0.050</td>
<td>0.061</td>
<td>0.050</td>
<td>0.109</td>
<td>0.135</td>
</tr>
<tr>
<td>$L. boringiae$</td>
<td>0.115</td>
<td>0.117</td>
<td>0.108</td>
<td>0.109</td>
<td>0.135</td>
</tr>
<tr>
<td>$L. promustache$</td>
<td>0.134</td>
<td>0.142</td>
<td>0.131</td>
<td>0.135</td>
<td>0.140</td>
</tr>
</tbody>
</table>

Fig. 3. Bayesian tree on 702 bp of the cyt $b$ gene, and including species of Pelobates with $Spea$ bombifrons and $S$. multiplicata as outgroups, used to date divergences between the four major clades (Lineages A–D) within $L. ailaonicum$. Bayesian posterior probabilities are given only for those nodes which obtained values > 0.95 with this method. $5.33 \pm 0.02$ Ma represents the calibration point described in the text. Inferred dates ± SE are given in rectangular boxes for the internal nodes of interest.

Table 2

Average pairwise genetic distances between the four main clades within $L. ailaonicum$ and between these clades and $L. boringiae$ and $L. promustache$. Bold values on diagonal represent average distances among haplotypes within each clade.
Lineage B was estimated at 7.35 ± 1.66 Ma (lower and upper 95% credibility interval = 4.41–10.77 Ma) (Fig. 3). Assuming an estimated rate of cyt b divergence in amphibians of 0.7–1.6%/Myr (as mentioned above), the splits between the four clades would be dated as follows: between Lineages A and B haplotypes at 2.69–6.14 Ma, Lineages A and C at 3.38–7.71 Ma, Lineages A and D at 3.13–7.14 Ma, Lineages B and C at 4.00–9.14, Lineages B and D at 3.81–8.71 Ma, and Lineages C and D at 3.13–7.14 Ma based on the net divergences. Therefore, age estimates averaged 1–2 Ma higher in the MULTIDIVTIME analysis than in using a molecular clock analyses, but 95% confidence intervals (CI) of the MULTIDIVTIME analysis normally overlapped with the range of results using the molecular clock analysis. However, these date estimates from p-distances among clades by using a molecular clock analysis should be taken only as a possible indication of the geologic periods when genetic splits occurred, because the sequence were found not to evolve in clock-like fashion.

3.3. Haplotype network estimation

Haplotypes separated by up to 19 mutational steps formed four unconnected networks at a 95% confidence level that corresponded to the four described clades observed in the MP and Bayesian trees; i.e., 4-1 (East Range) = Lineage D, 4-2 (Middle Range) = Lineage C, 4-3 (West Range) = Lineage A, and 3-7 (South Range) = Lineage B (Fig. 4). Haplotype H34 is implied to be the ancestral form of the West Range (Fig. 4). The null hypothesis of no geographic association of clades was significantly rejected (p < 0.001) by clades 3-5, 4-1, 4-2, and 4-3 in the NCA (Table 3). Using the inference keys in Templeton (1998, 2004) we found the biogeographic interpretations of both clades 4-1 and 4-2 to be consistent with past fragmentation, and clade 4-3 with allopatric fragmentation, but the biogeographical interpretation of clade 3-5 was indiscriminate (Table 3).

3.4. Population and geographic structure

Diversity indices, h and π, are summarized in Table 1. Among all individuals, h = 0.974 and π = 0.03684, indicating a high haplotype but relatively low nucleotide diversity, and that populations of this species harbored a high frequency of private haplotypes that were sampled from a single locality (Table 1). No haplotypes were shared among regional groups. Only one of 51 haplotypes (1.9%, i.e. haplotype 34) was shared by individuals from two geographically proximate localities (localities JD and NJ), the others were all restricted to the same sampling locality (98.1%). This pattern suggests little contemporary gene flow among populations and indicates that local populations of L. ailaonicum have evolved relatively independently in geographic isolation. In the AMOVA, the highest amount of genetic variance among groups (FCT = 0.848, p < 0.001) was found when the four main groups [JD, NJ, LC] [XLS] [MAD, BMH, FSL] [PB, WS] corresponding to the West Range, South Range, Middle Range, and East Range, respectively (Fig. 1) were distinguished as the most parsimonious geographic subdivisions (Table 4). In other words, most of the observed molecular variance was due to genetic differences among these four main groups indicating a relatively strong geographic structure for this species. A long-term interruption of gene flow among all lineages was also evidenced by the relatively high FST values (Table 4).

3.5. Population demographic history

Historical expansion of the main groups West Range and Middle Range is suggested by unimodal mismatch distributions (Fig. 5) and significant negative values of Fs (Table 5). We estimated an approximate time of expansion for the two groups of 386–883 Ka for the West Range and 586–1139 Ka for the Middle Range depending on an estimate of divergence rates of between 0.7% and 1.6% per million years. However, the mismatch distribution
The biogeographical interpretation is based on Templeton’s (1998, 2004) inference keys for nested clade analysis. AF, allopatric fragmentation; IBC, inconclusive outcome; RGF, restricted gene flow; IBF, isolation by distance; RGD, restricted gene dispersal; LDC, long distance colonization; LDD, long distance dispersal; CRE, continuous range expansion; PF, past fragmentation; ?, indiscriminate.

### Table 3

<table>
<thead>
<tr>
<th>Clade</th>
<th>Chi-squared statistic</th>
<th>P</th>
<th>Interior clades</th>
<th>Biogeographical Inference</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.5</td>
<td>28</td>
<td>0.001</td>
<td>2–11(T)</td>
<td>33.4809</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2–12(I)</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2–13(T)</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>I–T</td>
<td>−30.691</td>
</tr>
<tr>
<td>4.1</td>
<td>20</td>
<td></td>
<td>3–1(C)</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3–2(I)</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3–9(T)</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>I–T</td>
<td>0</td>
</tr>
<tr>
<td>4.2</td>
<td>35.486</td>
<td>0</td>
<td>3–3(I)</td>
<td>S5.4852</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0</td>
<td>3–4(T)</td>
<td>S0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0</td>
<td>3–8(T)</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0</td>
<td>I–T</td>
<td>15.4852</td>
</tr>
<tr>
<td>4.3</td>
<td>31</td>
<td></td>
<td>3–5(T)</td>
<td>S11.3469</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3–6(T)</td>
<td>S18.1307</td>
</tr>
</tbody>
</table>

### Table 4

Hierarchical analysis of AMOVA of *L. ailaonicum*.

<table>
<thead>
<tr>
<th>Groups</th>
<th>FST</th>
<th>FSC</th>
<th>FCT</th>
<th>% Among groups</th>
<th>% Within populations</th>
</tr>
</thead>
<tbody>
<tr>
<td>All populations</td>
<td>0.859</td>
<td>0.809</td>
<td>0.435</td>
<td>85.87</td>
<td>14.13</td>
</tr>
<tr>
<td>2 Groups: [JD, NJ, LC, XLS, MAD, BMH, FSL]</td>
<td>0.890**</td>
<td>0.804**</td>
<td>0.438**</td>
<td>43.80</td>
<td>11.04</td>
</tr>
<tr>
<td>2 Groups: [JD, NJ, LC, PB, WS]</td>
<td>0.880**</td>
<td>0.816**</td>
<td>0.350</td>
<td>35.01</td>
<td>11.97</td>
</tr>
<tr>
<td>3 Groups: [JD, NJ, LC, XLS]</td>
<td>0.887**</td>
<td>0.385**</td>
<td>0.317**</td>
<td>81.68</td>
<td>11.26</td>
</tr>
<tr>
<td>4 Groups: [JD, NJ, LC, XLS]</td>
<td>0.884**</td>
<td>0.620**</td>
<td>0.649**</td>
<td>69.40</td>
<td>11.63</td>
</tr>
<tr>
<td>4 Groups: [JD, NJ, LC, XLS]</td>
<td>0.882**</td>
<td>0.224**</td>
<td>0.504**</td>
<td>84.79</td>
<td>11.81</td>
</tr>
<tr>
<td>4 Groups: [JD, NJ, LC, XLS]</td>
<td>0.887**</td>
<td>0.736**</td>
<td>0.317**</td>
<td>53.31</td>
<td>12.33</td>
</tr>
<tr>
<td>5 Groups: [JD, NJ, LC, XLS]</td>
<td>0.875**</td>
<td>0.736**</td>
<td>0.540**</td>
<td>53.68</td>
<td>12.50</td>
</tr>
<tr>
<td>5 Groups: [JD, NJ, LC, XLS]</td>
<td>0.873**</td>
<td>0.205**</td>
<td>0.840**</td>
<td>84.03</td>
<td>12.69</td>
</tr>
<tr>
<td>5 Groups: [JD, NJ, LC, XLS]</td>
<td>0.879**</td>
<td>0.238**</td>
<td>0.842**</td>
<td>84.22</td>
<td>12.02</td>
</tr>
<tr>
<td>5 Groups: [JD, NJ, LC, XLS]</td>
<td>0.871**</td>
<td>0.305**</td>
<td>0.815**</td>
<td>81.48</td>
<td>12.88</td>
</tr>
<tr>
<td>5 Groups: [JD, NJ, LC, XLS]</td>
<td>0.871**</td>
<td>0.488**</td>
<td>0.747**</td>
<td>74.75</td>
<td>12.93</td>
</tr>
<tr>
<td>6 Groups: [JD, NJ, LC, XLS]</td>
<td>0.871**</td>
<td>0.224**</td>
<td>0.834**</td>
<td>83.37</td>
<td>12.90</td>
</tr>
</tbody>
</table>

* P < 0.05. ** P < 0.01. *** P < 0.001.

4. Discussion

4.1. Pre-Pleistocene split and geologic history

Phylogenetic analyses using both Bayesian and parsimony methods recovered four distinct mtDNA clades across the populations of *L. ailaonicum* (Table 1; Fig. 2). Although the relationships among the four clades varied according to the tree-building method used, the general pattern that emerged from both analyses is that the deep genetic structure recovered within *L. ailaonicum* coincides with the major river valley systems in southwest China (Figs. 1 and 2). This pattern is strengthened by a lack of an unambiguous link between the four haplotype networks (Fig. 4) and high $F_{ST}$ values (Table 4), all suggesting that the clades diverged in allopatry from a common ancestor, that these divergences were deep, and related to specific geologic events within the region.

Mitochondrial DNA haplotypes in *L. ailaonicum* displayed range-restrictions of the main clades, as many of the haplotypes within a population clustered together. Such distributions for mitochondrial DNA haplotypes may be interpreted as being the result of population isolation due to specific habitat requirements, specifically breeding site requirements of this species. *Leptobrachium ailaonicum* occurs primarily near and breeds in mid-elevation mountain streams with clear, cold, slow-flowing water. These specific habitat requirements, which definitely limit its rate of dispersal and choice of migration routes, probably explain its absence from lowland areas and the structuring observed between the current populations in the complex land conditions of this region. If the lowland river valleys (i.e., the Red River valley and Tengtiaojiang River valley) are then considered natural barriers to dispersal, a past vicariance hypothesis is consistent with the geological history and topography within the range of *L. ailaonicum*. Using a Bayesian dating method, we estimate the divergence times between the major clades to be 4.41–12.09 Ma, suggesting that these clades separated during the late Miocene to mid-Pliocene. This divergence time is roughly congruent with the latest and most significant uplift of Mount Ailao, which occurred approximately 5 Ma (Wang et al.,
2006), and thus considered the cause of allopatric isolation and lineage split. The same pattern has been observed for several species of plants (Assefa et al., 2007), arthropods (Masta, 2000; Ponniah, 2002; DeChaine and Martin, 2004, 2005; Smith and Farrell, 2005; Finn et al., 2006; Mark and Jane, 2006), amphibians (Doak, 2005; Roberts et al., 2007), and mammals (Yuan et al., 2006), to name a few, indicating that lowland areas are barriers to contemporary dispersal and were effective barriers to ancestral gene flow.

4.2. Population demographic history and secondary contact in Quaternary

It should be noted that high haplotype diversity, deep phylogenetic trees, negative Fu’s $F_s$ values, and multimodal distribution shapes showed in mismatch distribution, all indicate a complicated demographic history of *L. ailaonicum* populations. Population expansions at approximately 386–1139 Ka were detected in the West Range and Middle Range clades. We can attribute this to climatic oscillations during glacial periods in the Quaternary, allowing some populations to expand while others remained stable.

Two divergent lineages coexist in two localities; Lineage A and B at XLS, and Lineage A and C at FSL (Fig. 1), implying secondary contact after initial divergence. Phylogenetic branching patterns have been used in studies on amphibians and reptiles with limited dispersal capabilities to suggest dispersal direction and location of origin of a population (Carranza et al., 2000; Fu et al., 2005; Huang et al., 2007). In the present study, a general trend of *L. ailaonicum* populations in Lineage A from the West Range dispersing to the two localities XLS and FSL in the South Range and Middle Range, respectively can be suggested by the phylogenetic tree, as some individuals from Lineage A were collected in these localities. We

Table 5
Mismatch distribution analysis. The parameters of the model of sudden expansion (Rogers and Harpending, 1992) are presented. *p*-values for rejection of the sudden expansion model are based on a comparison of the sum of squares of expected and observed distribution, using parametric bootstrapping with 10,000 replicates (Excoffier et al., 2005). Tajima’s $D$ and Fu’s $F_s$ and their statistical significance are shown.

<table>
<thead>
<tr>
<th>Lineage</th>
<th>$\tau$ (KY)</th>
<th>$T$ (KY)</th>
<th>Fu’s $F_s$</th>
<th><em>P</em>-value</th>
<th>Tajima’s $D$</th>
<th><em>P</em>-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>West Range</td>
<td>4.097</td>
<td>386.218–882.784</td>
<td>−9.648</td>
<td>&lt;0.001</td>
<td>−1.561</td>
<td>0.05</td>
</tr>
<tr>
<td>Middle Range</td>
<td>6.214</td>
<td>585.784–1338.936</td>
<td>−9.09</td>
<td>0.001</td>
<td>−1.998</td>
<td>0.136</td>
</tr>
<tr>
<td>East Range</td>
<td>8.963</td>
<td>1.085</td>
<td>0.707</td>
<td>0.529</td>
<td>0.76</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>73.725</td>
<td>1.417</td>
<td>0.75</td>
<td>1.561</td>
<td>0.957</td>
<td></td>
</tr>
</tbody>
</table>

Fig. 5. Observed and expected mismatch distributions showing the frequencies of pairwise differences: (a) within West Range; (b) within Middle Range; (c) within East Range; and (d) considering all the samples.
postulate that these dispersal events occurred in the latest Pleistocene glacial period when lower temperatures would have accommodated southward dispersal and range expansion (Fig. 1). In the following interglacial to the present, populations of L. ailaonicum would again have been isolated in the fragmented mountain habitats. The NCA inference within this region (Table 3) is indistinguishable, but the signals of population growth derived from Fst test and patterns of the phylogenetic trees support this interpretation. This pattern has been observed in other montane organisms that were hypothesized to have been originally isolated in sky island refugia during warm interglacials, such as the present, and expanded their ranges by dispersing down-slope and across the historical lowland barriers during the glacial periods, resulting in secondary contact between previously allopatric populations (DeChaine and Martin, 2004; Yuan et al., 2006).

Therefore, based on our above results, we hypothesize that the genetic history of L. ailaonicum included (i) initial divergence between four clades at 4.41–12.09 Ma, which would have been caused by the Late Cenozoic uplift of Mt. Aila and (ii) secondary contact occurring after the initial divergence, caused by the expansion of Lineage A during Pleistocene glacial cooling.

4.3. Taxonomic implications

There is a long-standing dispute with respect to the recognition of the species L. echinatum (Dubois and Ohler, 1998; Ho et al., 1999; Ohler et al., 2000; Grosjean, 2001; Rao and Wilkinson, 2008; Zheng et al., 2008). Dubois and Ohler (1998) distinguished specimens collected from Vietnam as L. echinatum based on a relatively higher number of maxillary spines and a green iris color as opposed to a lower number of spines and blue iris color in L. ailaonicum. Ho et al. (1999) disputed this new species recognition, considering a higher number of maxillary spines in Vietnamese specimens as the end of a range. They did not discuss iris color differences. Ohler et al. (2000) defended the recognition of L. echinatum, stating that Ho et al. (1999) did not observe the iris color of live specimens but merely assumed that they were blue and that there is no overlap in range of maxillary spines between L. ailaonicum from China and the specimens from Vietnam. Grosjean (2001) further supported L. echinatum as distinct from L. ailaonicum based on larval characters. Rao and Wilkinson (2008) concluded L. ailaonicum and L. echinatum as two separate but closely related species in their phylogenetic analysis, noting that the specimens of L. echinatum in their study could be recognized as such by a higher number of spines and green iris color. However, Zheng et al. (2008) synonymized L. echinatum into L. ailaonicum because their phylogeny showed a L. ailaonicum + L. echinatum clade in which some individuals of L. ailaonicum (from the same locality as JD in our present paper, Fig. 1) formed a subclade with individuals of L. echinatum from Fan Si Pan, Vietnam (type locality of L. echinatum).

We have demonstrated that at least four deeply divided clades exist within L. ailaonicum. We have also shown that Lineage A has expanded its range southward to the Southern and Middle ranges that border Vietnam. We propose that Zheng et al. (2008) may have sampled individuals of Lineage A thinking them to be L. echinatum, even though they stated that these individuals resemble the description of L. echinatum. We suggest this because the samples in their study were tissues borrowed from a museum of individuals that came from the type locality in northern Vietnam, which is close to our Middle and Southern ranges where Lineages A and B and A and C, respectively, occur in sympatry. This is plausible since, iris color, the main character defended by Ohler et al. (2000) to differentiate these two species, cannot be determined from museum specimens, and the number of labial spines, the other only adult character given by Dubois and Ohler (1998) to distinguish these species, was not mentioned by Zheng et al. (2008).

It is interesting that the topologies in both Rao and Wilkinson (2008) and Zheng et al. (2008) are similar in that individuals from east of the Red River (our Lineage D) are clearly separate (with high bootstrap and posterior probability support) from other individuals of L. ailaonicum. Based on our results, we can conclude that the question of whether L. ailaonicum in actuality constitutes just one or more than one species is still unresolved. However, the degree of genetic divergence between the four clades suggests that more than one species may be present under this name. We therefore suggest that a more thorough sampling and examination of this species, including a better representative sampling from Vietnam, in which voucher specimens have been clearly identified (number of labial spines, color of iris, etc.), needs to be performed to answer this question. This examination should include population genetic as well as phylogenetic analyses on mitochondrial markers alone or in combination with nuclear markers as has been done in this and other studies (Shaffer et al., 2004; Gamble et al., 2008).

4.4. Conservation and management implications

Two goals of any conservation program should be to ensure the survival of a species and maintain its genetic diversity for long-term evolutionary success (Hamrick and Godt, 1996). Our study has provided a means for assessing the evolutionary distinctiveness of populations of L. ailaonicum that may need conservation. The data can be used to establish management units (MUs) and/or evolutionary significant units (ESUs), two commonly used designations for threatened or endangered taxa (Moritz, 1994; Bidlack and Cook, 2001). Management units are defined by either reciprocal monophyly in mtDNA or substantial allele frequency divergence at nuclear loci; ESUs are defined by the presence of both (Moritz, 1994). Considering these criteria, populations with genotypes that are closely related to but not shared with other populations would be described as MUs; thus, the East Range, South Range, Middle Range and West Range populations of L. ailaonicum would be considered at least MUs. These three regional populations may represent important components in the evolutionary and adaptive structure of the species, and thus any conservation policy should concentrate on protecting these distinct populations. Similar management units have been proposed to conserve the Tibetan gazelle in China (Zhang and Jiang, 2006). Although the status of protection in China is only to protect endangered species and their habitats and not to enforce the MU or ESU concept, we recommend that associated Chinese law adopt the international standard (MU and ESU) to better conserve endangered species. As a conservative approach, we would recommend enhancing efforts to protect all populations of L. ailaonicum in China and avoid the introduction of genes to any of the populations unless genetic erosion occurs.

In addition, further genetic studies using both mtDNA and nuclear markers (e.g., microsatellite loci) with greater sampling throughout the distribution of L. ailaonicum is needed, and nonmolecular information such as morphology, ecology, behavior, etc., should be used to guide the design of these studies.

Acknowledgments

We thank J. Che, X.Y Chen, Y. Li, H.F. Tian, D.D. Xin, W. Wan, C.H. Wu, K. He, and X.B. Qi for database and statistical assistance. We also thank M. Koo for providing Fig. 1, and D. Neely for critically reviewing earlier portions of the manuscript. This research was supported by grants from the Knowledge Innovation Project of the Chinese Academy of Sciences (KSCX2-SW-119), Chinese National Natural Science Foundation (No. 30270175), National Basic Research Program of China (973 Program) (No. 2007CB411600), Key Laboratory of Cell and Molecular Evolution at Kunming...